

November 2014

“The impact of perceptions in averting-decision models:
An application of the special regressor method to drinking
water choices”

Christophe Bontemps & Céline Nauges

**The impact of perceptions in averting-decision models:
An application of the special regressor method to drinking water choices¹**

Christophe Bontemps & Céline Nauges

Abstract: Individuals are commonly surveyed about their perception or assessment of risk and these variables are often used to explain individuals' actions to protect themselves against these risks. Perceptions appear as endogenous variables in traditional theoretical averting-decision models but, quite surprisingly, endogeneity of perceived risk is not always controlled for in empirical studies. In this article, we present different models that can be useful to the practitioner when estimating binary averting-decision models featuring an endogenous discrete variable (such as risk perception). In particular we compare the traditional bivariate probit model with the special regressor model, which is less well known and relies on a different set of assumptions. In the empirical illustration using household data from Australia, Canada, and France, we study how the perceived health impacts of tap water affect a household's decision to drink water from the tap. Individuals' perceptions are found to be endogenous and significant for all models, but the estimated marginal effect is sensitive to the model and underlying assumptions. The special regressor appears to be a valuable alternative to the more common bivariate probit model.

1. Introduction

Risk perceptions are commonly recognised as important drivers of the behaviour of individuals including their choice of averting actions taken to prevent these risks. In typical theoretical averting-behaviour models, both actual risk and perceived risk are assumed to determine the individual's level of risk exposure (Courant and Porter, 1981; Dickie and Gerking, 1996). Perceived risk is endogenous since it depends not only on actual risks but also on attitudes towards safety, hence on the averting actions being taken. In empirical studies however, the potential endogeneity of perceived risk in models describing households' actions to mitigate this risk is not always accounted for, which might induce estimation bias (Whitehead, 2006). In this paper, we acknowledge that perceptions are endogenous to averting decisions and we present and compare those models that are available to the practitioner to handle this endogeneity.

¹ The data used in this work come from an OECD survey on Environmental Policy and Individual Behaviour Change (EPIC) which is periodically conducted by the Environment Directorate. The views expressed do not necessarily reflect those of the Organisation for Economic Cooperation and Development (OECD) or its member countries.

In surveys, perceived risk commonly takes the form of a categorical variable, either binary (e.g., whether some risk is perceived or not; see e.g. Abrahams et al., 2000; Um et al., 2002; Konishi and Adachi, 2011) or ordered when perceptions are assessed on a scale (e.g., Um et al., 2002; Whitehead, 2006; Janmaat, 2007; Johnstone and Serret, 2012). A number of averting decisions are also recorded as binary variables. In averting-behaviour models related to drinking water consumption for example, buying bottled water and using water filtering systems are the most frequently studied decisions (Abrahams et al., 2000; Um et al., 2002; Janmaat, 2007; Johnstone and Serret, 2012). In this article we focus on cases where perceptions are recorded as discrete variables and the averting decision is binary.

In such settings, two-step estimation procedures such as two-stage least squares or control functions methods do not produce consistent estimates (Wooldridge, 2010) and the traditional approach is the estimation of a bivariate probit model, where the averting decision and the variable measuring perceived risk are the dependent variables of two probit models. Estimation of a bivariate probit model requires the specification of functional forms for the two probit models, as well as an assumption of joint normality of the error terms of the two equations. Another approach, which (as far as we know) has not yet been used in averting-type behaviour models, is the special regressor method (Lewbel, 2014).² This methodology, which is based on earlier work by the same author (Lewbel, 2000), is relatively easy to implement and relies on assumptions that are quite different from those required for maximum likelihood estimators such as the bivariate probit. The special regressor approach does not impose restrictions on the model errors and does not require the relationship between the endogenous and exogenous regressors to be specified, but its consistency relies on the presence of one exogenous regressor that is conditionally independent of the model error, is continuously distributed, and has a large support. The special regressor thus provides a simple way of testing the validity of the modelling assumptions underlying more standard maximum likelihood approaches. In this article, we use these two estimators to measure the influence of the perceived health impacts of tap water on a household's decision to drink water from the tap, using household data collected in Australia, Canada, and France.

The endogeneity of variables measuring perceptions was acknowledged by Whitehead (2006) but we found few empirical studies of water-related averting behaviour in which the endogeneity of perceptions had been explicitly taken into account. Konishi and Adachi (2011) estimated willingness-to-pay (WTP) for a reduction in the permissible level of arsenic concentration in drinking water. Different levels of payments were proposed and a household's response was coded as one if it was willing to pay more than the bid value, and zero otherwise. In the article, the endogeneity of averting actions was taken into account through the estimation of a bivariate probit model featuring the WTP equation and an equation describing the household's decision to take averting actions (here using arsenic-

² The special regressor method has been applied in other domains. See, for example, Maurin (2002) and Goux and Maurin (2005) on a child's performance at school.

related treatment or purchasing bottled water). In Abrahams et al. (2000), averting decisions to buy bottled water or to use a filtering system were assumed to depend on perceived risk, itself specified as a function of attitudes towards water safety and knowledge of water safety issues. The authors acknowledged the endogeneity of perceived risk but, in their empirical application, removed the endogeneity problem by assuming that perceived risk and actual risk were the same (they indicated that the lack of data on actual risk precluded the use of a two-step regression to control for the endogeneity of perceived risk).

This paper is organised as follows. In Section 2, we derive the theoretical model describing a household's drinking water choices. In Section 3, we describe the special regressor method as well as other approaches commonly used to estimate binary choice models with a binary endogenous regressor. We present our data and empirical model in Section 4. In Section 5, we compare estimation results from different models and run some sensitivity analyses on the special regressor estimator. Section 6 concludes.

2. The model of a household's drinking water choices

Following traditional averting-behaviour models, we assume that a representative household will choose to drink tap water if its expected utility when drinking water from the tap (U_1) is higher than its expected utility when not drinking water from the tap (U_0):

$$U^* = U_1 - U_0 > 0 \text{ where } U^* \text{ is a latent variable.}$$

Along the lines of Abrahams et al. (2000), we assume that water enters a household's utility directly (as with any other consumption goods) as well as indirectly through a health production function. The latter is directly influenced by the health properties of the type of water the household is drinking, as well as by the household's risk exposure related to water consumption. Since households may not always be able to assess risks properly, perceived risks are usually used in place of actual risks. Perceived risks could be influenced by a number of factors including a household's knowledge and information about water quality and its attitudes towards water safety (including its choices of drinking water). Under such assumptions, perceived risks are endogenous to the decision about drinking water sources.

Assuming that the utility function U is linear in its parameters, β_j ($j=0, 1$), and can be written as the sum of a deterministic term ($\mathbf{X}'\beta_j$) and an error term of mean zero (ε_j), we have:

$$U^* = U_1 - U_0 = \mathbf{X}'(\beta_1 - \beta_0) + \varepsilon_1 - \varepsilon_0 \quad (1)$$

where \mathbf{X} is the vector of explanatory variables. The latent variable U^* is not observed by the econometrician; only the decision (D) to drink tap water (or not) is observed. Hence the model to be estimated takes the form:

$$D = \mathbf{I}(\mathbf{X}'\beta + \varepsilon \geq 0). \quad (2)$$

In the following empirical application, the vector \mathbf{X} includes a household's perception of the health impacts of tap water (among other socio-demographic characteristics of the household). Perceptions are measured using a binary variable which, as explained above, is potentially endogenous in the household's decision model. In the next section, we present various estimators that can be used to identify the set of unknown parameters $\boldsymbol{\beta}$ in binary decision models featuring a binary endogenous regressor.

3. Estimation of a binary decision model with binary endogenous regressors

We present in turn the special regressor model and the standard maximum likelihood approach (bivariate probit model). We also discuss the standard (but inconsistent in this case) control function approach.

3.1 The special regressor method

The special regressor approach proposed by Lewbel (2000) was designed in a broad context to improve the identification and estimation of parameters and their associated distributions in general threshold-crossing models such as the following binary decision model:

$$D = \mathbf{I}(\mathbf{X}'\boldsymbol{\beta} + \varepsilon \geq 0) \quad (3)$$

where D represents the binary decision variable, the vector \mathbf{X} includes all observable regressors, ε has a zero mean distribution, and $\mathbf{I}(\cdot)$ is the indicator function taking the value one if the latent variable $\mathbf{X}'\boldsymbol{\beta} + \varepsilon$ is positive and zero otherwise. Traditional probit and logit models correspond to the case of ε following a normal and exponential distribution, respectively.

The special regressor model has the same form as Model (3) but is rewritten so that one regressor V (that is called the special regressor) is separated from the other regressors \mathbf{X} and its coefficient is normalised to one:

$$D = \mathbf{I}(\mathbf{X}'\boldsymbol{\beta} + V + \varepsilon \geq 0). \quad (4)$$

The special regressor V must satisfy three fundamental conditions:

- V is continuously distributed and has a large support (i.e., V varies on a support that is as large as the support of $\mathbf{X}'\boldsymbol{\beta} + \varepsilon$);
- V is exogenous;
- $E(D | \mathbf{X}, V)$ increases with V .

If some elements of \mathbf{X} are endogenous, then the special regressor method requires instruments \mathbf{Z} satisfying the following usual properties: $E(\mathbf{Z}'\varepsilon) = 0$ and $E(\mathbf{Z}'\mathbf{X})$ has full rank. The special regressor V should not be included in the set of instruments \mathbf{Z} which implies that a suitable V should only affect the binary decision of interest, and not the endogenous (binary) variable.³ The basic idea behind the special regressor approach is described below (we refer readers to Appendix A1 and to Lewbel (2014) for additional details).

Define T as:

$$T = \frac{D - I(V \geq 0)}{f_{V|\mathbf{Z}}(V|\mathbf{Z})} \quad (5)$$

where $f_{V|\mathbf{Z}}(V|\mathbf{Z})$ denotes the conditional probability density function of V given \mathbf{Z} . Under the above assumption, it can be shown that $E(T|\mathbf{Z}) = E(\mathbf{X}'\boldsymbol{\beta} + \varepsilon|\mathbf{Z})$, assuming ε is independent of $V|\mathbf{Z}$ (Lewbel, 2014). It follows from the latter equality that $E(\mathbf{Z}\mathbf{X}')\boldsymbol{\beta} = E(\mathbf{Z}T)$, and so:

$$\boldsymbol{\beta} = \left[E(\mathbf{X}\mathbf{Z}')E(\mathbf{Z}\mathbf{Z}')^{-1}E(\mathbf{Z}\mathbf{X}') \right]^{-1} E(\mathbf{X}\mathbf{Z}')E(\mathbf{Z}\mathbf{Z}')^{-1}E(\mathbf{Z}T) \quad (6)$$

which is the definition of a linear two-stage least squares regression of T on \mathbf{X} using instruments \mathbf{Z} . We describe below the steps to be followed to estimate T and the parameters of interest $\boldsymbol{\beta}$.

Step 0: V must be of mean zero; if not, one must first de-mean it.

Step 1: Run an Ordinary Least Squares (OLS) regression of V on (\mathbf{X}, \mathbf{Z}) and, for each observation i , compute the residuals as the difference between the observed V_i and its prediction: $\hat{U}_i = V_i - (\mathbf{X}'\hat{\mathbf{b}}_{\mathbf{X}} + \mathbf{Z}'\hat{\mathbf{b}}_{\mathbf{Z}})$, where $\hat{\mathbf{b}}_{\mathbf{X}}$ and $\hat{\mathbf{b}}_{\mathbf{Z}}$ are the OLS estimated coefficients for variables in \mathbf{X} and \mathbf{Z} , respectively.^{4,5}

³ Note that only one special regressor is needed whatever the number of endogenous variables in the model.

⁴ As noted by Dong and Lewbel (forthcoming), the relationship between V and (\mathbf{X}, \mathbf{Z}) could be made more flexible and estimated by semi-parametric or non-parametric methods, since the only output of interest is the vector of estimated residuals $(\hat{U}_i)_{i=1, \dots, n}$.

⁵ In the presence of heteroskedasticity in the residuals \hat{U}_i , the classical correction based on the regression of the squared residuals on the variables causing heteroskedasticity can be applied (Dong and Lewbel, forthcoming).

Step 2: Compute \hat{f}_h as the non-parametric kernel estimator of the density f of \hat{U}_i , as:

$$\hat{f}_h(u) = \frac{1}{nh} \sum_{j=1}^n K\left(\frac{\hat{U}_j - u}{h}\right) \quad (7)$$

and for each i compute the estimates $\hat{f}_i = \hat{f}(\hat{U}_i)$. Note that this estimation requires a kernel $K(\cdot)$ and a bandwidth h to be chosen. Alternatively, one may use the ordered data estimator proposed by Lewbel and Schennach (2007). This is based on an ascending pre-ordering of \hat{U}_i , from the smallest to the highest value. Then for each i define \hat{U}_i^+ as the nearest neighbour within $(\hat{U}_j)_{j \neq i}$ that is greater than \hat{U}_i . Likewise, define \hat{U}_i^- as the nearest neighbour that is smaller than \hat{U}_i . One may then replace \hat{f}_i in the next step by \hat{f}_i^0 defined as:

$$\hat{f}_i^0 = \frac{2/n}{\hat{U}_i^+ - \hat{U}_i^-} \quad (8)$$

Note that \hat{f}_i^0 is not a consistent estimator of f_i but has the desired advantage of not being dependent of any bandwidth choice and is trivial to compute (see also Dong and Lewbel, forthcoming, for details).

Step 3: For each observation i construct \hat{T}_i as:⁶

$$\hat{T}_i = \frac{D_i - \mathbf{I}(V_i \geq 0)}{\hat{f}_i}. \quad (9)$$

Whatever the estimator used in the construction of \hat{T}_i (\hat{f}_i or \hat{f}_i^0), the denominator can take very small values especially for large absolute values of \hat{U}_i . As a consequence, some \hat{T}_i could have extremely large absolute values, which could induce large standard errors in the final two-stage least squares regression. Lewbel (2014) recommends either removing these extreme values using a trimming procedure or using an appropriate statistical procedure such as Winsorisation (Tukey, 1962).⁷ In the case of a non-parametric estimation of \hat{f}_i , the

⁶ In the presence of heteroskedasticity, the numerator in Equation (9) has to be multiplied by $\sqrt{\hat{U}_i^2}$, see Dong and Lewbel (forthcoming).

⁷ When the data are Winsorised, tail values are set equal to some specified percentile of the data. For example, for a 90 per cent Winsorisation, the bottom 5 per cent of the values are set equal to the value corresponding to the 5th percentile while the upper 5 per cent of the values are set equal to the value corresponding to the 95th percentile. For a 90 per cent trimming, values below the 5th percentile and values above the 95th percentile are removed.

choice of the bandwidth h may also affect the magnitude of \hat{T}_i through the estimated density \hat{f}_i .

Step 4: Compute $\hat{\beta}$ as the coefficient of a two-stage least squares regression of \hat{T} on \mathbf{X} using instruments \mathbf{Z} .

The estimation of the special regressor thus involves four steps that can be implemented relatively easily in most statistical software (a procedure is now available in STATA, see Baum, 2012). Interestingly, classical tests of instruments' validity can be applied at the final two-stage least squares regression knowing that $E(\mathbf{Z}\varepsilon) = E(\mathbf{Z}\tilde{\varepsilon})$ with $\tilde{\varepsilon} = T - \mathbf{X}'\beta$, the error term of the Step 4 regression of \hat{T} on \mathbf{X} using instruments \mathbf{Z} (Dong and Lewbel, forthcoming).

3.2 The traditional maximum likelihood approach (bivariate probit)

A bivariate probit model corresponds to the joint estimation of two probit models. In our particular setting, Equation (3), which is the main model of interest, would be estimated jointly with a model that specifies the relationship between the (binary) endogenous variable and a set of exogenous regressors. For ease of notation let us assume $\mathbf{X} = (\mathbf{X}_0, X_e)$ where \mathbf{X}_0 is the vector of exogenous regressors and X_e is one binary endogenous regressor. The second equation to be estimated is written:

$$D_e = I(\mathbf{X}_0'\gamma + \mathbf{Z}'\theta + \omega \geq 0) \quad (10)$$

where D_e represents the binary (endogenous) decision variable, γ and θ are the set of unknown parameters corresponding to the vector of exogenous variables \mathbf{X}_0 and the vector of instruments \mathbf{Z} , respectively; and ω is a random term. A further assumption that is required for Equations (3) and (10) to be estimated jointly by maximum likelihood, is that the error terms of the two equations (ε and ω) follow a joint normal distribution. Under the assumption that the relationship in (10) is correctly specified and the assumption on the joint normality of the error terms, maximum likelihood provides consistent estimates of the parameters of interest β in Equation (3).

3.3 Control function: an inconsistent approach

As in the case of the bivariate probit, the control function approach, originally proposed by Heckman (1976), requires the estimation of the relationship between the endogenous

regressor X_e and $(\mathbf{X}_o, \mathbf{Z})$. Contrary to the maximum likelihood estimator described in Section 3.2, the relationship can be parametrically or non-parametrically specified. In addition, control function methods usually do not require the joint normality assumption necessary in the bivariate probit model, which makes control functions somehow less constrained (or more general) than maximum likelihood estimators. However, the control function methodology necessitates solving for the vector of errors ω in the relationship between X_e and $(\mathbf{X}_o, \mathbf{Z})$, something which is not do-able if the function that describes the relationship is a probit (or any sort of threshold-crossing model). Hence the common control function methodology is not appropriate when the model of interest is non-linear (such as binary choice models) and the endogenous variable is discrete (Wooldridge, 2010).⁸ Assuming that the relationship between the endogenous regressor X_e and $(\mathbf{X}_o, \mathbf{Z})$ is linear or non-parametric would allow the residuals to be determined but would not be valid because the assumption of independence between the residuals and the regressors $(\mathbf{X}_o, \mathbf{Z})$ is violated when the dependent variable is binary. This is because, in a linear model of the form $X_e = \mathbf{X}'_o \gamma + \mathbf{Z}' \theta + \omega$, the error term can take only two values ($1 - \mathbf{X}'_o \gamma - \mathbf{Z}' \theta$ or $0 - \mathbf{X}'_o \gamma - \mathbf{Z}' \theta$) which are clearly not independent of $(\mathbf{X}_o, \mathbf{Z})$. For more details on the comparison of maximum likelihood and control function estimators, see Lewbel et al. (2012). In the empirical application, we estimate the model describing drinking water choices using both the special regressor and the bivariate probit estimators. For comparison purposes, we also include estimates obtained from a simple probit and a control function approach.

4. Description of the data and model specification

The data come from an OECD survey on Environmental Policy and Individual Behaviour Change (EPIC) which is periodically conducted by the OECD Environment Directorate.⁹ We use household cross-sectional data for three countries (Australia, Canada, and France) and two years: 2008 and 2011. Around 1,000 (different) households were surveyed each year in each country.¹⁰ In order to ensure representativity, the sample was stratified in each country according to age, gender, region, and socio-economic group. The two rounds were implemented using Internet-based surveys.¹¹ Households were surveyed about their

⁸ Control functions can be applied when the model of interest is linear in parameters and the endogenous variable is discrete: the relationship between the endogenous variable and the set of instruments is estimated with a probit model in the first stage and, in the second stage, the fitted probability is used as an instrumental variable for the endogenous regressor in the model of interest.

⁹ For more details on the EPIC surveys and related publications, see OECD (2011) and <http://www.oecd.org/env/consumption-innovation/households.htm> (accessed, 21 October 2014).

¹⁰ The EPIC survey covered 10 countries in 2008 (Australia, Canada, Czech Republic, France, Italy, Korea, Mexico, the Netherlands, Norway, and Sweden) and 11 countries in 2011 (Australia, Canada, Chile, France, Israel, Korea, Japan, the Netherlands, Spain, Sweden and Switzerland).

¹¹ For more details on the survey implementation, we refer readers to Annex B in OECD (2013).

behaviour in several domains including energy and water consumption, waste management, and transportation. In the following, we study a household's decision to drink water from the tap using a threshold-crossing model of the form:

$$D = I(\mathbf{X}'\boldsymbol{\beta} + \varepsilon \geq 0). \quad (11)$$

In the empirical application, the variable D represents a household's decision to drink water straight from the tap (i.e., it does not include tap water that has been purified, filtered or boiled). As explanatory variables we use the household's socio-demographic characteristics including: the age of the respondent (in years); the household's annual income (in thousands of Euro (EUR)); a dummy variable that takes the value one if at least one household member is below 18 years of age; a dummy variable that takes the value one if the household owns (or uses regularly) one or more cars; a dummy variable that takes the value one if the household lives in a town or suburban area; and dummies to control for the country of residence. The respondent's age and the composition of the household are commonly used to control for drinking habits that may vary across different age groups. Living in a town or suburban area, as well as car usage, is used to control for easier access to bottled water suppliers such as local supermarkets (Johnstone and Serret, 2012). Income is used here as a proxy for the household's level of knowledge and information about water quality. Due to collinearity issues, variables measuring the household's educational attainment could not be included in the model. We control for the respondent's level of satisfaction with respect to water quality in his/her local environment (quality of lakes, rivers, etc.) using an index that can take five different values: very dissatisfied (-2); dissatisfied (-1); no opinion (0); satisfied (1); and very satisfied (2). Using household data from France, Bontemps and Nauges (2009) showed that households which held a higher opinion of the quality of their local environment were more likely to drink water from the tap. We use the (average) regional water price at the tap in 2008 and expect tap water to be a normal good.¹² Finally, the main variable of interest is the respondent's perceived health impacts of drinking water from the tap. We expect a higher level of satisfaction to increase the probability of drinking water straight from the tap. Respondents had to indicate their level of satisfaction (about health impacts) on a scale from zero to 10. Instead of considering the 0-10 scale, we make a simple distinction between respondents who stated being satisfied (six or above) and those who were not satisfied with the tap water health impacts (five or less). We believe that respondents can easily state whether they are satisfied or not, but that it is much more difficult for them to actually specify their level of satisfaction on a scale from zero to 10. The variable measuring a respondent's satisfaction thus takes the value one if the respondent perceives the health impacts of tap water as satisfactory, and zero otherwise. The latter is a potential endogenous variable and instruments are required.

The model of interest (Equation 11) is estimated using data collected in 2011 but we use data from the 2008 survey to build two instruments for the endogenous variable. Different

¹² The price of water at the tap was not collected during the 2011 survey round.

households were surveyed in the two rounds but, for each household, we have information on its country and region of residence.¹³ We use (average) regional variables from the 2008 survey as instruments for the household’s perceived health impacts of tap water in 2011. The first instrument is the proportion of households who drank water from the tap (including tap water boiled, purified or filtered) in 2008 in the region.¹⁴ The second instrument is the average level of concern about water pollution (in general) in 2008 in the region. The level of concern was measured on a scale from zero to 4: 0 (no opinion); 1 (not concerned); 2 (fairly concerned); 3 (concerned); and 4 (very concerned).

In order to estimate the special regressor model, a special regressor V satisfying the assumptions defined in Section 3.1 is needed. We choose the average price of tap water in the region in 2008 since it is continuously distributed over a large support (varying from 0.69 to 4.22 EUR per m^3), it can be assumed to be exogenous to a household’s decision to drink tap water in 2011, and consumer theory suggests that it should have a monotonic (decreasing) relationship with tap water consumption. The special regressor should be such that $E(D|\mathbf{X},V)$ increases with V . Since the probability of drinking water from the tap decreases with an increase in the price of tap water, we define V as minus the price and we normalise V such that it is of mean zero. A fitted kernel-weighted local polynomial regression of D on V provides empirical evidence for a monotonic (increasing) relationship between the two variables (Figure 1).

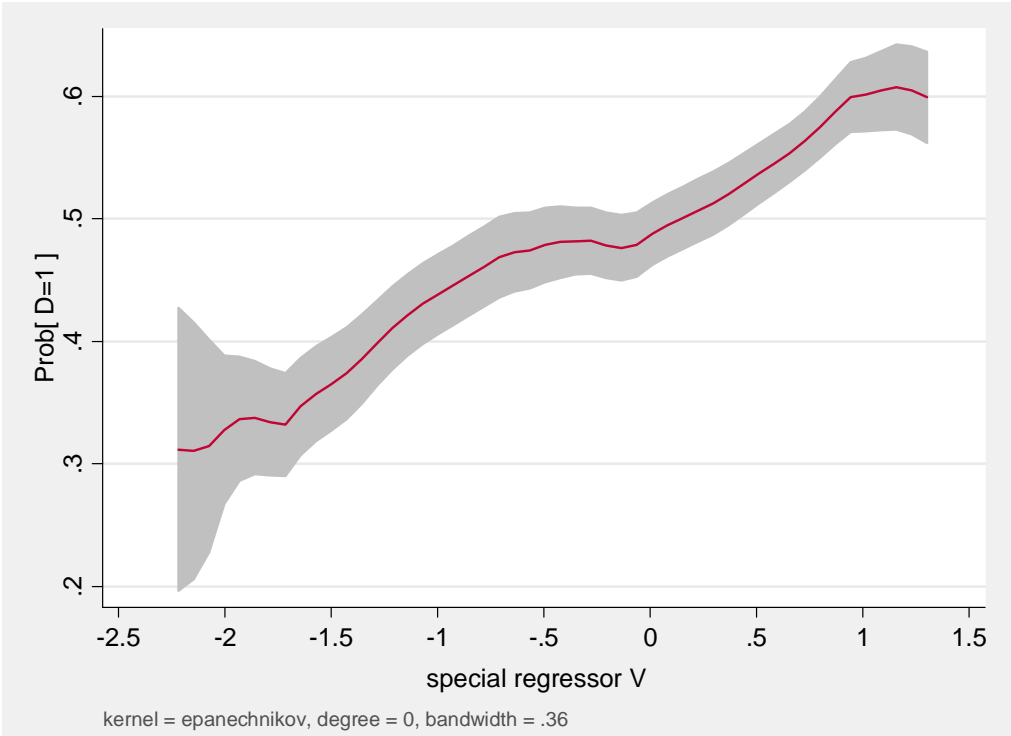


Figure 1: Kernel-weighted local polynomial regression of D on V .

¹³ There were eight regions in Australia, 10 in Canada, and 22 in France.

¹⁴ In the 2008 survey the respondents could not specify whether tap water was drunk straight from the tap, or filtered, or boiled, or purified.

Descriptive statistics (means) of the variables used in the empirical application are shown in Table 1. For the sample as a whole (2,771 observations), 50 per cent of the surveyed households stated drinking water straight from the tap in 2011, with differences across the three countries (59 per cent in Australia, 49 per cent in Canada, and 43 per cent in France). Overall 70 percent of the respondents indicated being satisfied with the health impacts of tap water. The average satisfaction level was lower in France (61 per cent), than in Australia (78 per cent) and Canada (74 per cent). Respondents' opinions about the water quality (of lakes, rivers etc.) in their local environment is quite different in the three countries: the perceived water quality was lower in France (the average index is equal to 0.35) than in Canada (0.45) and Australia (0.67). Finally, the average (tap) water price (in 2008) was higher in France (2.99 EUR per m3 on average) than in Canada (1.54 EUR) and Australia (1.37 EUR).

Table 1. Sample and country-specific averages

Variable	Survey year	Full sample	Australia	Canada	France
Drinking water from the tap ^a (0/1)	2011	0.50	0.59	0.49	0.43
Tap water satisfaction index (0/1)	2011	0.70	0.78	0.74	0.61
Respondent's age (years)	2011	42.8	41.8	43.3	43.1
Under 18 (0/1)	2011	0.35	0.38	0.29	0.39
Log(income)	2011	10.51	10.65	10.46	10.44
Town or suburban (0/1)	2011	0.67	0.83	0.74	0.46
Car (0/1)	2011	0.90	0.93	0.86	0.93
Water quality environmental index	2011	0.48	0.67	0.45	0.35
Average regional water price (EUR/m3)	2008	2.01	1.37	1.54	2.99
Drinking water from the tap ^b (0/1) (regional average)	2008	0.70	0.83	0.67	0.63
Concern about water pollution (regional average)	2008	3.31	3.26	3.36	3.30
No. of observations	-	2,771	819	952	1,000

^a: from 2011 survey; this variable takes the value 1 if the household drinks water straight from the tap, and 0 otherwise (i.e. it takes the value 0 if water is purified, filtered, or boiled, or if bottled water is used instead).

^b: from 2008 survey; this variable takes the value 1 if the household drinks water from the tap, and 0 otherwise (i.e. it takes the value 0 only if bottled water is used).

5. Estimation results and sensitivity analyses

In Section 5.1, we compare the special regressor estimates to estimates obtained using the bivariate probit approach. For comparison purposes we also report estimates of the simple probit model and the estimates obtained using the (inconsistent) control function approach. As is usually the case in binary choice models, the estimated coefficients $\hat{\beta}$ do not have a direct interpretation and marginal effects should be reported instead. Following Lewbel et al. (2012), marginal effects are calculated using the average index function $M(\mathbf{X}'\beta + V) = E(D|\mathbf{X}'\beta + V)$. Denoting by $m(\cdot)$ the derivative of the function $M(\cdot)$, the marginal effects of a change in one regressor X on the choice probabilities are:

$$\frac{\partial E(D|\mathbf{X}'\beta + V)}{\partial X} = m(\mathbf{X}'\beta + V)\beta. \quad (12)$$

For details on the estimation of marginal effects, see Appendix A2.

In Section 5.2, we test the sensitivity of the special regressor model to the method of estimation of the density (i.e., Step 2 in Section 3.1) and the method used to control for outliers.

5.1 Comparison of alternative models

Table 2 shows (average) marginal effects estimated with four different models: (1) the special regressor model; (2) the bivariate probit model; (3) the simple probit model; and (4) the model based on the control function approach (also called IV probit).¹⁵

For the special regressor model, we used the standard kernel density approach to estimate the density in Step 2.¹⁶ The outcome of a White's general test for heteroskedasticity in the first step regression of V (the special regressor) on \mathbf{X} and \mathbf{Z} indicates rejection of the null hypothesis of homoskedasticity at the one per cent level of significance. We thus used the version of the special regressor estimator that allows for an unknown form of heteroskedasticity in V . Ninety-five per cent Winsorisation was applied on to the data to remove outliers at Step 4. For comparison purposes, the three other models (bivariate probit, simple probit, and the IV probit) were estimated on the Winsorised sample (2,771 observations overall).

¹⁵ These four models have been estimated using Stata procedures: *sspecialreg* (Baum, 2012) for Model (1), *biprobit* for Model (2), *probit* for Model (3), and *ivprobit* for Model (4).

¹⁶ To estimate the density, the Epanechnikov kernel function was used and the bandwidth was given by Silverman's rule of thumb.

In the bivariate probit model, the likelihood-ratio test that the two probit models are unrelated is strongly rejected at the one per cent level, hence confirming the endogeneity of the perceived health impacts of tap water (also called the “tap water satisfaction index” in the following). The Anderson under-identification test (p-value: 0.001) and the Sargan-Hansen test of over-identification (p-value: 0.3091) were performed at the last stage of the special regressor model estimation and confirmed the validity of the two instruments.¹⁷ The IV probit is known to produce inconsistent estimates when the endogenous variable is binary, and the simple probit estimates are biased when endogeneity is not controlled for, but these two models are shown in Table 2 for comparison purposes.

Table 2. Estimated marginal effects (standard errors in parentheses),^a 2,771 observations

Dependent variable: drinking water from the tap (0/1)	Special regressor (1)	Bivariate probit (2)	Probit (3)	IV probit (control function) (4)
Tap water satisfaction index (0/1)	0.362** (0.162)	0.626** (0.246)	0.290*** (0.019)	0.620*** (0.005)
Water price in 2008	-0.243*** (0.072)	-0.048** (0.022)	-0.120*** (0.021)	0.012 (0.015)
Respondent's age	-0.001*** (0.000)	-0.002* (0.001)	-0.001 (0.001)	-0.001*** (0.000)
Under 18 (0/1)	-0.022** (0.011)	-0.038** (0.015)	-0.054*** (0.020)	-0.016 (0.012)
Log(income)	-0.030** (0.013)	-0.075*** (0.022)	-0.082*** (0.017)	-0.050*** (0.010)
Town or suburban (0/1)	-0.003 (0.010)	-0.013 (0.014)	-0.010 (0.021)	-0.011 (0.011)
Car (0/1)	-0.029 (0.018)	-0.044* (0.024)	-0.077** (0.033)	-0.012 (0.019)
Water quality environmental index	-0.013 (0.013)	-0.024 (0.022)	0.012 (0.008)	-0.041*** (0.005)
Australia (0/1) - reference	-	-	-	-
Canada (0/1)	-0.018 (0.013)	-0.057*** (0.017)	-0.091*** (0.023)	-0.017 (0.014)
France (0/1)	0.293***	0.056*	0.078*	0.030

¹⁷ The null hypothesis of the Anderson test of underidentification is that the equation is underidentified, i.e., that the instruments are not valid in the sense that they are not correlated with the endogenous regressor. The null hypothesis of the Sargan-Hansen test of overidentifying restrictions is that the instruments are valid, i.e., uncorrelated with the error term and correctly excluded from the estimated equation. These tests for instrument validity were incorporated by the authors into the *specialreg* procedure (Baum, 2012). The updated procedure is available upon request.

	(0.075)	(0.031)	(0.042)	(0.024)
<i>In-sample predictions</i>				
Correct predictions ^b	62.6%	62.9%	64.6%	62.7%
Sensitivity	79.1%	82.7%	71.8%	83.2%
Specificity	46.3%	43.4%	57.5%	42.4%
<i>Out-of-sample predictions</i>				
Correct predictions (average over 1,000)	52.3%	63.3%	64.1%	63.1%

^a: standard errors for the special regressor and bivariate probit models were calculated using bootstrap techniques (a total of 399 bootstraps were used). Standard errors for the probit and control function models (IV probit) were obtained using the margins procedure in Stata. *, **, *** indicate significance at the 10, 5, and 1 per cent levels, respectively.

^b: The predicted probabilities for the special regressor model were calculated from the estimated index $\mathbf{X}'\hat{\boldsymbol{\beta}} + V$, as suggested by Lewbel et al. (2012). The predicted probability was set at 1 when the index was greater than 0, and 0 when the index was lower than 0.

Despite the simple probit and IV probit producing biased estimates, the percentage of in-sample correct predictions is comparable in all four models (higher by two percentage points in the simple probit). The simple probit is found to better predict the negative outcomes (specificity) than the other three models but its predictive power for the positive outcomes (sensitivity) is lower. We also report in Table 2 outcomes of out-of-sample predictions: the four models were estimated using a sub-sample that covers 85 per cent of the original sample and the estimated coefficients were used to predict the probability of drinking water from the tap on the remaining 15 per cent of the sample. Each sub-sample was drawn randomly and the exercise was repeated 1,000 times. The mean percentage of correct out-of-sample predictions is comparable to the in-sample percentage of good predictions in all models except for the special regressor, which has an average score of out-of-sample correct predictions of 52 per cent (compared to 63 per cent in-sample).

When significant, the signs of the marginal effects are consistent across the four models but the magnitude of the estimates varies. The comparison of the four models provides some insights. The bivariate probit estimates are found to be similar to estimates obtained with the IV probit model which is known to be inconsistent, and to differ significantly from the special regressor estimates, which might be an indication that the assumptions underlying the bivariate probit model may not hold for our data.¹⁸

As expected, a respondent's satisfaction with the perceived health impacts of tap water influences positively the probability of drinking water from the tap. The estimated marginal

¹⁸ It is important to keep in mind that the special regressor model also relies on assumptions; in particular, it is important for identification purposes that V has a support which is as large as $\mathbf{X}'\boldsymbol{\beta} + \mathcal{E}$. This assumption, unfortunately, is not testable since the residuals \mathcal{E} cannot be recovered.

effect varies from 0.29 in the simple probit model to 0.62-0.63 in the bivariate probit and IV probit models, while the special regressor model produces an estimate of 0.36. The IV probit and the bivariate probit produce an estimated marginal effect of the (endogenous) perception variable which is almost twice as large as the marginal effect obtained with the special regressor. This large difference between the estimated marginal effects in the two models illustrates the importance of checking the robustness to the underlying assumptions required for identification. The price of tap water (in 2008) has a negative and significant effect in all models except in the IV probit (non-significant in this case). Other significant variables include the presence of household members under 18 years of age (which decreases the probability of drinking tap water by 2 to 5 per cent depending on the model), and income and car ownership (which have a negative influence on the probability of drinking tap water). These findings are generally in line with those reported in Johnstone and Serret (2012): using the entire OECD EPIC 2008 survey data (i.e., covering ten countries), these authors found that income and car ownership influenced positively the probability that households drink bottled water. Concerns about the health impacts of tap water (treated as an exogenous variable) had a significant impact on a household's likelihood of drinking bottled water, with an elasticity estimated at 0.35.

5.2 The special regressor model: sensitivity tests

We test the sensitivity of the special regressor model to: i) the method used to estimate the density in Step 2 (ordered choice versus kernel density); ii) the choice of the bandwidth when the kernel density method is chosen; and iii) the method used to remove outliers (trimming versus Winsorisation).

In Table 3 we report estimated coefficients and standard errors for six special regressor models. In Model (1) the density in Step 2 was estimated using the ordered data estimator proposed by Lewbel and Schennach (2007) while in Models (2) to (6) the density was estimated using non-parametric kernel methods. This requires choosing a bandwidth. We compare estimates obtained using different levels of the bandwidth: 0.10, 0.23 (obtained from Silverman's rule of thumb), 0.30, 0.50, and 1.00 (obtained using cross-validation techniques). In all cases a 95 per cent Winsorisation was applied to the data.

Table 3. Estimated coefficients (standard errors in parentheses)^a for six special regressor models (2,771 observations)

Model	(1)	(2)	(3)	(4)	(5)	(6)
Density estimation	Sorted data			Kernel		
Bandwidth	-	0.10	0.23	0.30	0.50	1.00
Tap water satisfaction index	6.326*** (1.891)	4.541*** (1.357)	4.599*** (1.394)	5.142*** (1.560)	3.430*** (1.091)	2.561*** (0.848)
Respondent's age	-0.022*** (0.006)	-0.013*** (0.004)	-0.014*** (0.005)	-0.015*** (0.005)	-0.010*** (0.004)	-0.0071** (0.003)
Under 18	-0.207 (0.129)	-0.165* (0.092)	-0.167* (0.095)	-0.187* (0.106)	-0.140* (0.074)	-0.136** (0.058)
Log(income)	-0.502*** (0.158)	-0.361*** (0.113)	-0.368*** (0.116)	-0.412*** (0.130)	-0.311*** (0.091)	-0.257*** (0.071)
Town or suburban	-0.228* (0.136)	-0.144 (0.097)	-0.152 (0.100)	-0.170 (0.112)	-0.0952 (0.078)	-0.0406 (0.061)
Car	0.121 (0.210)	-0.0102 (0.150)	-0.0597 (0.155)	-0.0732 (0.173)	-0.0855 (0.121)	-0.138 (0.094)
Water quality environ. index	-0.424*** (0.154)	-0.290*** (0.111)	-0.291** (0.114)	-0.325** (0.127)	-0.194** (0.089)	-0.115* (0.069)
Australia (ref.)	-	-	-	-	-	-
Canada	0.259* (0.149)	0.157 (0.107)	0.151 (0.110)	0.166 (0.123)	0.0759 (0.086)	-0.0401 (0.067)
France	2.516*** (0.282)	2.044*** (0.203)	2.179*** (0.208)	2.447*** (0.233)	2.029*** (0.163)	1.941*** (0.127)

^a: the coefficient of the special regressor V is not shown here. Because of the normalisation of V , its coefficient is equal to one. *, **, *** indicate significance at the 10, 5, and 1 per cent levels, respectively.

Estimated coefficients, when statistically significant, are of the same sign in all six models. The magnitude varies but coefficients in the six models are not statistically different from each other. Coefficients estimated in Model (1), i.e. when the density is estimated using the ordered data estimator, are higher in magnitude than the coefficients estimated using kernel densities. Estimates obtained using the ordered data estimator of the density are known to be less efficient than estimates obtained using the kernel density approach (Lewbel et al., 2012), which is confirmed in our data. In our sample, estimates from the special regressor model thus appear to be quite robust to the method of estimation of the density in Step 2

and to the choice of the bandwidth when kernel density is used. This is probably due to the fact that a Winsorisation procedure was applied to control for outliers in T .

In Table 4, we report estimated coefficients for six special regressor models, all estimated using the kernel density approach in Step 2 (with a bandwidth chosen following Silverman's rule of thumb). We compare estimated coefficients and standard errors under trimming and Winsorisation techniques at different levels (0.5, 2.5, and 5 per cent). The size of the sample remains constant when Winsorisation is applied, whatever the level, but the number of observations in the sample varies when trimming is applied.

Table 4. Estimated coefficients (standard errors in parentheses)^a for six special regressor models with different levels of trimming and Winsorisation

Method	Winsor.	Trim.	Winsor.	Trim.	Winsor.	Trim.
Level (%)	0.5	0.5	2.5	2.5	5	5
Tap water satisfaction index	11.912*** (3.548)	3.498** (1.668)	4.599*** (1.394)	3.411*** (1.005)	2.762*** (0.839)	3.177*** (0.855)
Respondent's age	-0.034*** (0.012)	-0.012** (0.005)	-0.014*** (0.005)	-0.009*** (0.003)	-0.008*** (0.003)	-0.010*** (0.003)
Under 18	-0.123 (0.241)	-0.179* (0.107)	-0.167* (0.095)	-0.181*** (0.070)	-0.132** (0.057)	-0.142** (0.062)
Log(income)	-0.919*** (0.296)	-0.318** (0.135)	-0.368*** (0.116)	-0.297*** (0.086)	-0.236*** (0.070)	-0.237*** (0.076)
Town or suburban	-0.318 (0.255)	-0.208* (0.112)	-0.152 (0.100)	-0.058 (0.072)	-0.075 (0.060)	-0.061 (0.064)
Car	0.229 (0.393)	-0.023 (0.174)	-0.060 (0.155)	-0.136 (0.113)	-0.094 (0.093)	-0.138 (0.100)
Water quality environ. index	-0.891*** (0.290)	-0.193 (0.136)	-0.291** (0.114)	-0.192** (0.081)	-0.153** (0.068)	-0.179** (0.070)
Australia (ref.)	-	-	-	-	-	-
Canada	0.382 (0.279)	0.048 (0.123)	0.151 (0.110)	0.071 (0.080)	0.061 (0.066)	-0.010 (0.071)
France	3.616*** (0.530)	2.184*** (0.249)	2.180*** (0.208)	1.705*** (0.156)	1.659*** (0.125)	1.540*** (0.140)
No. of obs.	2,771	2,758	2,771	2,702	2,771	2,633

Even if the 95 per cent confidence intervals overlap and estimated coefficients are not statistically different across the six models, the magnitude of the estimated coefficients can vary significantly depending on the level of trimming and Winsorisation. For example, the

estimated coefficient of the respondent's tap water satisfaction index in the model with 99 per cent Winsorisation (level set at 0.5 per cent) is about three times as large as the coefficients estimated in the other models (and similarly for all variables). This may be an indication that the level (0.5 per cent) is too low and that some outliers are still present in the sample. In such a situation, trimming the data appears to be a better solution than Winsorisation. When the level is set at 2.5 or 5 per cent, trimming and Winsorisation produce more comparable estimates and standard errors. It thus appears advisable to estimate the special regressor model under different levels of trimming and/or Winsorisation.

5.3 Strengths and weaknesses of the special regressor

Our empirical application illustrates some of the strengths and weaknesses of the special regressor approach. This approach is a multi-step procedure that is relatively easy to implement. It involves the estimation of a density of first-stage residuals and we found in our data that estimates were quite robust to the choice of the bandwidth when using a kernel method and when outliers were controlled for through Winsorisation. However, Lewbel (2000) showed that the estimation of the density can generate some non-negligible biases (see results of simulations in Lewbel, 2000). This might also be one reason behind the lower performance of the special regressor estimator in terms of out-of-sample predictions.

Very low values of the estimated density can also produce very large values of the variable T used in the final regression. Since such outliers could induce very large standard errors in the final stage regression, outliers should be detected and removed. In our sample, a 99 per cent Winsorisation (level set at 0.5 per cent) was not sufficient to exclude all outliers and led to biased estimates, so we recommend checking the robustness of estimates obtained under different levels of trimming/Winsorisation.

An interesting feature of the special regressor model is that only one special regressor is needed whatever the number of endogenous regressors and that usual tests for instruments' validity (under- and over-identification tests) can be applied at the final stage of the procedure.

The special regressor also avoids making strong assumptions about the relationship between the endogenous variable and the exogenous ones (including the instruments) and is free from assumptions on the distribution of the model error terms. The special regressor approach requires a special regressor which needs to fulfil a set of different assumptions. The monotonicity condition on the relationship between the variable D (describing the binary decision of interest) and the special regressor can be tested on the data but the large support condition is not testable, which is probably the main drawback of the special regressor method.

6. Conclusion

In this article we argue that variables measuring individuals' perceptions should be treated as endogenous in empirical averting-behaviour models. We discuss an approach that has not been considered so far in the literature and that could be seen as an interesting alternative to the traditional bivariate probit model when averting decisions are binary and perceptions are recorded as discrete variables.

In our sample, the perception variable (or tap water satisfaction index) was found to be endogenous and tests confirmed the validity of the chosen instruments. The average price of water in the region in 2008 was used as the special regressor, which is exogenous to a household's decision to drink tap water in 2011. The monotonic relationship between the special regressor and the probability of drinking tap water was confirmed in our data and, because the price of water varied significantly across the three countries and covered a large support, we believe the chosen special regressor satisfied all the required conditions for the proper identification of the parameters of interest. Estimates obtained with the special regressor model and the bivariate probit model differed significantly. The similarity between estimates from the bivariate probit model and from the IV probit model which is known to be inconsistent, thus cast doubt on the validity of the assumptions underlying the bivariate probit model. We also show that, even if the estimation of a density is involved, the special regressor estimates are quite robust to the choice of the bandwidth, providing that outliers are properly handled using either trimming or Winsorisation.

We encourage practitioners to use the special regressor model if a special regressor is available. The special regressor model relies on very different assumptions from the bivariate probit model, it is relatively easy to implement, and can thus be useful for robustness checks. It is also important to note that in this article we focus on the case of an endogenous binary variable, but the special regressor approach can also be applied if the endogenous variable is categorical or continuous.

Acknowledgements

We are grateful to the Environment Directorate of the OECD for allowing us to use the 2008 and 2011 EPIC survey data. We are also grateful to Thierry Magnac for comments and suggestions on an earlier version of this draft.

References

- Abrahams, N.A, Hubbell, B.J., and J.L. Jordan, 2000. Joint production and averting expenditure measures of willingness to pay: Do water expenditures really measure avoidance costs? *American Journal of Agricultural Economics* 82(2): 427–437.
- Baum, C.F., 2012. *Sspecialreg: Stata module to estimate binary choice model with discrete endogenous regressor via special regressor method*. Available at <http://ideas.repec.org/c/boc/bocode/s457546.html> [accessed: 1 October 2014].
- Bontemps, C., and C. Nauges, 2009. Carafe ou bouteille ? Le rôle de la qualité de l'environnement dans la décision du consommateur. *Economie et Prévision* 2009-2, 188: 61-79.
- Courant, P.N., and R.C. Porter, 1981. Averting expenditure and the cost of pollution. *Journal of Environmental Economics and Management* 8(4): 321–329.
- Dickie, M., and S. Gerking, 1996. Formation of risk beliefs, joint production and willingness to pay to avoid skin cancer. *The Review of Economics and Statistics* 78(3): 451–463.
- Dong, Y., and A. Lewbel, forthcoming. A simple estimator for binary choice models with endogenous regressors. *Econometric Reviews*.
- Goux, D., and E. Maurin, 2005. The effect of overcrowded housing on children's performance at school. *Journal of Public Economics* 89(5-6): 797–819.
- Heckman, J.J., 1976. *Simultaneous equation models with both continuous and discrete endogenous variables with and without structural shift in the equations*. In Steven Goldfeld and Richard Quandt (Eds.), *Studies in Nonlinear Estimation*, Ballinger.
- Janmaat, J., 2007. A little knowledge...: Household water quality investment in the Annapolis valley. *Canadian Journal of Agricultural Economics* 55(2): 233–253.
- Johnstone, N., and Y. Serret, 2012. Determinants of bottled and purified water consumption: results based on an OECD survey. *Water Policy* 14(4): 668–679.
- Konishi, Y., and K. Adachi, 2011. A framework for estimating willingness-to-pay to avoid endogenous environmental risks. *Resource and Energy Economics* 33(1): 130–154.
- Lewbel, A., 2000. Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables. *Journal of Econometrics* 97(1): 145–177.
- Lewbel, A., 2014. *An overview of the special regressor method*. In the *Oxford Handbook of Applied Nonparametric and Semiparametric Econometrics and Statistics*, co-edited by A. Ullah, J. Racine, and L. Su: 38–62, Oxford University Press.
- Lewbel, A., and S. Schennach, 2007. A simple ordered data estimator for inverse density weighted functions. *Journal of Econometrics* 136(1): 189–211.

- Lewbel, A., Dong, Y., and T.T. Yang, 2012. Viewpoint: Comparing features of convenient estimators for binary choice models with endogenous regressors. *Canadian Journal of Economics* 45: 809–829.
- Maurin, E., 2002. The impact of parental income on early schooling transitions: a re-examination using data over three generations. *Journal of Public Economics* 85(3): 301–332.
- OECD, 2011. *Greening Household Behaviour: The Role of Public Policy*, OECD, Paris.
- OECD, 2013. *Greening Household Behaviour: Overview from the 2011 Survey*, OECD, Paris.
- Tukey, J.W., 1962. The future of data analysis. *Annals of Mathematical Statistics* 33: 1–67.
- Um, M.J., Kwak, S.J., and T.Y. Kim, 2002. Estimating willingness to pay for improved drinking water quality using averting behavior method with perception measure. *Environmental and Resource Economics* 21(3): 287–302.
- Whitehead, J.C., 2006. Improving willingness to pay estimates for quality improvements through joint estimation with quality perceptions. *Southern Economic Journal* 73(1): 100–111.
- Wooldridge, J.M., 2010. *Econometric Analysis of Cross Section and Panel Data*, second edition. Cambridge, MA: MIT Press.

Appendix A1. Latent variable distribution identification and computation of T

Lewbel (2014) proposes an intuitive illustration of the special regressor method that we use here to derive the computation of the (unconditional) special regressor. Suppose we run an experiment in which we ask each participant whether or not he/she would be willing to pay p dollars to preserve a wetland or any other natural resources. We then observe the outcome as a *yes* or *no* through the binary decision variable D :

$$D = \mathbf{I}(W^* \geq P) \quad (\text{A1.1})$$

where W^* is a latent variable representing the individual (unobserved) willingness to pay. In such an experiment, P is chosen independently from W^* so that, for each proposed price p :

$$E[D | P = p] = \Pr[D = 1 | P = p] = \Pr[W^* \geq p] = 1 - \Pr[W^* < p] = 1 - F_{w^*}(p)$$

where $F_{w^*}(\cdot)$ is the probability density function of the latent variable W^* . Equivalently, this means that:

$$F_{w^*}(p) = 1 - E[D | P = p] \quad (= \Pr[D = 0 | P = p]) \quad . \quad (\text{A1.2})$$

So for each value of p , the observed probability of *not* accepting paying p dollars gives an unbiased estimation of one value in the distribution of the willingness to pay $F_{w^*}(p)$. Proposing different values for p such that the (unknown) support of W^* is covered, would provide additional information on F_{w^*} and hence “draw a picture” of the distribution of the willingness to pay distribution.

In this example, the variable P is a special regressor. If P is continuously distributed, then one could use the Equation (A1.2) and estimate the distribution function $\hat{F}_{w^*}(p)$ using the estimation of the probability of not willing to pay p dollars, that is using $\hat{E}[1 - D | P = p]$, a non-parametric estimator of the regression of $1 - D$ on P .

Then, the average willingness to pay in the population $E(W^*)$ is given by:

$$E(W^*) = \int_{w_L}^{w_U} w^* f_{w^*}(w^*) dw^* = \int_{w_L}^{w_U} w^* \frac{\partial F_{w^*}(w^*)}{\partial w^*} dw^* \quad (\text{A1.3})$$

where $[w_L, w_U]$ is the (unknown) support of W^* . This shows that $E(W^*)$ is identified and can be estimated by replacing $F_{w^*}(\cdot)$ with its non-parametric estimator $\hat{F}_{w^*}(\cdot)$ in Equation (A1.3).

In practice, deriving the average willingness to pay from (A1.3) would require computing the derivative of a non-parametric regression function, and taking its first moment. Fortunately, there exists an alternative way of computing $E[W^*]$, and this computation will lead to the expression of the (unconditional) special regressor estimator T . Rewriting Equation (A1.3) and integrating by parts gives:

$$E(W^*) = \int_{w_L}^{w_U} w^* \frac{\partial F_{w^*}(w^*)}{\partial w^*} dw^* = - \int_{w_L}^{w_U} F_{w^*}(w^*) dw^* + [w F_{w^*}(w^*)]_{w_L}^{w_U}. \quad (\text{A1.4})$$

Let us define $V = -P$.¹⁹ Now we introduce $H(v) = E[D|V=v]$ in Equation (A1.2) so that $F_{w^*}(-v) = 1 - H(v)$. With this new notation, $E(W^*)$ becomes:

$$\begin{aligned} E(W^*) &= - \int_{w_L}^{w_U} (1 - H(-w^*)) dw^* + [w F_{w^*}(w^*)]_{w_L}^{w_U} \\ &= \int_{w_L}^{w_U} H(-w^*) dw^* - \int_{w_L}^{w_U} 1 \cdot dw^* + [w F_{w^*}(w^*)]_{w_L}^{w_U}. \end{aligned}$$

Now let us choose an arbitrary value $-c$ and split the second integral at that threshold.²⁰

$$E(W^*) = \int_{w_L}^{w_U} H(-w^*) dw^* - \int_{w_L}^{-c} 1 \cdot dw^* - \int_{-c}^{w_U} 1 \cdot dw^* + [w F_{w^*}(w^*)]_{w_L}^{w_U}$$

or, by re-arranging things:

$$\begin{aligned} E(W^*) &= \int_{w_L}^{w_U} H(-w^*) dw^* - \int_{w_L}^{w_U} \mathbf{I}(w^* < -c) dw^* \\ &\quad - \int_{w_L}^{w_U} \mathbf{I}(w^* \geq -c) \cdot dw^* + [w F_{w^*}(w^*)]_{w_L}^{w_U} \end{aligned} \quad (\text{A1.5})$$

The last two terms of the previous equation are mostly boundary related terms that will disappear thanks to the fact that $F_{w^*}(w_L) = 0$ and $F_{w^*}(w_U) = 1$.²¹ So $E(W^*)$ can be rearranged:

¹⁹ Note that V is continuous if P is continuous, and is independent from w^* , and therefore is also a special regressor in the latent variable model re-defined from Equation (A1.1) to become $D = \mathbf{I}(V + W^* \geq 0)$.

²⁰ Later on, we will choose $-c = 0$.

²¹ See Lewbel (2014) for details in computing these boundary terms. One may note that these terms sum to zero whatever the value of c .

$$E(W^*) = \int_{w_L}^{w_U} H(-w^*)dw^* - \int_{w_L}^{w_U} \mathbf{I}(w^* < -c) dw^* = \int_{w_L}^{w_U} H(-w^*)dw^* - \int_{w_L}^{w_U} \mathbf{I}(-w^* > c)dw^*$$

by changing variable $-w^* \rightarrow v$

$$E(W^*) = \int_{-w_U}^{-w_L} H(v)dv - \int_{-w_U}^{-w_L} \mathbf{I}(v \geq c) dv = \int_{-w_U}^{-w_L} (H(v) - \mathbf{I}(v \geq c)) dv$$

or,

$$E(W^*) = \int_{-w_U}^{-w_L} \frac{H(v) - \mathbf{I}(v \geq c)}{f_v(v)} f_v(v) dv. \quad (\text{A1.6})$$

If we recall that $H(v) = E[D | V = v]$ and replacing H in Equation (A1.6):

$$\begin{aligned} &= \int_{-w_U}^{-w_L} \frac{E[D | V = v] - \mathbf{I}(v \geq c)}{f_v(v)} f_v(v) dv \\ &= \int_{-w_U}^{-w_L} E \left[\frac{D - \mathbf{I}(v \geq c)}{f_v(V)} | V = v \right] f_v(v) dv \end{aligned} \quad (\text{A1.7})$$

If we set $c = 0$, by the law of iterative expectations, this last term is also $E(T)$ for T defined as:

$$T = \frac{D - \mathbf{I}(V \geq 0)}{f_v(V)}.$$

These elements of computation show how the estimator T is constructed, and how it relates to the identification of the latent variable W^* in this unconditional setting. In a conditional framework, where the latent variable is defined as $W^* = X'\beta + \varepsilon$, the intuition is the same and the computation is made along the same lines leading to the special regressor estimator:

$$T_2 = \frac{D - \mathbf{I}(V \geq 0)}{f_{v|x}(V | X)}.$$

Appendix A2. Estimation of the marginal effects (special regressor approach)

The marginal effects of a change in a regressor X on the choice probabilities are:

$$\frac{\partial E(D | \mathbf{X}'\boldsymbol{\beta} + V)}{\partial X} = m(\mathbf{X}'\boldsymbol{\beta} + V)\boldsymbol{\beta}, \quad (\text{A2.1})$$

and the estimated (average) marginal effects are given by $\bar{m}\hat{\boldsymbol{\beta}} = \frac{1}{n} \sum_{i=1}^n \hat{m}_i \hat{\boldsymbol{\beta}}$ where i is the index for the observation. The estimator of m_i is:

$$\hat{m}_i = \frac{\frac{1}{h} \sum_{j=1}^n (D_j - \hat{M}_j) K' \left(\frac{\mathbf{x}_i' \hat{\boldsymbol{\beta}} - \mathbf{x}_j' \hat{\boldsymbol{\beta}}}{h} \right)}{\sum_{j=1}^n K \left(\frac{\mathbf{x}_i' \hat{\boldsymbol{\beta}} - \mathbf{x}_j' \hat{\boldsymbol{\beta}}}{h} \right)}, \quad (\text{A2.2})$$

where $K(\cdot)$ and $K'(\cdot)$ are a standard kernel function and its derivative, respectively, and h denotes the bandwidth. The estimated average index function for observation i is:

$$\hat{M}_i = \frac{\sum_{j=1}^n D_j K \left(\frac{\mathbf{x}_i' \hat{\boldsymbol{\beta}} - \mathbf{x}_j' \hat{\boldsymbol{\beta}}}{h} \right)}{\sum_{j=1}^n K \left(\frac{\mathbf{x}_i' \hat{\boldsymbol{\beta}} - \mathbf{x}_j' \hat{\boldsymbol{\beta}}}{h} \right)}. \quad (\text{A2.3})$$