Using multiple imputation for a zero-inflated contingent valuation with potentially biased sampling

Marouene MBAREK[*], Damien ROUSSELIERE[†], Julien SALANIE[‡]

December 2015

Abstract

The contingent valuation steels a recommended tools for the environment policies valuation. It helps decision makers to choose the relevant choice concerning the environment policy through the valuation of non market goods. By this article, we try to determine the Willingness To Pay (WTP) of a marine protected area (MPA) throw a contingent valuation survey. To estimate the WTP, we use the Zero Inflated Ordered Probit (ZIOP) model which had been developed by Harris and Zhao (2007) to overtake the excess zeros problem. The ZIOP model has shown that 29% of null wtp with 54,4% protest response and the covariates which influence the WTP amount are the education level, residency and income. Meanwhile, the project (MPA) acceptability is widely influenced by the age and education level.

Key words: Contingent valuation, Multiple Imputation, Sample and Non-response biases

1 Introduction

The determination of non-market goods' value has allowed to assess the economic effects of environmental policies (Carson et al., 2001). To give a monetary value of non marketed amenities is very important for the decisions makers. It helps them to choose the appropriate environmental and conservation policies.

To evaluate the non market goods, several technical tools have been proposed such as stated preference (SP) methods. The SP methods have ensured the evaluation of an important part of environmental goods (Andersson and Svensson, 2008). Among the SP methods, we find the contingent valuation method (CVM) which has been developed like a non market valuation

[*]GRANEM, Angers University, *mbarekmarouene@gmail.com*

[†]GRANEM, Agrocampus-Ouest (Angers), *damienrousseliere@gmail.com*

[‡]UMR 5824,Université Jean Monnet GATE LSE Saint -Etienne, *julien.salanie@univ-st-etienne.fr*

and become increasingly known in environmental economics especially for amenities and damages evaluation (Noonan, 2003; Carson et al., 2001; Mitchell, 2002).

It allows to determine the use and existence value of assets (Mitchell and Carson, 1989). It consists to go directly to individuals to determine the price of a natural asset (example: "*What would you pay for environmental safety?*", "*What would you pay for wilderness?*" (Hanemann, 1994)), which represented by the willingness to pay (WTP) or the willingness to accept (WTA). This method has succeeded in some measure to help decision-makers to choose whether or not such an environmental and conservation policies. However, its reliability to determine the economic value of non market goods, steels a persisting debate which doesn't have an end (Carson et al., 2001). This is due to the existing of several biases and limitations of the CVM (Diamond and Hausman, 1994; Ariely et al., 2003; Blumenschein et al., 2008). The recent works have considered the sources of bias regarding the use of CVM which conducted with bad survey design in many cases, such attention has enhanced the CV results (Blumenschein et al., 2008; Bateman et al., 2008; Zhang et al., 2010; Lusk and Hudson, 2004; Clinch and Murphy, 2001; Andersson and Svensson, 2008).

In several surveys which had been conducted using the CVM, some respondents may do not give their real value for the good, and thereafter, the SP methods will determine incorrectly the value of the good in question (Meyerhoff and Liebe, 2008). Indeed, it's difficult to know if these respondents, which give a false value for the good, have a zero or positive WTP. In which case, the good's economic value may be underestimate or overestimate. In addition, there is no consensus about the procedure to distinguish the true zero from protest responses. There are a lot of reasons that respondents state a wrong value of the good such as they may not understand the survey and give a false valuation, they may prefer that the change in the good will paid by other people, etc (Meyerhoff and Liebe, 2008). To overcome this problem, the solution which has been proposed is to keep only the true responses in the sample. But this solution, which means to exclude protest responses, hasn't any theoretical conception (Meyerhoff and Liebe, 2008).
To improve the CVM's results and overstep its hypothetical bias, the practitioners have used an additional techniques for the robustness and reliability tests such as "split sample approach" and "entreaties". The split sample aproach means to state the same questionnaires to independent samples at two different points in time for comparing the twice individual preferences (exemple: WTP) if they will change. Some surveys in this sense had shown the same results (WTP) in different times (Carson et al., 2001). Among the entreaties, we find the "cheap talk"which sends in "the costless of transmission of signals and information" (Cummings and Taylor, 1999; Atkinson et al., 2012). These additional techniques to CVM may be effective to decrease the hypothetical bias and can partially tackle protest responses. In this way, how the advocates of CVM can distinguish the true responses from the protest responses which, technically, are represented by an excess zeros in the datasets?

Some work in medical research have first highlighted the datasets with excess zeros (Moulton and Halsey, 1995; Aitchison, 1955). This statistical analysis issue is very common in other fields like finance, economics, insurance, etc. It's admitted as a recurring problem (Lachenbruch, 2002). To overtake this problem, an alternative models have been introduced like Heckman model[1]. However these models encounter a difficulties to achieve their goal and give a skewness estimations especially with a specific datasets like health insurance data (Duan et al., 1983; Lachenbruch, 2002). Excess zeros in datasets allows for overdispersion witch makes standard errors and chi-square statistics are worthless and maximum likelihood estimates inefficient. As mentioned above, the conventional models have failed to resolve the overdispersion and fit poorly when the number of individuals with a count of zero is large in datasets (Allison, 2012, chap 9). To overcome this problem, a zero inflated models[2] have been presented to give treatment to the excess zeros (Allison, 2012; Greene, 2007). The datasets with excess zeros called *zero-inflated data*. To mention the excess zeros, the term "inflation" has been used (Pimentel et al., 2015). The zero inflated model is introduced in the case where are two regimes of zeros in the datasets. It's considered like a latent class model (Greene, 2007).

In mostly, zero inflated models have been considered as an appropriate estimates to analyze the skewness datasets with excess zeros which have two regimes. Moreover, they have been suggested to overtake the problem of incomplete and missing data (Xie et al., 2013). Their success to reach this goal still partial because the treatment of missing data for example needs another statistical techniques like multiple imputation and bootstrap, besides zero inflated models.

Our case study is a contingent valuation method on the WTP (willingness to Pay) for accessing a MPA (Marine Protected Area) in a developing country (Tunisia). To our knowledge, previous studies did not correct for biased sampling (e.g. (Petrosillo et al., 2007)). In july and august 2012, 315 among 40000 Tourists were surveyed randomly during their visit to Kuriat islands (Monastir, Tunisia). Our econometric model are the ZIOP (Zero-inflated ordered probit) which is a double-hurdle combination of a split probit model and an ordered probit model (Harris, 2007) and its extension the ZIOPC which assumes that the two errors terms are correlated (Bagozzi et al., 2015). This model address the problem of two distinct data generating process for the zero. One type of zero corresponds to individuals who will always refuse to pay for MPA (because of ideological reasons), whereas the other type refers to a corner solution. The ZIOP models will be completed by a sensitive analysis using the multiple imputation (MI) to address a potentially biased sampling (Van Buuren, 2012). Before the presentation of the econometric models, we present the MPA in the section 2. The

[1] Among the alternative models we can cite: Heckman model or two part models (Lachenbruch, 2001, 2002; Berk, 2002), probit regression for the individuals with or not a count of zero and multiple regression for the nonzero part (Duan et al., 1983), censored log-gamma (Moulton and Halsey, 1996), censored lognormal mixture model (Taylor et al., 2001), lognormal distribution (Xiao-Hua and Tu, 1999).

[2] Among zero inflated models, we can evoke the zero inflated poisson, zero inflated negative binomial, zero inflated ordered probit model, etc. For our database, we use a zero inflated ordered probit model because the dependent variable in the second stage is a categorical variable and not continuously. For more detail, see section 3

case study will be presented in the section 4 and the results and conclusion take place in the sections 5 and 6 respectively.

2 MPA and the CVM

Due to the convention on biological diversity in 1992, the number of marines protected areas around the world has been increased to attend more than 5000 in 2012[3]. The aim of the establishment of MPA is the protection of marine and coastal ecosystems and the development of the several human activities such as coastal fishery, ecotourism, scuba diving and other recreational activities. Most of the MPA didn't reach their goal because the actors involved are various and numerous, which case mentions a complex system governing by MPA. Also, they don't find a satisfactory funds to ensure their sustainable management especially in the developing countries which receive an external short funds for taking place MPAs(Jentoft, 2007; Baral et al., 2008). Face to this condition, the practitioners have suggested a self financing of MPA to overtake their failure. In the most cases, the MPA management warranted by the fees access paid by the visitors. This solution helps MPA to be financed and managed for long run.
To make up a consistent management system of MPA, it's recommended to carry out a fee amount equals to the willingness to pay (WTP) declared by the visitors. This fee amount may be determined by the CVM which require to conduct a direct interviews with visitors. The choice to use CVM returns to the sort of MPA benefits which, in most of the time, don't have a market value[4]. Among the first who used the CVM to assess the WTP of visitors and help the decision makers to establish various types of MPA, we find the National Oceanic and Atmospheric Association (NOAA)[5] which had conducted their CVM survey in 1989(Hall et al., 2002). From that, the use of the CVM has been increased to determine the non use value of MPA such as the improvement of environmental area which evaluated as tourist entry fees(Baral et al., 2008; Hall et al., 2002; Bhat, 2003; Asafu-Adjaye and Tapsuwan, 2008; Lee and Han, 2002; Dharmaratne et al., 2000; Sumaila et al., 2000; Arin and Kramer, 2002).
To take place of MPA, some surveys conducted using the CVM may have a missing data. Such case makes the conventional models and statistical techniques unable to give a fair estimate of WTP especially with missing data or excess zeros. For this reason, we use for our study one of the zero inflated models: zero inflated ordered probit model which considers the weakness of CVM evoked above.

[3]Information collected from the web site of the International Union for Conservation of Nature (IUCN) www.IUCN.org, seen in September 16th 2015.

[4]Here we mention that much of MPA benefits don't have market value and need a hypothetical market to be valued such as the improvement of the coastal ans marine ecosystem which attract more visitors, increase of fish quality, upturn of environment, protection of threatened species, etc.

[5]For more detail, see the work of Hall et al. 2002 which presents an exhaustive list of the surveys of NOAA and other using the CVM to determine the WTP of visitors

3 Econometric models

The zero inflated models[6] represent an alternative to the conventional models using databases with missing data or excess zeros. As we explained above, there are different zero inflated models such as zero inflated Poisson, zero inflated negative binomial, zero inflated ordered probit. The choice to one of these models depends on the nature of the datasets, particularly the dependent variable in the model. For our case, the dependent variable - WTP - is a discrete ordered variable which characterized by excess zero. At this level, the conventional truncated ordered probit model may encounter the problem of overdispersion and give an inconsistent estimates(Weiss, 1993). Thus, we choose the zero inflated ordered probit model which best corresponds to our database.

3.1 Zero inflated ordered probit (ZIOP) model

Like the zero inflated (ZI) models, the ZIOP model can be implemented on two stages. The ZI model has a logistic regression in the first stage and a linear regression in the second stage but the ZIOP model has, respectively, a probit "splitting" model and an ordered probit model[7]. The two different sets of covariates split the observations into two regimes (Harris, 2007). We denote l a binary variable which determines the two regimes (regime 0: non participant if $l = 0$ and regime 1: participant if $l = 1$). The propensity for participation is l^* which $l^* > 0$ if $l = 1$ and $l^* \leq 0$. It is defined as $l^* = \Gamma'\alpha + \mu$, where Γ is a vector of variables which determine the two regimes, α a vector of coefficients, and μ the error term. The probability of an individual can be in regime 1 is:

$$Pr(l = 1|\Gamma) = Pr(l^* > 0|\Gamma) = \Psi(\Gamma'\alpha)$$

where $\Psi(.)$ is the cumulative distribution function $(c.d.f)$ of the univariate standard normal distribution. For our case, the WTP seems like a discrete random variable λ which has discrete ordered values. In the case that $l = 1$, the WTP under regime 1 given by a discrete variable $\tilde{\lambda}$ $(\tilde{\lambda} = 0, 1, \ldots, J)$ that is estimated by an ordered probit in the second stage in ZIOP and given by $\tilde{\lambda}^*$; $\tilde{\lambda}^* = \tau'\beta + \nu$ with τ is a vector of covariates with unknown weights β and ν the error term. The mapping between $\tilde{\lambda}^*$ and $\tilde{\lambda}$ may written by

$$\tilde{\lambda} = \begin{cases} 0 & \text{if } \tilde{\lambda}^* \leq 0 \\ j & \text{if } \nu_{j-1} < \tilde{\lambda}^* \leq \nu_j (j = 1, \ldots, J-1) \\ J & \text{if } \nu_{j-1} \leq \tilde{\lambda}^* \end{cases}$$

where ν_j $(j = 1, \ldots, J-1)$ are parameters to be estimated in addition to β. The ordered

[6]Allison P, 2012 detailed the different zero inflated models and carries out their comparison with the conventional models (Allison, 2012, chap 9)

[7]The probit and ordered probit models are estimated in the same time with different sets of variables.

probit probabilities are represented by (Harris, 2007)

$$Pr(\tilde{\lambda}) = \begin{cases} Pr(\tilde{\lambda} = 0|\tau, l = 1) = \Psi(-\tau'\beta) \\ Pr(\tilde{\lambda} = j|\tau, l = 1) = \Psi(\nu_j - \tau'\beta) - \Psi(\nu_{j-1} - \tau'\beta)(j = 1, \dots, J-1) \\ Pr(\tilde{\lambda} = J|\tau, l = 1) = 1 - \Psi(\nu_{J-1} - \tau'\beta) \end{cases}$$

with $\lambda = l\tilde{\lambda}$. There are two cases where the individual mentions a null WTP; $l = 0$ leads that $\lambda = 0$ and $l = 1$ may lead to $\tilde{\lambda} = 0$. The case when WTP is positive, $l = 1$ and $\tilde{\lambda}^* > 0$. The probabilities for λ may be presented by

$$Pr(\lambda) = \begin{cases} Pr(\lambda = 0|\tau, \Gamma) = Pr(l = 0|\Gamma) + Pr(l = 1|\Gamma)Pr(\tilde{\lambda} = 0|\tau, l = 1) \\ Pr(\lambda = j|\tau, \Gamma) = Pr(l = 1|\Gamma)Pr(\tilde{\lambda} = j|\tau, l = 1)(j = 1, \dots, J) \end{cases}$$

$$= \begin{cases} Pr(\lambda = 0|\tau, \Gamma) = [1 - \Psi(\Gamma'\alpha)] + \Psi(\Gamma'\alpha)\Psi(-\tau'\beta) \\ Pr(\lambda = j|\tau, \Gamma) = \Psi(\Gamma'\alpha)[\Psi(\nu_j - \tau'\beta) - \Psi(\nu_{j-1} - \tau'\beta)](j = 1, \dots, J) \\ Pr(\lambda = J|\tau, \Gamma) = \Psi(\Gamma'\alpha)[1 - \Psi(\nu_{J-1} - \tau'\beta)] \end{cases}$$

with the condition that μ and ν follow the standard Gaussian distributions. By the ZIOP, the probability of a zero observation has been presented like a combination of the probability of "zero consumption" from the OP process and the probability of "non-participation" from the "split probit model" (Harris, 2007). The parameters of the model $\omega = (\alpha', \beta', \nu')'$ can be estimated by maximum likelihood criteria. The log-likelihood function can be represented as follows:

$$l(\omega) = \sum_{i=1}^{N} \sum_{j=0}^{J} \zeta_{ij} ln[Pr(\lambda_i = j|, \beta, \nu, \omega)]$$

with $i = 1, \dots, N$ and $j = 0, 1, \dots, J$. The function ζ is defined as:

$$\zeta_{ij} = \begin{cases} 1 & \text{if individual } i \text{ chooses outcome } j \\ 0 & \text{otherwise} \end{cases}$$

3.2 ZIOP model's extension: ZIOP Correlated

The result of the two separate equations, $l^* = \Gamma'\alpha + \mu$ and $\tilde{\lambda}^* = \tau'\beta + \nu$, with uncorrelated terms, give the observed variable λ. According to Harris and Zhao (2007), the two error terms μ and ν are related in view of λ, which given by "two separate latent equations", tie in the same individual. The two authors had extended the model to include (μ,ν). These two stochastic terms follow a bivariate normal distribution with correlation coefficient ρ. Thus, the observations can be presented as follows:

$$\lambda = l\tilde{\lambda} = \begin{cases} 0 & \text{if } (l* \leq 0) \text{ or } (l* \succ 0) \text{ and } (\tilde{\lambda}^* \leq 0) \\ j & \text{if } (l* \succ 0) \text{ and } (\nu_{j-1} \prec \tilde{\lambda}^* \leq \nu_j) \; (j = 1, \dots, J-1) \\ J & \text{if } (l* \succ 0) \text{ and } (\nu_{J-1} \prec \tilde{\lambda}^*) \end{cases}$$

and the probabilities become:

$$Pr(\lambda) = \begin{cases} Pr(\lambda = 0|\tau, \Gamma) = [1 - \Psi(\Gamma'\alpha)] + \Psi_2(\Gamma'\alpha, -\tau'\beta; -\rho) \\ Pr(\lambda = j|\tau, \Gamma) = \Psi_2(\Gamma'\alpha, \nu_j - \tau'\beta; -\rho) - \Psi_2(\Gamma'\alpha, \nu_{j-1} - \tau'\beta; \rho)](j = 1, \ldots, J-1) \\ Pr(\lambda = J|\tau, \Gamma) = \Psi_2(\Gamma'\alpha, \tau'\beta - \nu_{J-1}; \rho)] \end{cases}$$

where Ψ_2 is the cumulative distribution function of the bivariate normal distribution with correlation coefficient ρ between the two "univariate random elements". The parameters of the model ω are redefined and and become$\omega = (\alpha', \beta', \nu', \rho)'$. Harris and Zhao (2007) had suggested a Wald test of $\rho = 0$ like a test for knowing if the two error terms are correlated or no. This test may be conducted by the ZIOPC model against the null of ZIOP model.

3.3 Multiple imputation (MI)

Despite it's critics, multiple imputation considered in mostly as relevant means to deal with the incomplete observations which encounter the survey data (Van Buuren, 2012, chap 2). It has been introduced and developed as an important statistical means applied to data which had been collected and it'll be analyzed (Meng, 1994).
For each missing element in the data matrix, MI impute z values. Like that we obtain z "completed" database which holds in observed ans unobserved data (Honaker et al., 2012). We denote $\boldsymbol{D}$ includes observed ans missing data; $\boldsymbol{D} = \left\langle \boldsymbol{D}^{obs}, \boldsymbol{D}^{mis} \right\rangle$. $\boldsymbol{D}$ is multivariate normal distribution with mean vector σ and covariance matrix θ, $\boldsymbol{D} \sim \mathcal{N}(\sigma, \theta)$.
T represents the missing data, with elements $z_{ij} = 1$ if $d_{ij} \in \mathbf{D}^{mis}$ and $z_{ij} = 0$ otherwise. The data are *missing at random* (MAR) in multiple imputation, which means that T may be imputed only by the observed data $\boldsymbol{D}^{obs}$. The MAR assumption is defined as $p(T|\boldsymbol{D}) = p(T|\boldsymbol{D}^{obs})$. It includes the missing values that created randomly and would be better assumption when the missingness is related to the data. The data are *missing completely at random* (MCAR)[8] in the case when the missingness is not dependent on the data (Honaker et al., 2012; Honaker and King, 2010; Van Buuren, 2012). In our case, we use the MAR assumption which required by $AMELIA$ beside the multivariate normality especially the missigness is dependent on the data.
The likelihood of observed data is $p(\boldsymbol{D}^{obs}, T|\Delta)$, which $\boldsymbol{D}^{obs}$ is the observed data, T is the missingness matrix and Δ is the complete data parameters ($\Delta = (\sigma, \theta)$). By the MAR assumption, it can be written as

$$p(\boldsymbol{D}^{obs}, T|\Delta) = p(T|\boldsymbol{D}^{obs})p(\boldsymbol{D}^{obs}|\Delta)$$

The likelihood for complete data is:

$$L(\Delta|D^{obs}) \infty p(\boldsymbol{D}^{obs}|\Delta)$$

[8]See Van Buuren 2012 (chapter 1) for more detail about missing data problems

By iterated expectations, we can rewrite it as:

$$p(\boldsymbol{D}^{obs}|\Delta) = \int p(\boldsymbol{D}|\Delta)\, dD^{miss}$$

The posterior can be determined by the likelihood for complete data and a flat prior on Δ

$$p(\Delta|\boldsymbol{D}^{obs}) \infty p(\boldsymbol{D}^{obs}|\Delta) = \int p(\boldsymbol{D}|\Delta)\, dD^{miss}$$

This posterior helps to analyze the incomplete data. The mode of the posterior may be found by *expectation maximization* (EM) algorithm. The first algorithms used to impute the missing data are the *imputation-posterior* (IP), which represents a Markov chain, and the *expectation maximization importance sampling* (EMis) (Honaker and King, 2010). The implementation of these algorithms encounters many difficulties such as long run-times, software kinks,and not able to impute in the case of large data sets with high number of variables. For this reason, the combination of expectation maximization and bootstrapping (EMB) has been proposed as an alternative of IP and EMis. In fact, EMB algorithm gives the same results like IP and EMis and outstrips their problems (Honaker and King, 2010; Honaker et al., 2012).

4 Case study

Since 2000, the environment institutions in Tunisia continue to make much effort to establish a marine protected area (MPA) in Kuriat islands (Monastir, Tunisia) which threatened by both natural and anthropic pressures (elevation of sea level, marine and coastal erosion, overfishing, pollutions, etc.). Their MPA project highlights the environment effects but lays aside the economic interests of stakeholders which their professional and recreational activities related there (CPPA, 2000; CPPA and RAC/SPA, 2010). Among these stakeholders, we find nearly 40000 persons (residents and non residents) which come visiting for one day on average the Kuriat island for swimming and benefit from their virgin nature.

Once an MPA will be established, the visitors behavior may be change if the islands' access price will be increased for a possible means of funding the project. For this reason, the objective of our survey is to determine the willingness to pay (WTP) of visitors. Our survey consists of two main parts: socio-economic data and visitor's perceptions of MPA and WTP[9]. 315 direct interviews had been conducted randomly between july and august 2012 with visitors during their visit to Kuriat islands. 78.43% of the interviews have the age between 20 and 50 years and 54.3% of them are women. Almost all the interviews have a high educational school (secondary and university school). 44.76% among them had visited Kuriat island by means of local touristic agencies. The residents visitors in our sample represent 51.43% and the rest are foreign (21.59% are french, 7.94% are german, 3.81% are

[9] see appendix A

russian, 4.44% are belgium, 8.57% are italian and 2.22% are canadian).

	<20	[20-30[	[30-40[	[40-50[	[50-60[	≥ 60	Total
Women	2.54%	11.43%	22.23%	13.34%	1.90%	2.86%	54.30%
Man	3.80%	9.84%	9.84%	11.75%	8.57%	1.90%	45.70%
Total	6.34%	21.27%	32.07%	25.09%	10.47%	4.76%	100%

Table 1: *Distribution of interviewees by age and gender*

50.8% of the interviewees consider that the natural resources are overexploited especially the marine resources. About the question if the interviewee had listen of MPA before, 54.92% of them mentioned that they know what it means. Half of them are for the establishment of a MPA in Kuriat islands and they consider that a such project will contribute to protect their environment richness and improve their coastal and marine ecosystems. On the contrary, 21.91% of the interviewees refuse categorically the project and they display a null willingness to pay (WTP). The rest mention a WTP varies between 5 and 50 DT[10]. The WTP is as a payment card which had been fixed and chosen according to the preliminary interviews. To determine the WTP amount, a dichotomous choice question had been asked as follows:
Would you be willing to pay $5, 10, 15, 20, 30, 40, 50$ DT in additional per visit? (The interviewee chooses one of the payment card's amounts)[11]

	0 DT	5 DT	10 DT	15 DT	20 DT	30 DT	40 DT	50 DT
WTP	21.90%	3.81%	19.04%	17.14%	15.24%	10.48%	10.48%	1.91%

Table 2: *WTP of interviewees*

5 Results

As evoked above, 21.90% of the interviewees had mentioned a null WTP, which represent an excess zeros. Moreover, we highlight here the weakness of our sample which represent 0.7% of all the population (40000 visitors to Kuriat island annually). It my suffer of missing data and misrepresentation which can give a bias estimations. To sum up, we encounter main difficulty: excess zero. The choice of ZIOP model is to resolve the excess zero and the use of multiple imputation[12] to extend our sample size for reaching 500 in place of 315. To carry out this objective, we use the official statistics and informations on the tourism in Tunisia

[10]Tunisian Dinar, local currency; 1DT≈0.5 euro

[11]The question for determining the WTP has been inspired of Hall et al's (2002) study(Hall et al., 2002)

[12]The sample is augmented by new artificial records using various hypotheses and priors. In spite of its known limits, our MI benchmark algorithm will be $AMELIA$ (Honaker et al., 2012) because informative Bayesian priors about individual missing data cells can be included. $AMELIA$ has also an advantage in its flexibility and its efficiency. The incorporation of priors follows basic Bayesian analysis where the imputation turns out to be a weighted average of the model-based imputation and the prior mean, with weights depending of the relative strength of the data and prior (Honaker and King, 2010). These informative prior can come from the elicitation of expert belief or from the analysis of previous studies (Garthwaite et al., 2005). There are helpful for handling sparse data (Lenk and Orme, 2009), which can appear if the sample is strongly biased, as more additional artificial records are needed.

about the nationalities of the population for generating 185 additional articial record. The ordered dependent variable, WTP, has two types of zeros. The first type refers to economic conditions of the interviewee who is not able to pay for an environment improvement. The second type represents the false zero; the interviewee refuse to pay because he has a strategic behavior (the improvement will be paid by the other, there is difficult to understand the hypothetical survey...). Compared to a standard OP, the ZIOP model allows to get relevant estimates which have a huge coverage of the probabilities and they have a low level of bias (Hill et al., 2011).

To estimate the conditional OP, ZIOP and ZIOPC, we used the R program and code which had been developed by Bagozzi et al (2011) (Hill et al., 2011). The covariates which have been incorporated in standard OP, have been selected by stepwise method. Those have been included in OP, are the same in ZIOP(C). We chose the same covariates for the splitting and ordered equations. All maximum likelihood estimated models must begin with an initial set of starting parameters. For our application, the goal is to provide a naive but reasonable set of starting parameter values, where the term "reasonable", which introduced by Bagozzi et al (2011), typically means values between -1 and 1, so as not to provide too large an initial starting parameter value. Seeing that the ZIOP model is complicated, it can lead to non-convergence or non-finite initial values which case represents an issue especially in the case of our sample. In which case, it is recommended to estimate a simpler model (the standard OP in our case), and then use the OP estimates as starting parameter values (in ordered equation) for the more complicated model (the ZIOP in our case), and then "reasonable" values (we used 0.1) for any additional parameters estimated for the covariates in splitting equation.

The stepwise method has chosen seven significant covariates; gender, age, matrimonial situation, education level, income, how the visit organized and the residency of the interviewees. We have excluded the covairates which linked to environment and MPA perceptions to sidestep an endogeneity problem.Table 4 mentions the results of OP, ZIOP and ZIOPC models. For the three models, the income, education level and the residency are significant in ordered equation. They explain well what influences the WTP and determines its amount. It is quite normal since the literature has shown the importance of income in the contingent labor. The age is significant in the splitting equation for ZIOP(C) models. The ZIOPC shows the non significance of covariates gender and marital situation in the splitting equation and the manner of visit's organization in ordered equation. Referring to the information criterion, the ZIOP and ZIOPC models have been chosen such as a relevant models than the OP model. To same up, the acceptability of the MPA project may be explained by the age and the education level of the interviewees. Thus has been proved by the ZIOP(C) models. Furthermore, the three models (OP, ZIOP, and ZIOPC) choose the same covariates which determine the WTP amounts. These covariates are the income, education level and residency of the interviewees.

Harris and Zhao (2007) have proposed the Vuong test (Vuong, 1989) to compare the ZIOP and ZIOPC models to the OP model. The Vuong test described by the equation

$$v = \frac{\sqrt{N}(\frac{1}{N}\sum_i^N m_i)}{\sqrt{\frac{1}{N}\sum_i^N (m_i - \overline{m})^2}}$$

with m_i the natural logarithm of the ratio of the predicted probability that $\lambda_i = j$ of OP model (in the numerator) and the ZIOP(C) model in the denominator. where $v < -1.96$; ZIOP or ZIOPC model is favor, $-1.96 < v < 1.96$ means no support any model, and $v > 1.96$ favors the OP model (Vuong, 1989; Hill et al., 2011; Harris, 2007). The vuong test results favor the ZIOP model over the OP model ($v = -2.872$) and similarly favors ZIOPC model over the OP model ($v = -3.270$). These results show that ZIOPC and ZIOP model are superior to the OP model, which result had been mentioned by Harris and Zhao (2003, 2007), and Bagozzi et al (2011). The zeros in our sample represent 21.9%, and the predicted probability of the ZIOP mentioned that 29% of null wtp with 54.4% protest response or structural 0.

For an average individual who would be willing to pay, the WTP will be 41.8 DT (conditional WTP). Besides that, for an average individual whom we know nothing a priori, the WTP will be 35.036 DT (41.8*0.83).
Reverting to the results of ZIOP(C) models, who have a high education level among the interviewees, are more able to accept the MPA project than the other, and pay an additional amount to benefit of the environment change in the two islands. Furthermore, the residency of the interviewees, plays an important role for the accepting of the project and the WTP declared. In fact, the interviewees are residents (Tunisians) and non residents (foreign). The foreign are more able to accept the MPA project and to declare a high willingness to pay than the residents. Thus may be a kind of source of bias and a part of them proclaim a positive WTP due to that they are on holiday and visit the Kuriat islands once and it's not safe to come back another time. On the other hand, the non residents have a high income level than the resident and their culture allows them to be more involved in the environment conservation projects. Thence, we can evoke that the contingent valuation surveys more reliable in the developed countries than in the developing countries, which result had been proved by the literature (Whittington, 2002).

Table 3: *Estimated Coefficients for OP, ZIOP and ZIOPC*

	OP	ZIOP	ZIOPC
Splitting Parameters			
INTERCEPT		−1.235 (0.397)***	−1.240 (0.376)***
GENDER1		0.146 (0.196)	0.259 (0.206)
AGE2		1.506 (0.406)***	1.232 (0.425)***
AGE3		2.324 (0.447)***	2.153 (0.446)***
AGE4		2.197 (0.480)***	1.922 (0.521)***
AGE5		1.465 (0.476)***	1.268 (0.476)***
AGE6		1.619 (0.556)***	1.454 (0.532)***
MS1		−0.178 (0.257)	0.017 (0.300)
EL1		0.756 (0.191)***	0.845 (0.197)***
Ordered Parameters			
0|1	3.077 (0.561)***	0.517 (0.603)	−0.021 (0.637)
1|2	−1.738 (0.280)***	1.561 (0.202)***	1.369 (0.300)
2|3	−0.246 (0.116)**	0.552 (0.119)***	0.498 (0.139)***
3|4	−0.473 (0.124)***	−0.185 (0.121)	−0.247 (0.139)*
4|5	−0.549 (0.132)***	−0.385 (0.129)***	−0.460 (0.157)***
5|6	−0.636 (0.161)***	−0.506 (0.160)***	−0.582 (0.184)***
GENDER1	0.217 (0.134)*		
AGE2	0.799 (0.386)**		
AGE3	1.448 (0.423)***		
AGE4	1.392 (0.438)***		
AGE5	0.951 (0.445)**		
AGE6	0.959 (0.493)*		
MS1	0.087 (0.174)		
EL1	0.794 (0.134)***	0.668 (0.150)***	0.459 (0.201)***
OV1	−0.033 (0.242)	−0.000 (0.243)	−0.021 (0.228)
NAT1	0.551 (0.254)**	0.627 (0.255)***	0.605 (0.239)***
lnINCOME	0.236 (0.056)***	0.719 (0.090)***	0.597 (0.140)***
ρ			−0.645 (0.331)**
AIC	1016.019	964.928	965.046
BIC	1079.813	1036.227	1040.098
IBIC	1151.609	1112.686	1120.810
CAIC	1096.813	1055.227	1060.098

Standard errors are in parentheses.***, ** and * indicate significance at 1%, 5% and 10% sizes, respectively.

As we evoked above, the MI has been used like a sensitive analysis for our estimation, especially with and extension of the sample which reached 500. Using the imputed data, we have obtained approximately the some results (Tab.5) for OP and ZIOP(C) models with a sleazy difference. For the OP model, the covariate "gender" becomes significant. The manner of visit's organization (OV1) has been selected in the ordered parameters for the ZIOP model. For the ZIOPC model, the covariates "gender" and the visit's organization become significant, but the education level has been neglected. According the information criterion, the BIC, IBIC and CAIC have selected the OP model, besides the AIC has chosen the ZIOPC. The vuong test results favor the ZIOP ($v = -2.623$) and ZIOPC ($v = -3.662$) models over the OP model. For the sample with imputed data, the zeros represent 21%, and the predicted probability of the ZIOP mentioned that 49% of null WTP with 58.7% protest response or structural 0. With the imputed data, the WTP of an average individual will be 28.4 DT (conditional WTP). Moreover, for an average individual whom we know nothing a priori, the WTP will be 20.16 DT (28.4*0.71).

Table 4: *Estimated Coefficients for OP, ZIOP and ZIOPC with imputed data*

	OP	ZIOP	ZIOPC
Splitting Parameters			
INTERCEPT		−0.486 (0.293)*	−0.537 (0.252)**
GENDER1		0.249 (0.189)	0.327 (0.152)**
AGE2		0.953 (0.281)***	0.863 (0.225)***
AGE3		1.786 (0.375)***	1.620 (0.282)***
AGE4		1.564 (0.343)***	1.459 (0.276)***
AGE5		1.145 (0.335)***	1.152 (0.262)***
AGE6		1.271 (0.426)***	1.203 (0.335)***
MS1		−0.139 (0.208)	0.019 (0.162)
EL1		0.617 (0.196)***	0.576 (0.205)***
Ordered Parameters			
0|1	1.233 (0.273)***	−0.763 (0.314)***	−0.844 (0.305)***
1|2	−1.332 (0.162)***	−0.688 (0.248)***	−0.773 (0.237)***
2|3	−0.515 (0.094)**	−0.245 (0.115)**	0.325 (0.111)***
3|4	−0.699 (0.102)***	−0.595 (0.105)***	−0.694 (0.104)***
4|5	−0.744 (0.110)***	−0.686 (0.110)***	−0.820 (0.114)***
5|6	−0.785 (0.131)***	−0.744 (0.130)***	−0.912 (0.137)***
GENDER1	0.173 (0.098)*		
AGE2	0.706 (0.215)***		
AGE3	1.300 (0.215)***		
AGE4	1.232 (0.219)***		
AGE5	1.020 (0.237)***		
AGE6	1.192 (0.283)***		
MS1	0.025 (0.112)		
EL1	0.536 (0.101)***	0.416 (0.120)***	0.201 (0.139)
OV1	−0.053 (0.117)	−0.311 (0.130)**	−0.221 (0.109)**
NAT1	0.479 (0.118)***	0.593 (0.132)***	0.515 (0.109)***
lnINCOME	0.053 (0.014)***	0.047 (0.016)***	0.054 (0.014)***
ρ			−0.879 (0.068)***
AIC	1780.294	1797.532	1779.469
BIC	1851.942	1877.609	1863.761
IBIC	1937.042	1960.624	1958.193
CAIC	1868.942	1896.609	1883.761

Standard errors are in parentheses.***, ** and * indicate significance at 1%, 5% and 10% sizes, respectively.

6 Conclusion

The contingent survey suffers from several bias especially the protest responses like in our sample. The ZIOP model seems able to treat this kind of issue; excessive zeros in the datasets ($\approx 21.9\%$). It uses a system of two latent equations that allows for the zero observations to be generated by two different attitudinal regimes. Also, Harris and Zhao (2007) have given an importance to the correlation between the two latent equations. For this reason, they have introduced the ZIOPC model which, applying the vuong test, has more performance than the OP and ZIOP models (same results with imputed data). The ZIOP(C) model mention that the acceptability (in splitting equation)of the environment change (MPA establishment) depends on attitude and behavioral criteria such as the age, education level, gender, etc. Moreover, the ordered dependent variable (WTP) may better explained by the economic criteria such as the income, residency and education level too. The ZIOP mentioned that 29% (49% with imputed data) of null WTP with 54,4% (58.7%with imputed data) structural 0.

By this contingent survey, we tried to highlight the role of the contingent valuation for measuring the WTP for an environmental project (MPA) and to detect which criteria influences the interviewees declarations. The contingent valuation surveys conducted in the developing countries had been widely criticized. For instance, Whittington (2002) criticized them for three reasons; (i) the CV surveys are "poorly executed and implemented", the CV scenarios had "poorly crafted" and the failure to conduct a robustness test such as using the "split-sample experiments" (Whittington, 2002). On the other side, some practitioners had mentioned the reliability of the CV surveys in the developing countries one time the settlements, which had been suggested by the literature, are followed. At this level, we evoke the reliability of CV survey which had been conducted by Memon and Matsuoka (2002)in a developing country (Pakistan). They shown that their CV survey had been conducted successfully because it had been well executed, its scenario had been lavishly crafted and the sensibility test had been done without difficult (Memon and Matsuoka, 2002). In our case, the CV survey had been well implemented and its scenario had been well explained to the interviewees, but it does not stop to have a biased sampling. The sensitivity analysis has been done using the sample with imputed data. We have obtained approximately the same results in comparison with initial sample. The results of the MI strengthen our analysis especially the CV method has been largely criticized.

References

John Aitchison. On the distribution of a positive random variable having a discrete probability mass at the origin*. *Journal of the american statistical association*, 50(271):901–908, 1955.

Paul D Allison. *Logistic regression using SAS: Theory and application.* SAS Institute, 2012.

Henrik Andersson and Mikael Svensson. Cognitive ability and scale bias in the contingent valuation method. *Environmental and Resource Economics*, 39(4):481–495, 2008.

Dan Ariely, George Loewenstein, and Drazen Prelec. "coherent arbitrariness": Stable demand curves without stable preferences*. *Technology*, 73, 2003.

Tijen Arin and Randall A Kramer. Divers' willingness to pay to visit marine sanctuaries: an exploratory study. *Ocean & Coastal Management*, 45(2):171–183, 2002.

John Asafu-Adjaye and Sorada Tapsuwan. A contingent valuation study of scuba diving benefits: Case study in mu ko similan marine national park, thailand. *Tourism Management*, 29(6):1122–1130, 2008.

Giles Atkinson, Sian Morse-Jones, Susana Mourato, and Allan Provins. 'when to take "no" for an answer'? using entreaties to reduce protests in contingent valuation studies. *Environmental and Resource Economics*, 51(4):497–523, 2012.

B.E. Bagozzi, D.H. Hill, W.H. Moore, and B. Mukherjee. Modeling two types of peace: The zero-inflated ordered probit (ziop) model in conflict research. *Journal of Conflict Resolution*, page forthcoming, 2015.

Nabin Baral, Marc J Stern, and Ranju Bhattarai. Contingent valuation of ecotourism in annapurna conservation area, nepal: Implications for sustainable park finance and local development. *Ecological Economics*, 66(2):218–227, 2008.

Ian J Bateman, Diane Burgess, W George Hutchinson, and David I Matthews. Learning design contingent valuation (ldcv): Noaa guidelines, preference learning and coherent arbitrariness. *Journal of environmental economics and management*, 55(2):127–141, 2008.

Kenneth N Berk. Repeated measures with zeros. *Statistical methods in medical research*, 11 (4):303–316, 2002.

Mahadev G Bhat. Application of non-market valuation to the florida keys marine reserve management. *Journal of Environmental Management*, 67(4):315–325, 2003.

Karen Blumenschein, Glenn C Blomquist, Magnus Johannesson, Nancy Horn, and Patricia Freeman. Eliciting willingness to pay without bias: Evidence from a field experiment*. *The Economic Journal*, 118(525):114–137, 2008.

Richard T Carson, Nicholas E Flores, and Norman F Meade. Contingent valuation: controversies and evidence. *Environmental and resource economics*, 19(2):173–210, 2001.

J Clinch and Anthony Murphy. Modelling winners and losers in contingent valuation of public goods: appropriate welfare measures and econometric analysis. *The Economic Journal*, 111(470):420–443, 2001.

CPPA. Gestion des zones sensibles littorales : Les îles kuriat. Technical report, Coastal Protection and Planning Agency of Tunisia, 2000.

CPPA and RAC/SPA. Rapport de la mission d'étude des habitats marins et des principales espèces des îles kuriat (tunisie). Technical report, Coastal Protection and Planning Agency of Tunisia, Regional Activity Center for Specially Protected Areas, 2010.

Ronald G Cummings and Laura O Taylor. Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. *American Economic Review*, pages 649–665, 1999.

Gerard S Dharmaratne, Francine Yee Sang, and Leslie J Walling. Tourism potentials for financing protected areas. *Annals of Tourism Research*, 27(3):590–610, 2000.

Peter A Diamond and Jerry A Hausman. Contingent valuation: is some number better than no number? *The Journal of economic perspectives*, pages 45–64, 1994.

Naihua Duan, Willard G Manning, Carl N Morris, and Joseph P Newhouse. A comparison of alternative models for the demand for medical care. *Journal of business & economic statistics*, 1(2):115–126, 1983.

P.H. Garthwaite, J.B. Kadane, and A. O'Hagan. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, 100(470):690–701, 2005.

William Greene. *Functional form and heterogeneity in models for count data*. Now Publishers Inc, 2007.

Darwin C Hall, Jane V Hall, and Steven N Murray. Contingent valuation of marine protected areas: Southern california rocky intertidal ecosystems. *Natural Resource Modeling*, 15(3): 335–368, 2002.

W Michael Hanemann. Valuing the environment through contingent valuation. *The Journal of Economic Perspectives*, pages 19–43, 1994.

M.N. Harris. A zero-inflated ordered probit model, with application to modelling tobacco consumption. *Journal of Econometrics*, 141(2):1073–1099, 2007.

Daniel W Hill, Benjamin E Bagozzi, Will H Moore, and Bumba Mukherjee. Strategic incentives and modeling bias in ordinal data: the zero-inflated ordered probit (ziop) model in political science. In *new faces in political methodology meeting, Penn State*, volume 30, 2011.

J. Honaker and G. King. What to do about missing values in time-series cross-section data. *American Journal of Political Science*, 54(2):561–581, 2010.

J. Honaker, G. King, and M. Blackwell. *Amelia II: A Program for Missing Data*. http://gking.harvard.edu/amelia, 2012.

Svein Jentoft. Limits of governability: Institutional implications for fisheries and coastal governance. *Marine Policy*, 31(4):360–370, 2007.

Peter A Lachenbruch. Comparisons of two-part models with competitors. *Statistics in medicine*, 20(8):1215–1234, 2001.

Peter A Lachenbruch. Analysis of data with excess zeros. *Statistical methods in medical research*, 11(4):297–302, 2002.

Choong-Ki Lee and Sang-Yoel Han. Estimating the use and preservation values of national parks' tourism resources using a contingent valuation method. *Tourism management*, 23 (5):531–540, 2002.

P. Lenk and B. Orme. The value of informative priors in bayesian inference with sparse data. *Journal of Marketing Research*, 46:832–845, 2009.

Jayson L Lusk and Darren Hudson. Willingness-to-pay estimates and their relevance to agribusiness decision making. *Applied Economic Perspectives and Policy*, 26(2):152–169, 2004.

Mushtaq Ahmed Memon and Shunji Matsuoka. Validity of contingent valuation estimates from developing countries: scope sensitivity analysis. *Environmental economics and policy studies*, 5(1):39–61, 2002.

X.L. Meng. Multiple-imputation inferences with uncongenial sourecs of input(with discussion). *Statistical Science*, 10:538–573, 1994.

Jürgen Meyerhoff and Ulf Liebe. Do protest responses to a contingent valuation question and a choice experiment differ? *Environmental and Resource Economics*, 39(4):433–446, 2008.

Robert Cameron Mitchell. On designing constructed markets in valuation surveys. *Environmental and Resource Economics*, 22(1-2):297–321, 2002.

Robert Cameron Mitchell and Richard T Carson. *Using surveys to value public goods: the contingent valuation method*. Resources for the Future, 1989.

Lawrence H Moulton and Neal A Halsey. A mixture model with detection limits for regression analyses of antibody response to vaccine. *Biometrics*, pages 1570–1578, 1995.

Lawrence H Moulton and Neal A Halsey. A mixed gamma model for regression analyses of quantitative assay data. *Vaccine*, 14(12):1154–1158, 1996.

Douglas S Noonan. Contingent valuation and cultural resources: a meta-analytic review of the literature. *Journal of cultural economics*, 27(3-4):159–176, 2003.

I Petrosillo, G Zurlini, ME Corliano, N Zaccarelli, and M Dadamo. Tourist perception of recreational environment and management in a marine protected area. *Landscape and Urban Planning*, 79(1):29–37, 2007.

Ronald S Pimentel, Magdalena Niewiadomska-Bugaj, and Jung-Chao Wang. Association of zero-inflated continuous variables. *Statistics & Probability Letters*, 96:61–67, 2015.

Ussif Rashid Sumaila, Sylvie Guénette, Jackie Alder, and Ratana Chuenpagdee. Addressing ecosystem effects of fishing using marine protected areas. *ICES Journal of Marine Science: Journal du Conseil*, 57(3):752–760, 2000.

Douglas J Taylor, Lawrence L Kupper, Stephen M Rappaport, and Robert H Lyles. A mixture model for occupational exposure mean testing with a limit of detection. *Biometrics*, 57(3):681–688, 2001.

S. Van Buuren. *Flexible Imputation of Missing Data.* CRC Press, Boca Raton, 2012.

Quang H Vuong. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, pages 307–333, 1989.

Andrew A Weiss. A bivariate ordered probit model with truncation: Helmet use and motorcycle injuries. *Applied statistics*, pages 487–499, 1993.

Dale Whittington. Improving the performance of contingent valuation studies in developing countries. *Environmental and resource economics*, 22(1-2):323–367, 2002.

Zhou Xiao-Hua and Wanzhu Tu. Comparison of several independent population means when their samples contain log-normal and possibly zero observations. *Biometrics*, 55(2): 645–651, 1999.

Haiyi Xie, Jill Tao, Gregory J McHugo, and Robert E Drake. Comparing statistical methods for analyzing skewed longitudinal count data with many zeros: An example of smoking cessation. *Journal of substance abuse treatment*, 45(1):99–108, 2013.

Huifang Zhang, R Karina Gallardo, Jill J McCluskey, and Eugene M Kupferman. Consumers' willingness to pay for treatment-induced quality attributes in anjou pears. *Journal of Agricultural and Resource Economics*, pages 105–117, 2010.

A

Variables signification included in the survey

Variables	Label	Description	n	Min	Mean	Max	Std. dev.
Part 1: socioeconomic data							
Gender		1= Man	144				
		0= Women	171				
NT	Nationality	0= Tunisian	162				
		1= French	68				
		2= German	25				
		3= Russian	12				
		4= Belgium	14				
		5= Italian	27				
		6= Canadian	7				
Age		1= <20 years	20				
		2= [20 - 30[	67				
		3= [30 - 40[	101				
		4= [40 - 50[	79				
		5= [50 - 60[	33				
		6= $\geq$ 60	15				
MS	Marital Status	0=Single	107				
		1= Not single	208				
EL	Education Level	0= primary and secondary school	140				
		1= University	175				
lnINCOME	natural logarithm of income			0	9.178	11.156	2.171
INCOME	Income in dummies	1= <5000DT*	54				
		2= [5000 - 10000[	56				
		3= [10000 - 15000[	29				
		4= [15000 - 20000[	20				
		5= [20000 - 30000[	60				
		6= [30000 - 40000[	42				
		7= [40000 - 60000[	45				
		8= $\geq$ 60000	9				
MPA perception and WTP declared							
OV	If the visit is organized	0= by agency	141				
		0= by self	174				
NRU	Nature resources Usage	0= underexploited	38				
		1= operating under standards	117				
		2= overexploited	160				
THMPA	If the Kuriat islands are threatened	0= no	155				
		1= yes	160				
KMPA	Knowing MPA before	0= no	142				
		1= yes	173				
TMPA	If the MPA will attract more visitors	0= no	85				
		1= yes	230				
NAT	The interviwee is resident or not	0= resident	162				
		1= no resident	153				
WTP	Willigness to pay	0= 0 DT	69				
		1= 5 DT	12				
		2= 10 DT	60				
		3= 15 DT	54				
		4= 20 DT	48				
		5= 30 DT	33				
		6>= 40 DT	39				

$*$=1 DT$\approx$ 0.5$euro$

Regressors' signification

B

Standard OP model

	Value	Std. Error	t value
0|1	3.077	0.560	5.487
1|2	3.253	0.562	5.781
2|3	4.035	0.574	7.024
3|4	4.658	0.582	8.000
4|5	5.235	0.588	8.895
5|6	5.764	0.595	9.676
GENDER1	0.217	0.133	1.624
AGE2	0.799	0.385	2.073
AGE3	1.447	0.423	3.420
AGE4	1.392	0.438	3.175
AGE5	0.951	0.445	2.135
AGE6	0.959	0.493	1.945
MS1	0.087	0.174	0.500
EL1	0.793	0.134	5.918
OV1	-0.033	0.242	-0.137
NAT1	0.550	0.253	2.170
lnINCOME	0.235	0.055	4.218
Residual Deviance	982.019		
AIC	1016.019		

C

OP results

Variables	estimate	sdbeta	t value
0|1	3.077	0.561	5.486
1|2	−1.738	0.280	−6.197
2|3	−0.246	0.116	−2.125
3|4	−0.473	0.124	−3.828
4|5	−0.549	0.132	−4.175
5|6	−0.636	0.161	−3.944
GENRE1	0.217	0.134	1.624
AGE2	0.799	0.386	2.074
AGE3	1.448	0.423	3.421
AGE4	1.392	0.438	3.175
AGE5	0.951	0.445	2.136
AGE6	0.959	0.493	1.946
MS1	0.087	0.174	0.500
EL1	0.794	0.134	5.918
OV1	−0.033	0.242	−0.138
NAT1	0.551	0.254	2.170
lnINCOME	0.236	0.056	4.218

D

ZIOP results

variable	β	sdbeta	t test
INTERCEPT	-1.235	0.397	-3.111
GENDER1	0.146	0.196	0.742
AGE2	1.506	0.406	3.705
AGE3	2.324	0.447	5.199
AGE4	2.197	0.480	4.580
AGE5	1.465	0.476	3.077
AGE6	1.619	0.556	2.910
MS1	-0.178	0.257	-0.692
EL1	0.756	0.191	3.948
0|1	0.517	0.603	0.858
1|2	1.561	0.202	7.726
2|3	0.552	0.119	4.630
3|4	-0.185	0.121	-1.527
4|5	-0.385	0.129	-2.976
5|6	-0.506	0.160	-3.171
EL1	0.668	0.150	4.458
OV1	0.000	0.243	0.001
NAT1	0.627	0.255	2.456
lnINCOME	0.719	0.090	8.026

E

ZIOPC results

variable	β	sdbeta	t test
INTERCEPT	−1.240	0.376	−3.295
GENDER1	0.259	0.206	1.257
AGE2	1.232	0.425	2.901
AGE3	2.153	0.446	4.832
AGE4	1.922	0.521	3.692
AGE5	1.268	0.476	2.663
AGE6	1.454	0.532	2.732
MS1	0.017	0.300	0.056
EL1	0.845	0.197	4.278
0|1	−0.021	0.637	−0.033
1|2	1.369	0.300	4.558
2|3	0.498	0.139	3.569
3|4	−0.247	0.139	−1.772
4|5	−0.460	0.157	−2.931
5|6	−0.582	0.184	−3.157
EL1	0.459	0.201	2.287
OV1	−0.021	0.228	−0.091
NAT1	0.605	0.239	2.526
lnINCOME	0.597	0.140	4.262
ρ	−0.645	0.331	−1.949

Information criterion	OP	ZIOP	ZIOPC
AIC	1016.019	964.928	965.046
BIC	1079.813	1036.227	1040.098
IBIC	1151.609	1112.686	1120.810
CAIC	1096.813	1055.227	1060.098

F

OP results with imputed data

variable	β	sdbeta	t test
0|1	1.233	0.273	4.524
1|2	−1.332	0.162	−8.238
2|3	−0.515	0.094	−5.460
3|4	−0.699	0.102	−6.878
4|5	−0.744	0.110	−6.780
5|6	−0.785	0.131	−6.000
GENDER	0.173	0.098	1.769
AGE2	0.706	0.215	3.279
AGE3	1.300	0.215	6.053
AGE4	1.232	0.219	5.622
AGE5	1.020	0.237	4.295
AGE6	1.192	0.283	4.206
MS1	0.025	0.112	0.223
EL1	0.536	0.101	5.299
lnINCOME	0.053	0.014	3.695
OV1	−0.053	0.117	−0.454
NAT1	0.479	0.118	4.053

G

ZIOP results with imputed data

variable	β	sdbeta	t test
INTERCEPT	−0.486	0.293	−1.662
GENDER	0.249	0.189	1.313
AGE2	0.953	0.281	3.387
AGE3	1.786	0.375	4.759
AGE4	1.564	0.343	4.553
AGE5	1.145	0.335	3.417
AGE6	1.271	0.426	2.983
MS1	−0.139	0.208	−0.666
EL1	0.617	0.196	3.154
0|1	−0.763	0.314	−2.434
1|2	−0.688	0.248	−2.781
2|3	−0.245	0.115	−2.130
3|4	−0.595	0.105	−5.697
4|5	−0.686	0.110	−6.225
5|6	−0.744	0.130	−5.706
EL1	0.416	0.120	3.457
lnINCOME	0.047	0.016	2.854
OV1	−0.311	0.130	−2.389
NAT1	0.593	0.132	4.502

H

ZIOPC results with imputed data

variable	β	sdbeta	t test
INTERCEPT	-0.537	0.252	-2.131
GENDER	0.327	0.152	2.157
AGE2	0.863	0.225	3.826
AGE3	1.620	0.282	5.735
AGE4	1.459	0.276	5.282
AGE5	1.152	0.262	4.399
AGE6	1.203	0.335	3.591
MS1	0.019	0.162	0.116
EL1	0.576	0.205	2.811
0|1	-0.844	0.305	-2.769
1|2	-0.773	0.237	-3.258
2|3	-0.325	0.111	-2.934
3|4	-0.694	0.104	-6.649
4|5	-0.820	0.114	-7.210
5|6	-0.912	0.137	-6.644
EL1	0.201	0.139	1.441
lnINCOME	0.054	0.014	3.796
OV1	-0.221	0.109	-2.021
NAT1	0.515	0.109	4.742
ρ	-0.879	0.068	-13.014

Information criterion	OP	ZIOP	ZIOPC
AIC	1780.294	1797.532	1779.469
BIC	1851.942	1877.609	1863.761
IBIC	1937.042	1960.624	1958.193
CAIC	1868.942	1896.609	1883.761

Information criterion for the imputed data