Selected Poster prepared for presentation at the 2016 Agricultural & Applied Economics Association Annual Meeting, Boston, MA, July 31 – Aug. 2.

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

Marouene Mbarek, Phd Student, AGROCAMPUS OUEST, Angers, France

Damien Rousselière, Professor, AGROCAMPUS OUEST, Angers, France

Julien Salanié, Associate Professor, University of Saint-Etienne, Saint-Etienne, France

Objectives of Contingent Valuation (CV)

- Degradation of renewable natural resources → establishment conservation policies
- Valuation of conservation policies → Success ⇌ Failure → Help with the decision
- Various means for measuring non-market goods or public goods : direct (Contingent valuation) or non-direct methods (revealed preferences).
- CV : Estimation of the Willingness to pay for an environmental public goods using surveys
- Objectives for public policies :
  - Providing an aggregate economic estimates.
  - Identify the variables influencing the WTP

Issues in contingent valuation

Two main critics (See discussion in a 2012 issue of the *Journal of Economic Perspectives*)

- Issues in sampling design (Hausmann 1994, 2012) : small samples, biased samples...
- Taking into account the heterogeneity of the population (especially the opponents for an environmental policy)

Current researches

- Good practices for survey design
- Bayesian econometric models

Alternative to solve these problems :

- Multiple Imputation (MI) (Schafer 1999, Buuren 2012)
- Zero Inflated Ordered Probit (ZIOP) model (Harris and Zhao (2007))

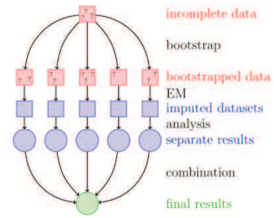
Missing data and biased sampling

Proposition based on Van Buuren (2012, chapter 8) : **Correcting for non-response** and using MI to make a sample representative.
Thinking in term of **potential outcomes** (Rubin 2004) : What if non-sampled peoples have been sampled ? What would have they answered ?
Generation of additional observations (hypothetical respondents) based on auxiliary data
Assumption that the bias depends on known covariates which are fully observable (MAR assumption). Assumptions are not different from those which are usually necessary for the classical method (post-stratification or weighting to known population totals)

- Advantages according to Van Buuren (2012) "Imputation provides fine-grained control over the correction process"
- MI can help us therefore to introduce **sensitivity analysis** (Leamer 1985)

Benchmark Multiple Imputation Algorithm : Amelia II

High efficient algorithm **EMB (Expectation-Maximization with Bootstrapping)** (Honaker and King 2010)



Benchmark Multiple Imputation Algorithm : Amelia II

Divide the data matrix D into an observed and a missing part, with $D = \{D^{obs}, D^{mis}\}$. D is assume to be multivariate normal $D \hookrightarrow N(\mu, \Sigma)$ with mean μ and variance Σ.

Widely use in social sciences (political sciences, economics, sociology...)

Simplicity of sensitivity analysis : incorporations of priors following basic Bayesian Analysis (ridge prior, bounds or logical priors, elicited priors) : imputation is simply a weighted average of model-based imputation and the prior mean (Honaker and King 2010)

Known limits (as it is based on multivariate normal assumption) (Kropko et al. 2014, Audigier et al. 2015)
Note : By transforming variables, this method can also be applied to non-linearly distributed data

The econometric model : ZIOP

ZIOP (Zero-Inflated Ordered Probit) : **double-hurdle combination of a split probit model and an ordered probit model**
Extension to ZIOPC which assumes that the two errors terms are correlated.
These models address the problem of **two distinct data generating processes for the zeros**.

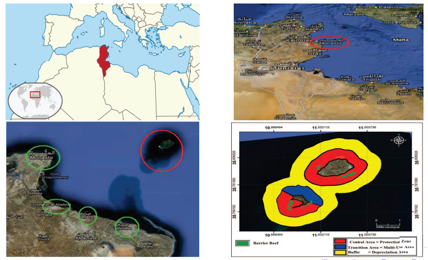
- One type corresponds to individuals who will always refuse to pay for MPA (because of non economic reasons) (inflation stage)
- The other type refers to a corner solution (outcome stage).

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

Application : tobacco use, environmental or medical policies

Case Study : Marine Protected Area Project for the Kuriat Islands in Monastir (Tunisia)

Figure: Protected and multiusages areas



Data

Our questionnaire consists of two main parts : socio-economic data and visitor's perceptions of environment and MPA.
315 direct interviews had been conducted randomly between july and august 2012 with visitors during their visit to Kuriat islands.

Table: Distribution of nationality

	Tunisian	French	German	Russian	Belgian	Italian	Canadian
Observed sample	**51.4%**	21.6%	7.9%	3.8%	4.4%	**8.6%**	2.2%
Official Statistics	46.8%	19.8%	**10.2%**	**13.0%**	3.4%	5.4%	1.4%

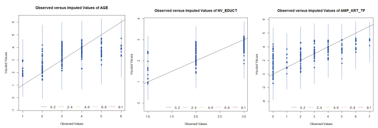
Table: *WTP of respondents*

	0 DT	5 DT	10 DT	15 DT	20 DT	30 DT	$\geq 40DT$
WTP	21.90%	3.81%	19.04%	17.14%	15.24%	10.48%	12.39

Language effect (interviews in French, Arabic and English) ?

Imputation

Generation of 185 additional observations
Priors : Ridge prior =0.01, Logarithmic distribution for Income (bounded by 0 and long tail)
Congenial imputation model : attitudes variables included (and excluded from the econometric model)
Overimputation test



First Results ZIOP

Table: Results for uncorrected and corrected samples

	Uncorrected Sample				Corrected sample			
	Inflation stage		outcome stage		Inflation stage		outcome stage	
Gender	-1.235				0.249			
AGE_20-30y	1.506	***			0.953	***		
AGE_30-40y	2.324	***			1.786	***		
AGE_40-50y	2,197	***			1.564	***		
AGE_50-60y	1,465	***			1.145	***		
AGE_60y	1,619	***			1.271	***		
MS_Not Married	-0.178				-0.139			
Organized			0				-0.311	***
Nationality			0,627	***			0,593	***
Educ_Univ	0.756	***	0.668	***	0.617	***	0,416	***
Income			0,719	***			0,047	***
Intercept	-1.235	***			-0.486	*		
0|1			0,517	***			-0.763	***
1|2			1,561	***			-0.688	***
2|3			0.552	***			-0.245	**
3|4			-0.185				-0.595	***
4|5			-0.385	***			-0.686	***
5|6			-0.506	***			-0.744	***

Economic estimations

Importance difference on the estimation of WTP
Uncorrected sample

- Estimation of 29% (21%) of zeros with 54.5% of structural zeros (inflation stage)
- Average WTP = 35.036 TD

Corrected sample

- Estimation of 49% (21%) of zeros with 58.7% of structural zeros (inflation stage)
- Average WTP = 20.16 TD (a decrease of 42.36% !)

Conclusion

- ZIOP : tool that can process data from contingent valuation ; the case of excess zeros in general (heterogeneity problem)
- The inflation and outcome stages are not explained by the same covariates
- The MI (sensitive analysis) we may allow to overtake the missing data problem and improve estimation results