AJAE appendix for Modeling Starting Point Bias as Unobserved Heterogeneity in Contingent Valuation Surveys: an Application to Air Pollution

Frédéric Aprahamian   Olivier Chanel   Stéphane Luchini

June 2006

The material contained herein is supplementary to the article named in the title and published in the American Journal of Agricultural Economics (AJAE).

Misspecification biases with censoring

In this appendix, we perform a set of Monte-Carlo simulations which account for censoring of the WTP distribution:

1. We consider that respondent i's willingness to pay is generated such that

$$W_{i1}^* = 200X_{i1} + 200X_{i2} + 200X_{i3} + u_i \tag{1}$$

   where X_1 is a constant term, X_2 is a continuous variable generated using a normal distribution with mean 0.5 and variance 4, X_3 is a dichotomous variable generated from a Bernoulli distribution with values 0 and 1 with probability 0.5 and u are the error terms normally distributed with zero mean and standard deviation $\sigma_0 = 50$. The variance of W_{i1}^* is given by

$$V_0(W_{i1}^*) = 200^2 V(X_{i2}) + 200^2 V(X_{i3}) + V(u_i) = 172500 \tag{2}$$

   with standard-error 415.33.

2. 600 observations are drawn according to equation (1). When W_{i1}^* is negative, it is considered as a null WTP. According to parameter values, the censoring rate equals 16.17% and the unconditional mean WTP is given by

$$E_0((W_{i1}^*) = \Phi(400/415.33) \times 400 + 415.33 \times \phi(400/415.33) = 437.10 \tag{3}$$

3. The bid design is such that $b_i = 300$ for the observations $i = 1, \ldots, 200$, $b_i = 400$ for the observations $i = 201, \ldots, 400$ and $b_i = 500$ for the observations $i = 401, \ldots, 600$.

4. In keeping with the econometric models, respondent i's answer to the open-ended follow-up question is generated such that:

$$W_{i2} \begin{cases} (1-\gamma_i)W_{i1}^* + \gamma_i b_i & \text{if} \quad W_{i1}^* > 0 \\ 0 \quad \text{otherwise} \end{cases} \tag{4}$$

where γ_i is a random variable with values 0.2 and 0.8. Its probability distribution is conditionally defined with respect to X_3 as in the simulations without censoring.

We draw 500 artificial datasets of 600 observations. In a first step, we estimate by maximum likelihood the homogeneous anchoring model, and ignore the fact that anchoring is heterogeneous in the population as well as we take into account censoring. In a second step, we estimate the Bernoulli heterogeneous anchoring model used to generate the data taking account censoring. Table 1 presents the mean and the variance of the prior willingness to pay W_{i1}^* as well as the mean and the variance of parameters in equation (1) for the 500 simulations and the two models.

First note that the estimations of the homogeneous anchoring model when $p = 0.5$ lead to results that differ from the case without censoring. Although, the standard deviation of the error term is still highly biased (368.32 while the true value is 50), the parameter estimates associated with the covariates and the constant term in the WTP equation also are highly biased, and consequently also the mean WTP. This is not the case when one does not take into account censoring as shown in Table 1 in the article. Second, in keeping with the non censored case, the parameter estimates associated with the WTP equation covariates, the standard deviation $\hat{\sigma}$ and the mean WTP are also biased for all other values of p. According to these results, the consequences of a mistaken assumption of homogeneous anchoring become more accentuated if one censors the WTP distribution at zero.

Finally, the estimation of the heterogeneous anchoring model leads to unbiased parameter estimates and thus an unbiased mean WTP whatever the value of p as in the non censored simulations. Therefore, the heterogeneous anchoring model with censoring yields accurate parameter estimates when the true WTP $W_{i1}^{\star}$ is considered to be positive or null.

Table 1: Monte Carlo Simulations (Variance in Parentheses)

Case	Model		$E[\widehat{W_1^*}]$	X_1 parameter	X_2 parameter	X_3 parameter	Std. Dev. $\hat{\sigma}$
$p = 0.5$	Homogeneous	Mean	550.33	43.63	343.84	342.74	368.32
		Variance	(2018.44)	(1833.10)	(1714.43)	(2544.71)	(3072.59)
	Heterogeneous	Mean	437.45	200.25	199.96	199.86	49.77
		Variance	(228.96)	(15.89)	(2.07)	(22.96)	(2.89)
$p = 0.6$	Homogeneous	Mean	557.56	30.21	344.57	386.38	360.40
		Variance	(2390.10)	(2447.24)	(2023.77)	(3685.51)	(3542.68)
	Heterogeneous	Mean	437.53	200.37	199.97	199.80	49.83
		Variance	(229.46)	(15.28)	(2.04)	(21.74)	(2.77)
$p = 0.7$	Homogeneous	Mean	559.74	23.28	340.32	420.60	345.90
		Variance	(2032.45)	(2235.35)	(1698.81)	(3538.59)	(2857.97)
	Heterogeneous	Mean	437.51	200.36	199.97	199.76	49.82
		Variance	(229.96)	(15.02)	(2.02)	(21.34)	(2.78)
$p = 0.8$	Homogeneous	Mean	559.96	19.01	334.07	450.70	327.07
		Variance	(1936.93)	(2222.02)	(1512.64)	(4074.11)	(2481.91)
	Heterogeneous	Mean	437.49	200.32	199.97	199.79	49.82
		Variance	(230.14)	(14.75)	(2.01)	(20.88)	(2.69)
$p = 0.9$	Homogeneous	Mean	560.21	16.72	327.88	477.15	306.43
		Variance	(1825.52)	(2147.45)	(1352.05)	(4404.66)	(2083.63)
	Heterogeneous	Mean	437.49	200.30	199.97	199.83	49.81
		Variance	(229.35)	(14.27)	(2.01)	(20.72)	(2.65)
$p = 1$	Homogeneous	Mean	560.20	14.78	321.94	501.92	286.09
		Variance	(1450.28)	(1803.02)	(1135.36)	(4121.87)	(1519.31)
	Heterogeneous	Mean	437.30	200.08	200.05	200.12	49.85
		Variance	(249.38)	(13.76)	(1.94)	(22.53)	(2.59)