CONTRIBUTI DI RICERCA CRENOS

CRENoS
CENTRO RICERCHE
ECONOMICHE NORD SUD
Università di Cagliari
Università di Sassari

MODELING ELICITATION EFFECTS IN CONTINGENT VALUATION STUDIES: A MONTE CARLO ANALYSIS OF THE BIVARIATE APPROACH

Margarita Genius
Elisabetta Strazzera

WORKING PAPERS

CUEC

2005/02

Margarita Genius
Dept. of Economics, University of Crete, Greece
genius@ermis.soc.uoc.gr

Elisabetta Strazzera
DRES and CRENoS, University of Cagliari, Italy
strazzera@unica.it

MODELING ELICITATION EFFECTS IN CONTINGENT VALUATION STUDIES: A MONTE CARLO ANALYSIS OF THE BIVARIATE APPROACH

Abstract
A Monte Carlo analysis is conducted to assess the validity of the bivariate modeling approach for detection and correction of different forms of elicitation effects in Double Bound Contingent Valuation data. Alternative univariate and bivariate models are applied to several simulated data sets, each one characterized by a specific elicitation effect, and their performance is assessed using standard selection criteria. The bivariate models include the standard Bivariate Probit model, and an alternative specification, based on the Copula approach to multivariate modeling, which is shown to be useful in cases where the hypothesis of normality of the joint distribution is not supported by the data. It is found that the bivariate approach can effectively correct elicitation effects while maintaining an adequate level of efficiency in the estimation of the parameters of interest.

Keywords: Double Bound, Elicitation effects, Bivariate models, Probit, Joe Copula

January 2005

1. Introduction

In the conclusions of their extensive overview of the state of the art of the Contingent Valuation method, Carson, Flores and Meade (2001) remark that "… at this point in the development of CV, the key objective in terms of methodological development should shift to trying to determine how to reduce the cost of conducting CV studies while still maintaining most of the quality of the very best studies now being conducted." Since costs are mainly driven by the survey administration, the crucial question is how to obtain valid and reliable estimates for the population WTP from smaller samples than those employed in benchmark high quality CV studies. A substantial stream of research has been conducted in the last decade aimed at finding an optimal method for elicitation of WTP, apt to combine the two properties of unbiasedness and statistical efficiency.

There seems to be a fairly general agreement in the literature (one notable exception is Green, Jacowitz, Kahneman and McFadden, 1998) that the single bound method is valid in terms of incentive compatibility, i.e. incentive for the respondent to truthfully reveal her preferences. Unfortunately, the single bound method is inefficient in terms of information conveyed by the elicitation process, and small size surveys can be particularly affected by this problem. As it is well known after the seminal paper by Hanemann, Loomis and Kanninen (1991), follow-up questions help to improve efficiency of the estimates: the double bound procedure shows a dramatic increase in the precision of the estimates. The problem is that, as pointed out by Carson, Groves and Machina (1999), and discussed more thoroughly in the following of the paper, iteration of the elicitation question may lead respondents to misrepresent their true preferences.

Yet, the double bound elicitation method may still be a preferred choice if the statistical analysis could detect the presence of elicitation effects in a reliable manner; and if, after correction of the estimates, which takes into account such effects, the statistical efficiency of the method were preserved. A recent trend is to base the statistical analysis of double bound data on the use of bivariate probit models. This path of research was initiated by Cameron and Quiggin (1994), who adopt several competing specifications based on the bivariate probit, and compare them to the univariate double bound model to analyze CV data. The approach is further pursued by Alberini, Carson and Kanninen (1997), who model, using bivariate probit specifications, different behavioral patterns induced by the reiteration of the elicitation question after the first answer has been obtained. Such different behavioral hypotheses are then tested as competing models by means of standard specification tests.

In a neoclassical theoretical framework, such as that considered by Carson, Groves and Machina, elicitation effects are deemed to affect the response to the second bid offer, no matter if the response to the first bid had been positive or negative. A different approach has been taken by DeShazo (2002), who theorizes that respondents who reject the first bid tender have no incentive to misrepresent their valuation when facing the second bid, while the opposite holds for the others. For data affected by this form of distortion ("framing effects"), DeShazo proposes that only the single bound response should be used. If needed, information from the second elicitation question should be limited to the sub-sample of respondents who say "no" to the first price offer. In an application, DeShazo models the resulting censored data by means of a bivariate probit with sample selection.

The present paper is aimed at assessing the validity of the bivariate approach to modeling double bound data for detection and correction of different forms of elicitation effects. The analysis is based on Monte Carlo methods, involving generation of several simulated data sets, each one characterized by a specific form of elicitation problem. Alternative univariate and bivariate models are applied to the simulated data, and their performance is assessed using standard selection criteria. The univariate models considered in the present work are the single bound and the double bound estimators, plus a univariate censored model, applied to data characterized by DeShazo's framing effects. The bivariate models include the standard Bivariate Probit model, based upon the Bivariate Normal distribution, and the Bivariate Probit with sample selection proposed by DeShazo.

It is well known that the normality assumption for the distribution of WTP (or its logarithm) is often not supported by the data. In these cases use of a bivariate probit would result in biased estimates: we therefore extend our analysis to alternative bivariate models, namely Copula models, which are characterized by a great flexibility in the distributional shape of their marginals, and in their dependence structure. In particular, the Joe Copula, which is characterized by asymmetry and fat tails, is applied to the double bound data generated in our experiments, and its performance is compared to the models mentioned above.

The paper is organized as follows: in section 2 we examine the behavioral hypotheses underlying the decision to respond strategically to WTP questions. Section 3 presents the bivariate modeling approach to fit double bound data, which can be based on the conventional Bivariate Probit, or, alternatively, on Copula models, here introduced. Section 4 describes the experimental design of the Monte Carlo analysis: underlying behavioral assumptions, data construction, and statistical models applied to the simulated data; in section 5 we discuss results; section 6 concludes the paper.

2. **Behavioral hypotheses**

Carson, Groves and Machina (1999, henceforth CGM) set a theoretical framework to analyze the incentive properties of different formats for elicitation of WTP. The single bound protocol is deemed valid, as long as some conditions hold: the valuation procedure should be presented as a referendum, involving the furniture of a new public good with a coercive contingent payment (for example, a tax), or the choice between two new public goods. Moreover, the survey should be seen by respondents as *consequential*, i.e. having the potential effect of influencing the agency decisions.

Unfortunately, the incentive properties of the single bound are not shared by the double bound method. CGM consider several behavioral hypotheses why a follow-up question may produce elicitation effects. For example, if the initial bid is considered by the respondent as informative about the real cost of provision of the good, the second bid may induce uncertainty about the cost distribution. For risk averse agents this would result in a lower reservation price (because of the risk premium), so that the WTP elicited after the second question would be shifted downward[1].

The same consequence on the final WTP value could also be produced by a strategic type of reaction to the new bid tender: agents may consider the iterative procedure as a bargaining game, where it would be strategically optimal not to reveal a higher WTP than the second price offered.

A third behavioral hypothesis considered by CGM is that individuals take the offered bids as informative data on the true cost of provision of the good: if, as it often is the case in contingent valuation studies, the respondent does not possess some *a priori* information, the bid offered in the first elicitation question can be taken as conveying information on the price distribution of the good, and the second bid serves as a Bayesian update of that information. CGM observe that this could have either positive or negative effects on the underlying WTP distribution, according to the sign of the answer to the first elicitation question. This problem, often referred to as anchoring or starting point bias, is analyzed for applications to double bound data by Herriges and Shogren (1996), who propose an econometric model that incorporates a Bayesian updating mechanism to detect and correct anchoring bias. They show in an application that correcting for starting point bias reduces the difference between the double bound and the single bound estimator, both for punctual estimates and confidence intervals. Analogous results are obtained by McLeod and Bergland (1999) in two other applications of the updating model, and they conclude that "the increased precision in the estimated WTP by asking a follow-up question is not as large, or even non-existent, when Bayesian updating is accounted for in the estimation". This point of view seems largely subscribed by Whitehead (2002), but with an important qualification: the double bound model should be used, with appropriate corrections for anchoring or other kinds of bias, when the initial bids are poorly chosen, i.e. they do not represent adequately the population's WTP distribution. In such a case the double bound still allows relevant efficiency gains, even after controlling for elicitation effects.

[1] By the same argument, some doubts may be cast on the One and One Half bound method, recently proposed by Hanemann, Loomis and Signorello (2002), which confronts the individual with an initial bid distribution, rather than a single first bid.

The updating econometric models described above are specified as univariate distributions. Cameron and Quiggin (1994) argue that when the underlying WTP values from the two elicitation questions are different, the assumption of a unique distribution is unduly restrictive, and propose a bivariate modeling approach, based on the bivariate probit model, which is also applied by Alberini (1995) and Alberini, Carson and Kanninen (1997) to several data sets. In general, these papers empirically support the view that the two underlying WTP values obtained through the double bound method are not identical, and while for the most part the data seems unaffected by a systematic form of bias, some (Alaska study, in Alberini, Carson, Kanninen) show evidence of a downward shift in the WTP elicited through the follow-up question.

An alternative theoretical framework is proposed by DeShazo (2002). Building upon theoretical results from Prospect Theory (Kahneman and Tversky, 1979), he argues that strategic behavior may only emerge for respondents who answer "yes" to the first price offer, while incentives for the others would be unaltered across elicitation questions. The reason is that if the respondent's true WTP is above the first bid, she expects to gain some consumer surplus, which may be taken as a reference point. Conversely, no consumer's surplus is expected, and no reference point is created, by people whose WTP is below the first bid tender. Under the double bound protocol, a "yes" response to the first bid question leads to a higher second bid (ascending sequence), while the converse holds in the case of a negative answer (descending sequence). According to DeShazo's theory, ascending sequences are susceptible to strategic behavior, induced by the creation of a reference point (framing effect), which does not affect descending sequences. He devises a test to verify if such assumptions are tenable: if so, the suggestion is to use single bound data, or to use the double bound data from descending sequences only, using the single bound response for the rest of the sample.

3. Bivariate Models

The bivariate model for discrete dependent variables is a two-equation system:

$$Y_{1i} = x_{1i}^{'}\beta_1 + u_{1i}$$
$$Y_{2i} = x_{2i}^{'}\beta_2 + u_{2i}$$
(1)

the dichotomous dependent variables are $y_{1i} = 1$ if $Y_{1i} > 0$; $y_{2i} = 1$ if $Y_{2i} > 0$; x_i are vectors of exogenous variables; $\beta_{1,2}$ are vectors of unknown parameters; and $u_{1,2i}$ are error terms with zero means, variances $\sigma_{1,2}^2$, marginal distribution functions F_1, F_2 and with a joint distribution function *H*.

The bivariate probit model as applied in the contingent valuation literature is defined by equation (1) and the following:

$y_{1i} = 1$ if $Y_{1i} > t_{1i}$; $y_{2i} = 1$ if $Y_{2i} > t_{2i}$, where t_1, t_2 are the bids proposed and H is a bivariate normal with zero vector of means, unit variances and correlation coefficient ρ. Denoting the bivariate normal with zero means, unit variances and correlation ρ by $\Phi(.,.,\rho)$ and defining $a_{1i} = \dfrac{t_{1i} - x'_{1i}\beta_1}{\sigma_1}$ and $a_{2i} = \dfrac{t_{2i} - x'_{2i}\beta_2}{\sigma_2}$ the log-likelihood for the bivariate probit is given by,

$$\ln L(\beta_1, \beta_2, \sigma_1, \sigma_2, \rho) = \sum_{i-1}^{n} (1 - y_{1i})(1 - y_{2i})\Phi(a_{1i}, a_{2i}, \rho) + y_{1i} y_{2i} \Phi(-a_{1i}, -a_{2i}, \rho)$$
$$+ \sum_{i-1}^{n} (1 - y_{1i}) y_{2i} \Phi(a_{1i}, -a_{2i}, -\rho) + y_{1i}(1 - y_{2i})\Phi(-a_{1i}, a_{2i}, -\rho) \quad (2)$$

The four terms in the log-likelihood correspond to “no-no”, “yes-yes”, “no-yes” and “yes-no” responses to the two bid tenders respectively.

The double bound model results from (2) if ρ=1 and the parameters are the same across equations. If the two error terms are not perfectly correlated then the responses to the two bids are governed by a bivariate model, with parameters that may or may not be the same across equations.

The assumption of normality for the WTP distribution is often not supported by the data, and this may give rise to serious misspecification problems. When the model is univariate, the analyst can pick from a wide range of possible distributions the one that better fits the data. When the model is bivariate, the choice is usually constrained to the bivariate normal distribution, and a bivariate probit as above is applied. In practice, WTP is often assumed to have a lognormal distribution, which accounts for the skewness that generally characterizes WTP distributions, and the bivariate normal is applied to the logarithm of WTP. Unfortunately, also the lognormal assumption may not be supported by the data, which implies that the bivariate probit would not be a valid estimator (as also seen in Alberini, 1995): distributional misspecification of the marginals will, in general, result in inconsistent estimates of the parameters since it implies misspecification of the model for the conditional mean of the binary dependent variable (see Ruud, 1983). On the other hand, alternative bivariate distributions, such as the bivariate logistic or the bivariate extreme value, are not as flexible, in terms of correlation allowed between marginals, as the bivariate normal.

As suggested by Hanemann and Kanninen (1999) a possible solution to the problem could be the following: even if the stochastic parts of the two equations are specified as non-normal, they can be transformed into random variables that are characterized by the bivariate normal distribution. This transform, which involves the use of the inverse standard normal distribution, is a special case of a bivariate *copula function*, and is known in econometrics after Lee's (1982, 1983) applications to sample selection models. A general definition for bivariate copulas is:

Definition: A *2-dimensional copula* is a function $\mathbf{C : [0,1]^2 \rightarrow [0,1]}$, with the following properties:

For every $\boldsymbol{u} \in [0,1]$, $\boldsymbol{C(0,u) = C(u,0) = 0}$;

For every $\boldsymbol{u} \in [0,1]$, $\boldsymbol{C(u,1) = u \quad and \quad C(1,u) = u}$;

For every

$(u_1, v_1), (u_2, v_2) \in [0,1] \times [0,1]$ *with* $u_1 \le u_2$ *and* $v_1 \le v_2$:

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \ge 0 .$$

The last condition is the two-dimensional analogue of a nondecreasing one-dimensional function.

The theoretical basis for multivariate modeling through copulas is provided by a theorem due to Sklar (1959).

Sklar's Theorem

Let $\boldsymbol{H}$ *be a joint distribution function with margins* $\boldsymbol{F_1}$ *and* $\boldsymbol{F_2}$*, which are, respectively, the cumulative distribution functions of the random variables* $\boldsymbol{x_1}$ *and* $\boldsymbol{x_2}$*. Then there exists a function* $\boldsymbol{C}$ *such that* $\boldsymbol{H(x_1, x_2) = C(F_1(x_1), F_2(x_2))}$*, for every* $\boldsymbol{x_1, x_2 \in \overline{R}}$ *, where* $\boldsymbol{\overline{R}}$ *represents the extended real line. Conversely, if C is a copula and* $\boldsymbol{F_1}$ *and* $\boldsymbol{F_2}$ *are distribution functions, then the function* H *defined above is a joint distribution function with margins* $\boldsymbol{F_1}$ *and* $\boldsymbol{F_2}$*.*

Since the copula function "links a multidimensional distribution to its one-dimensional margins" (Sklar, 1996), the name "copula" (connection) is explained. The parametric copula approach ensures a high level of flexibility to the modeler, since the specification of the margins $\boldsymbol{F_1}$ and $\boldsymbol{F_2}$ can be separated from the specification of the dependence structure through the function *C* and an underlying parameter θ, which governs the intensity of the dependence.

Although the Lee copula allows flexibility in the choice of the marginals, yet it maintains some restrictive properties (for example, symmetry) of elliptical distributions. More interesting for applied work is the class of Archimedean copulas. These are functions generated by an additive continuous, convex decreasing function φ, with $\varphi(1)=0$. In general, Archimedean copulas have the following form:

$$\varphi(C_\theta(u,v)) = \varphi(u) + \varphi(v).$$

The additive structure of Archimedean copulas makes maximum likelihood estimation, and calculation of the score function, relatively easy. Furthermore, the family is sufficiently large so as to allow a wide range of distributional shapes (right or left skewness, fat or thin tails, etc.). A particular feature of most Archimedean copulas is monotonicity, i.e. they cannot accommodate negative dependence, and this may limit their application in some contexts. In the present application, where the marginals represent the underlying WTP distributions elicited by the double bound method, it is realistic to exclude negative dependence, and use of Archimedean copulas is warranted. Specifically, drawing from previous work (Genius and Strazzera, 2004), we choose the Joe copula, which is defined as follows:

$$C(u,v) = 1 - \left((1-u)^\theta + (1-v)^\theta - (1-u)^\theta(1-v)^\theta\right)^{1/\theta}, \quad \theta \in [1,\infty),$$

where u and v are univariate distributions, and θ is a dependency parameter.

A relevant part of our analysis deals with the estimation of the dependency between equations. When dealing with elliptical copulas (such as the BVN, or the Lee Copula) a valid measure of dependence is linear correlation; however, this does not hold when the bivariate distribution is not elliptical (see fig.1 for a comparison of distributional shapes: the Joe copula is not elliptical).

Alternative measures of dependence include Kendall's τ (K_τ) which is a measure of concordance. It is defined as follows:

$$K_\tau = P\left((X-\tilde{X})(Y-\tilde{Y}) > 0\right) - P\left((X-\tilde{X})(Y-\tilde{Y}) < 0\right),$$

where (X,Y) and $(\tilde{X},\tilde{Y})$ are two independent random vectors with a common distribution function H whose margins are F and G. Kendall's τ can also be expressed in terms of copulas (see Nelsen, 1999):

$$K_\tau = 4\iint_{[0,1]^2} C(u,v)\,dC(u,v) - 1.$$

For continuous random variables the above measure is a measure of concordance, which implies that it takes values in [-1,1], taking the value zero when we have independence. We recall that the linear (or Pearson) correlation is not a measure of dependence: for example, $\rho(x,y)=0$ does not imply independence of the two variables.

Since our estimations involve both elliptical (Normal) and not elliptical (Joe) bivariate distributions, for comparison purposes we report results for the Kendall's τ rather than for the correlation parameter ρ or the dependence parameter θ.

3. Experimental Design

The experiments presented in this paper are aimed at analyzing the performance of competing models when some specific forms of bias affect the responses to the second bid question in the double bound format. We start from a mild form of elicitation effect, where the underlying WTP elicited after the second question is the same as the first, but because of some disturbance in the elicitation process they are not perfectly correlated:

$$\begin{aligned} Y_{1i} &= x_i^{'}\beta + u_i \\ Y_{2i} &= x_i^{'}\beta + \tilde{u}_i \end{aligned}, \qquad 0 < \rho < 1, \tag{3}$$

where ρ is the correlation parameter, and $u_i, \tilde{u}_i$ are identically distributed random variables with mean zero, and variance σ^2. If the random errors are assumed to be distributed as a Normal, this specification gives rise to a BVN model, where parameters of the two equations are constrained to be equal, while the correlation parameter is unconstrained. In Cameron and Quiggin this was deemed as the best specification to fit their double bound data (a well known Australian study for the Kakadu area), which were modeled by means of the univariate interval data (or double bound) model, and several alternative specifications of the bivariate probit model. The simulated data are constructed using the following specification: intercept parameter α=10, slope coefficient β=3, standard deviation σ=5, and a BVN distribution, with correlation ϱ=0.7. The corresponding value of τ is given by $\tau = \frac{2}{\pi}\arcsin(\rho) = 0.493$. The variable x is generated from a uniform, with mean 3.95 and standard deviation 2.05.

The same BVN and interval data specifications used by Cameron and Quiggin were analyzed again by Alberini (1995) on a slightly different sample from the same data, and in this case the preferred specifications was as in 1), but with different estimated variances for the two latent dependent variables. The underlying behavioral hypothesis could be that the cognitive process after the second elicitation question is more “disturbed” –and indeed Alberini finds that the estimated standard deviation of the second equation is substantially higher than the first. This more general structure has been analyzed again in Alberini, Carson and Kanninen (1997). Our Monte Carlo analysis considers both models, which can be respectively referred to as restricted and unrestricted random effects models, with an equality restriction imposed on the standard deviation parameters of the former (experiment A). In the latter model, the standard deviation of the second equation is set at 7 (experiment B).

Another experiment (C) studies the performance of different models when the double bound elicitation method produces more serious forms of bias, leading to a downward shift of the second equation WTP: CGM indicate several possible causes of this effect, briefly reported in the introduction of this paper. In its simplest form the bivariate model with shift is structured as follows:

$$\begin{aligned} Y_{1i} &= x_i^{'}\beta + u_{1i} \\ Y_{2i} &= \delta + x_i^{'}\beta + u_{2i} \end{aligned}, \quad 0 < \rho < 1, \ \delta < 0 \qquad (4)$$

i.e. the shift is simply a leftward translation of the WTP distribution. While more complex specifications may model the shift effect as dependent on some covariates, in our experiment we hold to this basic model, setting $\delta = -2$, the other data being constructed as in experiment B.

The Framed model proposed by DeShazo is relatively new, and, to our knowledge, has not been as yet studied by means of Monte Carlo methods. Here the structure of the model is somewhat more complex, since the model involves a mechanism of sample selection. Theoretically responses to both questions should be dictated by model B but because of framing effects a percentage of respondents belonging to the “yes,yes” class produce responses in the “yes,no” class. In DeShazo’s proposed method to estimate such data affected by framing effects, follow-up responses from individuals who faced a downward sequence of bids enter the second equation, while for individuals facing an upward sequence only the first response is considered, as if it were a single bound elicitation. The bivariate model with selection for descending sequences proposed by DeShazo is the following:

$$\begin{aligned} Y_{1i} &= x_i^{'}\beta + u_{1i} \\ Y_{2i} &= x_i^{'}\beta + u_{2i} \end{aligned}, \quad 0 < \rho < 1, \qquad (5)$$

where Y_2 is modeled for respondents in a descending sequence only (i.e. individuals who responded No to the first elicitation question). The parameter values are constructed as in experiment B above but we switch randomly a percentage of "yes,yes" responses to "yes,no". In order to evaluate the performance of the bivariate estimator versus its univariate counterpart, we also estimate a univariate Framed model: it is a censored Double Bound model, with the second bid included in the equation for respondents in the descending sequence only.

In another experiment (E) we study the performance of different estimators when the initial bids are poorly chosen: in particular, the initial bids of this experiment leave uncovered the left tail of the true WTP distribution, being placed at the 45, 65, and 85 percentile of the distribution. As discussed in the introductory section, this is a case where the double bound method is generally deemed superior to the single bound, since the follow-up question allows inspection of the previously excluded part of the distribution. All other aspects of the data generation process are as in experiment B.

Finally, in experiments F and G we analyze the case where the DGP departs from the Bivariate Normal distribution. WTP distributions are commonly specified as non normal: logistic or extreme value (for WTP or its log) or gamma distributions are typical choices in CV modeling. Application of bivariate models to WTP data may induce a misspecification problem, if the standard bivariate probit is adopted. In this experiment we are especially interested in analyzing the performance of the bivariate probit estimator when the assumption of bivariate normality is wrong. We first simulate a bivariate distribution with normal marginals, but with a dependency structure different from linear correlation: i.e. the Joe copula distribution with normal marginals (JOE-N: experiment F). In experiment G, we use as a DGP a bivariate model with extreme value marginals, again linked by a Joe copula (JOE-E). The dependency parameter τ is set at the value 0.499 which corresponds to a value of θ equal to 2.85. All other settings are as in experiment B. The reader is referred to the Appendix for the algorithm used to generate the data by means of the JOE copula.

Given the above scenarios we analyze the performance of different estimators: the univariate single and double bound models; a univariate model for descending sequences (in experiment D only); the bivariate probit model; the bivariate model based on the Joe Copula with normal marginals for experiments A-F; a bivariate Joe Copula with extreme value marginals for experiment G; and the bivariate probit with sample selection for descending sequences. The respective log-likelihoods for each experiment are given in the table 1, where $a_{1i} = \dfrac{t_{1i} - x_i'\beta}{\sigma}$ and $a_{2i} = \dfrac{t_{2i} - x_i'\beta_2}{\sigma_2}$.

4. Results

The results of the experiments are based on the parameter estimates of the models reported in Table 1, under the different scenarios considered in the previous section. The Monte Carlo analysis involved 400 replications, and was performed using a Gauss 386-i code.

The maximum likelihood estimation required imposition of constraints on the dependence parameters for the bivariate models: specifically, for bivariate models based on the bivariate normal distribution $-1 < \rho < 1$ is imposed while for bivariate models based on the Joe copula the constraint is given by $\theta > 1$. Summary statistics of these results are reported in Tables 2-8. In particular, we show results for bias, standard deviation and empirical size of the α, β, and σ estimates, and the bias and standard deviation of the dependency parameter τ. Bivariate models are in general susceptible to convergence problems, so we also report the number of successful replications: it will be seen that convergence failures can be a serious problem in some circumstances. In our study the definition of failure includes cases where the maximum number of iterations was exceeded and cases where the Hessian failed to invert. For the BVN model the first type of failure was very often associated with the estimate of ρ being on the boundary of the parameter space.

No equality constraints are imposed on the α, β, σ coefficients across equations, since this would be the first approach that a modeler would take to check for those forms of elicitation bias that would result in different parameters across equations, in addition to imperfect correlation.

The results of experiment A (Table 2) show that even a minor elicitation effect, resulting in imperfectly correlated distributions of the error terms, produces some bias in the double bound estimates. On the other hand they are characterized by so small standard errors that the Mean Squared Error (MSE) criterion would rank the double bound estimator best for all sample sizes (as also found in Alberini, 1995). However, it should be noted that the empirical size of the parameters of the double bound model worsens for large samples. The BVN model, which is the correct specification, estimates correctly the correlation coefficient, and proves to be a valid instrument to detect an elicitation effect problem. For experiments A and B we also report the estimates obtained from the more parsimonious BVN model, where the α and β parameters are constrained to be equal across equations for both experiments while σ is constrained only in experiment A (so corresponding to the true DGP). It can be observed that the constrained BVN model brings a substantial gain in efficiency, such that the MSE of the BVN estimates gets close to that of the DOUBLE model. Unfortunately, the BVN model seems susceptible to convergence failure, especially in the large sample experiment. Using a Joe copula when estimating the bivariate model (JOE-N) produces good results in terms of relatively small bias and standard errors, but empirical sizes are higher than nominal, and the dependency parameter is overestimated. This is a general feature that can be observed in all experiments where the Joe copula is fitted instead of the correct BVN distribution. On the other hand, the copula model is often more robust to convergence problems, and may be useful for preliminary exploration of the data. A much higher rate of failures than the BVN characterizes the Bivariate Framed model proposed by DeShazo, which we assess also in the experiments with "unframed" data. The convergence problem is observed in all experiments, including experiment D with "framed" data: this fragility may be seen as a limit to its practical usefulness, although, as is discussed below for other experiments, it generally performs quite well in the estimation of the first equation parameters.

In experiment B (Table 3) variances are different across equations, the second being higher than the first. The relative performance of the double bound model decreases compared to experiment A and the empirical size is above 5% (reaching 0.998 for σ in the larger sample). Overall, the estimates of the bivariate models for the dependence and the first equation parameters are more correct and precise than the corresponding estimates in experiment A and are superior to the single bound model estimates in terms of the MSE criterion. As in the previous experiment, some convergence problems are observed, in this case especially for the small sample: however, they seem to reduce when the constrained model is estimated. Again, it can be noticed how the adoption of the constrained BVN model can effectively correct elicitation effects, while maintaining a high precision in the estimates of the first equation parameters.

Experiment C (Table 4) simulates a response bias characterized by a downward shift, which means that the intercept in the second WTP equation is smaller; and, as in B, by higher variance in the same equation. As discussed in section 2, a downward shift of the WTP distribution elicited after the second bid offer can be due to strategic behavior or risk aversion, behavioral attitudes that can be thought of as heterogeneously varied among individuals, giving rise to higher disturbance in the second WTP model. The double bound estimates are now more evidently biased, as it can also be seen from the empirical size values, which in the large sample are 32% for the constant, and 99.3% for the scale parameter. Both the bivariate probit and the JOE-N model perform quite well in the estimates of the first equation, and are superior to the SINGLE in terms of MSE for the relevant parameters. As expected, given the data generating process used for this experiment, the BVN is superior to the copula model in estimating the dependence parameter, but it is more susceptible to convergence problems, and such vulnerability seems independent of sample size. As usual, the FRAMED2 model is even more fragile in this respect than the others, with 37% of the replications lost in the small sample experiment. It can also be observed that the shift has produced an increase in bias for the estimates of the second equation, especially for FRAMED2.

The data created for experiment D (Table 5) incorporate framing effects in some of the second responses. In this case it could very well happen that the joint distribution of observed responses has changed with respect to the joint distribution of "unframed responses" or that the dependence structure has changed. The average estimate of Kendall's τ is negative for the BVN model (the average ϱ estimate is also negative), and smaller than in previous experiments for the JOE-N –while this does not apply to the sample selection FRAMED2 model. Two salient features of this experiment are that now the DOUBLE model is the worst one in terms of MSE for α and β, with empirical sizes often above 98% and that the estimate of τ is highly biased especially for BVN. For this experiment we tested the univariate censored model FRAMED1 described in the previous section: although its estimates for α and β are satisfactory, we observe that the estimate of σ is strongly biased, and that empirical sizes increase with sample size and are beyond 5%. All bivariate models perform quite well in estimating the first equation parameters, with JOE-N generally slightly better than BVN, which may suggest that the underlying distribution of observed responses might have changed. More importantly, the rate of successful replications of the JOE-N model in experiment D, compared to that of the other two bivariate models confirms the robustness of this copula model, and its usefulness at least at the exploratory stage of the analysis.

In experiment E (Table 6) the initial bids have been changed, while it is maintained the behavioral hypothesis underlying experiment B, resulting in imperfect correlation of the WTP equations plus different associated standard errors. Except for the DOUBLE, whose results are fairly close to the results in experiment B, the estimates of all other models have worsened in terms of bias, inefficiency and empirical sizes for all models, especially for the small sample case. The JOE-N model performs relatively better than the others, even though its empirical sizes tend to increase for larger samples. In any case, all bivariate models, and especially the correctly specified BVN, give a good estimate of the dependence parameter, and therefore are able to signal the presence of some response effect. The DOUBLE model is more robust to poorly chosen starting bids but it is biased, as can be seen from the estimate of the scale parameter which remains biased for large samples.

In experiments F (Table 7) and G (Table 8) we analyze two cases of departure from the assumption of bivariate normality for the WTP distributions. In F the data are generated by a bivariate distribution derived from a Joe copula with normal marginals, while in G the Joe copula links two extreme value distributions. All other settings are as in experiment B, with the qualification that the correct specification now is the JOE copula: in experiment F the Joe copula with normal marginals (JOE-N), in experiment G the Joe copula with extreme value marginals (JOE-E). It can be seen from Tables 7 and 8 that the two copula models estimate correctly all the parameters. While the BVN performs relatively well with normal marginals, its application to a bivariate with non normal marginals results in biased estimates of the first equation parameters. The bad performance of the BVN in this experiment is particularly evident for smaller sample sizes, but it can also be observed that the bias of the variance estimate increases with sample size. Note also that in experiment G the BVN is at high risk of convergence failure: over 15% of the replications for the small sample size and 10% for the large sample size are lost.

5. Conclusions

The need to obtain valid and reliable estimates for the value of non market goods under tighter budget constraints than those characterizing most benchmark CV studies has spurred research to a quest for an elicitation method that maximizes the amount of information obtained from respondents. More efficient elicitation methods allow smaller sample sizes, and this in turn results in less expensive surveys. The double bound method meets this efficiency requirement, but at the cost of potentially inducing elicitation effects, and hence unreliable estimates; on the other hand, the single bound method may not induce response effects, but it is statistically inefficient, and requires large samples in order to give reliable estimates. A solution could be at hand if appropriate statistical analysis could enable the modeler to detect and correct response bias from double bound data, so preserving the property of unbiasedness of the single bound method, and efficiency of the double bound. In this paper we have analyzed the performance of the bivariate modeling approach to the treatment of double bound data affected by elicitation bias. Several experiments were conducted, addressing different types of elicitation problems, and alternative distributional assumptions. We find that the bivariate approach can effectively detect the presence of elicitation effects, and produce correct estimates while maintaining a satisfactory level of efficiency. We have shown that application of a correctly specified bivariate model can lead to efficiency levels close to those reported for the double bound model. Clearly, the bivariate approach should be based on well specified models, since application of, say, a bivariate probit to non normal bivariate distributions would result in biased estimates. This was clearly shown in one experiment, where the bivariate probit was applied to a bivariate distribution with extreme value marginals.

We propose the adoption of copula models as a flexible instrument to fit non normal bivariate distributions. Specifically, in this analysis we adopted the Joe copula, which turns out to be quite robust to convergence problems, and therefore can be used in preliminary analysis of the data to check for presence of response effects.

Finally, we found that convergence problems often affect the bivariate probit model, and even more seriously the bivariate probit with sample selection proposed by DeShazo to model data affected by framed effects. We showed that even when data are characterized by this type of bias, the sample selection model does not perform better than the bivariate models with no selection, especially the Joe copula: it is possible that framing produces some effect on the dependency structure, which can be better modeled by asymmetric bivariate distributions, such as the Joe copula.

Obviously, our results only apply to the forms of response bias considered in the present work: further research is called for to explore the effectiveness of the bivariate approach for correction of other relevant sources of bias, such as anchoring effects.

References

Alberini, A. (1995), "Efficiency vs Bias of Willingness to Pay Estimates: Bivariate and Interval Data Models", Journal of Environmental Economics and Management, 29:169-80.

Alberini, A., R.T. Carson and B.J. Kanninen (1997), "Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data", Land Economics, 73:309-24.

Cameron, T.A. and J. Quiggin (1994), "Estimation Using Contingent Valuation Data from a 'Dichotomous Choice with Follow-up' Questionnaire", Journal of Environmental Economics and Management, 27:218-34.

Carson, R.T, N.E. Flores and N.F. Meade (2001), "Contingent Valuation: Controversies and Evidence", Environmental and Resource Economics, 19:173-210.

Carson, R.T., T. Groves and M.J. Machina (1999), "Incentive and Informational Properties of Preference Questions", Plenary Address to the European Association of Resource and Environmental Economists Conference, Oslo, June 1999.

Cooper, J.C., W.M. Hanemann, and G. Signorello (2002), "One and One Half Bound Dichotomous Choice Contingent Valuation", The Review of Economics and Statistics, 84:742-50.

DeShazo, J.R. (2002), "Designing Transactions without Framing Effects in Iterative Question Formats", Journal of Environmental Economics and Management, 43:360-85.

Embrechts P, F. Lindskog F and A.J. McNeil (2003), "Modelling dependence with copulas and applications to risk management," *Handbook of heavy tailed distributions in finance*, edited by Rachev S.T., Elsevier/North-Holland, Amsterdam.

Genius, M. and E. Strazzera (2004), "The Copula Approach to Sample Selection Modelling: an application to the Recreational Value of Forests", Nota di Lavoro 73-2004, FEEM.

Green, D., K.E. Jacowitz, D. Kahneman, D. McFadden (1998), "Referendum contingent valuation, anchoring, and willingness to pay for public goods", Resource and Energy Economics, 20:85-116.

Hanemann, W.M., J. Loomis and B.J. Kanninen (1991), "Statistical Efficiency of Double-bounded Dichotomous Choice Contingent Valuation", American Journal of Agricultural Economics, 73:1255-63.

Hanemann, W.M., and B.J. Kanninen (1999), "The Statistical Analysis of Discrete-Response CV Data", *Valuing Environmental Preferences*, edited by Bateman, I.J. and K.G. Willis, Oxford U.P.

Herriges, J.A. and J.F. Shogren (1996), "Starting Point Bias in Dichotomous Choice Valuation with Follow-up Questioning", Journal of Environmental Economics and Management, 30:112-31.

Kahneman, D. and A. Tversky (1979), "Prospect Theory: an Analysis of Decisions under Risk", Econometrica, 47:263-91

Lee, L.-F. (1983), "Generalized Econometric Models with Selectivity," *Econometrica* 51: 507–12.

Lee, L. F. (1982), "Some Approaches to the Correction of Selectivity Bias," *Review of Economic Studies*, 49: 355-372.

McLeod, D.M. and O. Bergland (1999), "Willingness to Pay Estimates Using the Double Bounded Dichotomous Choice Contingent Valuation Format: a Test for Validity and Precision in a Bayesian Framework", Land Economics, 75:115-25.

Nelsen, R. B. (1999). *An Introduction to Copulas*. Lecture Notes in Statistics. New York: Springer-Verlag.

Ruud, Paul A. (1983), "Sufficient Conditions for the Consistency of Maximum Likelihood Estimation Despite Misspecification of Distribution in Multinomial Discrete Choice Models", Econometrica, 51: 225-228.

Sklar, A. (1959), "Fonctions de Répartition à N Dimensions et Leurs Marges," *Publ. Inst. Statis. Univ. Paris,* 8: 229-231.

Sklar A. (1996), "Random Variables, Distribution Functions, and Copulas - a Personal Look Backward and Forward", published in: *Distributions with Fixed Marginals and Related Topics*, edited by L. Rüschendorf, B. Schweizer, and M.D. Taylor, Institute of Mathematical Statistics, Hayward, CA.

Whitehead, J.C. (2002), "Incentive Incompatibility and Starting Point Bias in Iterative Valuation Questions", Land Economics, 78:285-97.

Figure 1: Contour plots of Bivariate Normal and Joe Copula when marginals are N(0,1).

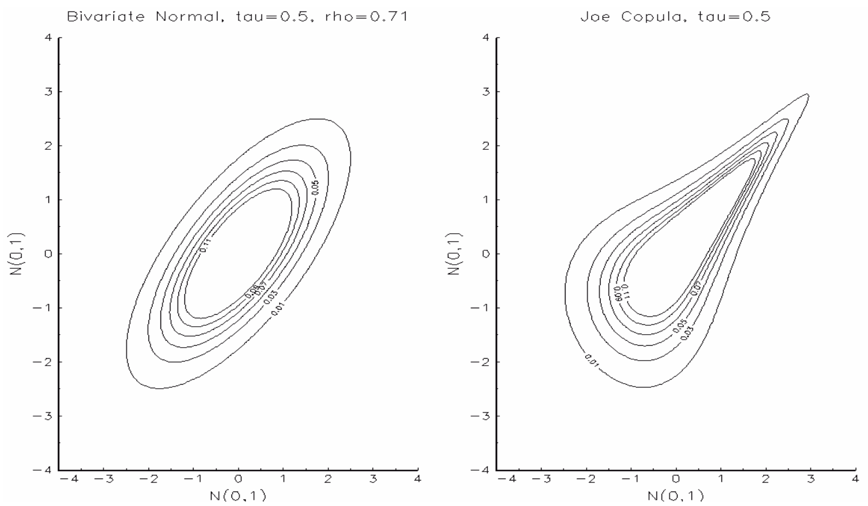

Table 1: Log-likelihood of estimated models

ESTIMATED MODELS	LOG-LIKELIHOOD TERMS: Experiments (A) to (F)	LOG-LIKELIHOOD TERMS: Experiment (G), where the cdf for extreme value is given by, $G(a)=\exp(-\exp(-a))$
SINGLE	P("yes"): $1-\Phi(a_{1i})$ P("no"): $\Phi(a_{1i})$	P("yes"): $1-G(-a_{1i})$ P("no"): $G(-a_{1i})$
DOUBLE	P("n,n"): $\Phi(a_{2i})$ P("n,y"): $\Phi(a_{1i})-\Phi(a_{2i})$ P("y,n"): $\Phi(a_{2i})-\Phi(a_{1i})$ P("y,y"): $1-\Phi(a_{2i})$	P("n,n"): $G(-a_{2i})$ P("n,y"): $G(a_{1i})-G(a_{2i})$ P("y,n"): $G(a_{2i})-G(a_{1i})$ P("y,y"): $1-G(a_{2i})$
FRAMED1 (experiment D only)	P("n,n"): $\Phi(a_{2i})$ P("n,y"): $\Phi(a_{1i})-\Phi(a_{2i})$ P("yes"): $1-\Phi(a_{1i})$	P("n,n"): $G(a_{2i})$ P("n,y"): $G(a_{1i})-G(a_{2i})$ P("yes"): $1-G(a_{1i})$
BVN	P("n,n"): $\Phi(a_{1i},a_{2i},\rho)$ P("n,y"): $\Phi(a_{1i},-a_{2i},-\rho)$ P("y,n"): $\Phi(-a_{1i},a_{2i},-\rho)$ P("y,y"): $\Phi(-a_{1i},-a_{2i},\rho)$	
JOE COPULA	P("n,n"): $C(\Phi(a_{1i}),\Phi(a_{2i}),\theta)$ P("n,y"): $\Phi(a_{1i})-C(\Phi(a_{1i}),\Phi(a_{2i}),\theta)$ P("y,n"): $\Phi(a_{2i})-C(\Phi(a_{1i}),\Phi(a_{2i}),\theta)$ P("y,y"):1-P("n,n")-P("n,y")-P("y,n")	P("n,n"): $C(G(a_{1i}),G(a_{2i}),\theta)$ P("n,y"): $G(a_{1i})-C(G(a_{1i}),G(a_{2i}),\theta)$ P("y,n"): $G(a_{2i})-C(G(a_{1i}),G(a_{2i}),\theta)$ P("y,y"):1-P("n,n")-P("n,y")-P("y,n")
FRAMED2	P("n,n"): $\Phi(a_{1i},a_{2i},\rho)$ P("n,y"): $\Phi(a_{1i},-a_{2i},-\rho)$ P("yes"): $1-\Phi(a_{1i})$	P("n,n"): $C(G(a_{1i}),G(a_{2i}),\theta)$ P("n,y"): $G(a_{1i})-C(G(a_{1i}),G(a_{2i}),\theta)$ P("yes"): $1-G(a_{1i})$

Table 2. Experiment A: $\alpha = \alpha_2 = 10, \beta = \beta_2 = 3, \sigma = \sigma_2 = 5, \rho = 0.7, \tau = 0.493.$

Size 200	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.117 1.971 (0.060)	0.043 0.492 (0.060)	0.010 0.855 (0.070)	-	-	-	-	-
DOUBLE	0.232 0.955 (0.067)	-0.061 0.231 (0.065)	-0.199 0.402 (0.120)	-	-	-	-	-
BVN	-0.235 1.984 (0.065)	0.075 0.490 (0.060)	0.083 0.847 (0.052)	-0.325 1.364 (0.039)	0.095 0.357 (0.039)	0.279 0.855 (0.034)	-0.089 0.252	382
BVN Restricted	-0.186 1.199 (0.045)	0.057 0.298 (0.045)	0.132 0.584 (0.042)	-	-	-	-0.021 0.259	381
JOE-N	-0.044 1.900 (0.084)	0.021 0.479 (0.079)	-0.093 0.864 (0.087)	-0.185 1.387 (0.040)	0,042 0.357 (0.047)	0.055 0.943 (0.069)	0.204 0.274	379
FRAMED2	-0.048 1.968 (0.064)	0.023 0.483 (0.067)	-0.041 0.860 (0.073)	-2.090 4.147 (0.027)	0.337 1.148 (0.049)	1.835 3.319 (0.034)	-0.284 0.469	328
Size 400	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.047 1.240 (0.068)	0.016 0.310 (0.068)	0.008 0.583 (0.065)	-	-	-	-	-
DOUBLE	0.201 0.705 (0.083)	-0.060 0.174 (0.078)	-0.186 0.297 (0.135)	-	-	-	-	-
BVN	-0.104 1.216 (0.057)	0.031 0.304 (0.054)	0.037 0.581 (0.054)	-0.169 0.955 (0.032)	0.041 0.240 (0.032)	0.113 0.600 (0.068)	-0.030 0.211	370
BVN Restricted	-0.107 0.797 (0.045)	0.026 0.199 (0.037)	0.056 0.396 (0.052)	-	-	-	-0.007 0.207	382
JOE-N	0.021 1.173 (0.078)	-0.008 0.299 (0.069)	-0.111 0.588 (0.098)	-0.059 0.900 (0.061)	0.003 0.225 (0.061)	-0.032 0.534 (0.111)	0.242 0.249	377
FRAMED2	0.025 1.157 (0.058)	0.001 0.295 (0.061)	-0.049 0.566 (0.073)	-1.721 4.182 (0.049)	0.283 0.919 (0.052)	1.443 3.651 (0.055)	-0.184 0.429	344
Size 1000	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.018 0.862 (0.050)	-0.001 0.204 (0.050)	-0.034 0.379 (0.055)	-	-	-	-	-
DOUBLE	0.200 0.490 (0.110)	-0.055 0.112 (0.105)	-0.165 0.171 (0.160)	-	-	-	-	-
BVN	-0.072 0.810 (0.042)	0.012 0.193 (0.045)	-0.022 0.366 (0.056)	-0.068 0.667 (0.076)	0.017 0.162 (0.064)	0.060 0.374 (0.050)	0.027 0.187	357
BVN Restricted	-0.069 0.557 (0.066)	0.014 0.130 (0.066)	0.026 0.252 (0.055)	-	-	-	0.029 0.180	363
JOE-N	0.058 0.791 (0.061)	-0.020 0.193 (0.071)	-0.109 0.388 (0.110)	0.051 0.640 (0.090)	-0.019 0.156 (0.087)	-0.068 0.342 (0.108)	0.277 0.217	379
FRAMED2	0.004 0.857 (0.050)	-0.004 0.202 (0.050)	-0.053 0.377 (0.068)	-0.835 2.072 (0.032)	0.124 0.507 (0.050)	0.699 1.675 (0.053)	-0.082 0.372	339

***number of replications where convergence is achieved and the Hessian is invertible.**

Table 3. Experiment B: $\alpha = \alpha_2 = 10, \beta = \beta_2 = 3, \sigma = 5, \sigma_2 = 7, \rho = 0.7, \tau = 0.493.$

Size 200	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.117 1.972 (0.060)	0.044 0.492 (0.060)	0.010 0.855 (0.070)	-	-	-	-	-
DOUBLE	0.359 1.080 (0.078)	-0.103 0.264 (0.063)	0.794 0.530 (0.323)	-	-	-	-	-
BVN	-0.180 1.916 (0.070)	0.065 0.476 (0.061)	0.069 0.822 (0.058)	-0.430 2.004 (0.029)	0.120 0.537 (0.029)	0.593 2.008 (0.093)	-0.041 0.197	345
BVN Restricted	-0.190 1.524 (0.053)	0.061 0.384 (0.045)	0.075 0.711 (0.035)	-	-	0.468 1.654 (0.088)	-0.040 0.195	375
JOE-N	0.023 1.895 (0.080)	0.013 0.470 (0.078)	-0.030 0.819 (0.080)	-0.672 2.939 (0.044)	0.175 0.839 (0.057)	0.843 3.886 (0.178)	0.190 0.296	387
FRAMED2	-0.101 2.030 (0.065)	0.039 0.501 (0.068)	-0.028 0.850 (0.078)	-1.211 3.693 (0.048)	0.131 1.341 (0.068)	1.613 4.167 (0.102)	-0.181 0.376	293
Size 400	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.047 1.240 (0.068)	0.016 0.309 (0.068)	0.008 0.583 (0.065)	-	-	-	-	-
DOUBLE	0.313 0.796 (0.100)	-0.097 0.195 (0.115)	0.809 0.344 (0.678)	-	-	-	-	-
BVN	-0.073 1.192 (0.058)	0.021 0.308 (0.061)	0.005 0.557 (0.058)	-0.173 1.195 (0.044)	0.040 0.304 (0.033)	0.177 1.210 (0.127)	0.005 0.171	362
BVN Restricted	-0.077 0.961 (0.046)	0.016 0.240 (0.056)	0.020 0.474 (0.046)	-	-	0.128 1.010 (0.116)	-0.002 0.156	372
JOE-N	0.145 1.136 (0.089)	-0.032 0.292 (0.095)	-0.074 0.524 (0.072)	-0.097 1.500 (0.081)	0.003 0.387 (0.075)	-0.087 1.693 (0.301)	0.255 0.250	359
FRAMED2	-0.062 1.263 (0.059)	0.019 0.315 (0.065)	-0.022 0.576 (0.068)	-1.346 3.811 (0.071)	0.216 1.018 (0.071)	1.639 4.358 (0.103)	-0.101 0.309	340
Size 1000	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.018 0.862 (0.050)	-0.001 0.204 (0.050)	-0.034 0.379 (0.055)	-	-	-	-	-
DOUBLE	0.259 0.549 (0.108)	-0.075 0.124 (0.100)	0.837 0.219 (0.998)	-	-	-	-	-
BVN	-0.063 0.819 (0.049)	0.011 0.194 (0.046)	-0.011 0.355 (0.057)	-0.108 0.837 (0.060)	0.026 0.201 (0.044)	0.081 0.772 (0.125)	0.011 0.133	367
BVN Restricted	-0.064 0.643 (0.059)	0.012 0.148 (0.051)	-0.009 0.310 (0.053)	-	-	0.047 0.623 (0.096)	0.007 0.120	375
JOE-N	0.185 0.762 (0.078)	-0.046 0.185 (0.086)	-0.069 0.345 (0.078)	0.102 1.094 (0.119)	-0.044 0.261 (0.127)	-0.372 1.126 (0.553)	0.313 0.208	371
FRAMED2	0.001 0.864 (0.064)	-0.006 0.202 (0.067)	-0.050 0.374 (0.052)	-0.654 2.209 (0.087)	0.083 0.566 (0.067)	0.750 2.510 (0.108)	-0.050 0.259	343

***number of replications where convergence is achieved and the Hessian is invertible.**

Table 4. Experiment C: $\alpha_1 = 10, \alpha_2 = 8, \beta = \beta_2 = 3, \sigma = 5, \sigma_2 = 7, \rho = 0.7, \tau = 0.493.$

Size 200	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.117 1.971 (0.060)	0.043 0.492 (0.060)	0.010 0.855 (0.070)	-	-	-	-	-
DOUBLE	-0.739 1.212 (0.080)	-0.104 0.274 (0.085)	0.764 0.533 (0.320)	-	-	-	-	-
BVN	-0.180 1.850 (0.057)	0.060 0.448 (0.049)	0.053 0.782 (0.049)	-0.672 2.641 (0.043)	0.136 0.541 (0.046)	0.509 1.810 (0.089)	-0.042 0.200	348
JOE-N	-0.096 1.905 (0.0804)	0.030 0.470 (0.076)	-0.058 0.837 (0.098)	-0.855 4.516 (0.092)	0.165 0.908 (0.095)	0.632 3.360 (0.211)	0.200 0.281	369
FRAMED2	-0.185 2.047 (0.067)	0.051 0.507 (0.063)	0.012 0.818 (0.067)	-0.643 3.856 (0.091)	-0.076 1.038 (0.079)	0.606 3.229 (0.143)	-0.116 0.360	252
Size 400	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.047 1.240 (0.068)	0.016 0.309 (0.068)	0.008 0.583 (0.065)	-	-	-	-	-
DOUBLE	-0.789 0.849 (0.130)	-0.099 0.199 (0.105)	0.802 0.361 (0.648)	-	-	-	-	-
BVN	-0.087 1.190 (0.049)	0.031 0.300 (0.060)	0.031 0.571 (0.057)	-0.206 1.619 (0.052)	0.039 0.340 (0.049)	0.150 1.099 (0.083)	0.001 0.154	348
JOE-N	0.005 1.162 (0.092)	-0.008 0.299 (0.094)	-0.082 0.561 (0.097)	0.167 2.134 (0.143)	-0.037 0.433 (0.129)	-0.162 1.632 (0.332)	0.242 0.233	371
FRAMED2	-0.040 1.234 (0.071)	0.015 0.307 (0.068)	-0.032 0.577 (0.065)	-1.528 4.419 (0.104)	0.088 0.834 (0.068)	1.430 4.173 (0.123)	-0.097 0.312	309
Size 1000	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.018 0.862 (0.050)	-0.001 0.204 (0.050)	-0.034 0.379 (0.055)	-	-	-	-	-
DOUBLE	-0.869 0.590 (0.320)	-0.072 0.129 (0.110)	0.841 0.220 (0.993)	-	-	-	-	-
BVN	-0.068 0.841 (0.052)	0.009 0.198 (0.055)	0.005 0.355 (0.057)	-0.093 1.060 (0.063)	0.014 0.209 (0.055)	0.086 0.731 (0.086)	0.005 0.123	348
JOE-N	0.004 0.798 (0.063)	-0.016 0.188 (0.058)	-0.042 0.347 (0.089)	0.434 1.474 (0.228)	-0.077 0.278 (0.189)	-0.325 1.186 (0.430)	0.266 0.206	381
FRAMED2	0.0170 0.838 (0.051)	-0.006 0.199 (0.044)	-0.041 0.370 (0.061)	-1.310 3.713 (0.105)	0.151 0.678 (0.088)	1.177 3.317 (0.136)	-0.074 0.247	295

***number of replications where convergence is achieved and the Hessian is invertible.**

Table 5. Experiment D: $\alpha = \alpha_2 = 10, \beta = \beta_2 = 3, \sigma = 5, \sigma_2 = 7, \rho = 0.7, \tau = 0.493.$

Size 200	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.117 1.972 (0.060)	0.043 0.492 (0.060)	0.010 0.855 (0.070)	-	-	-	-	-
DOUBLE	1.964 0.890 (0.613)	-0.928 0.202 (0.985)	0.006 0.347 (0.043)	-	-	-	-	-
FRAMED1	0.192 1.229 (0.050)	-0.011 0.357 (0.048)	0.598 0.628 (0.075)	-	-	-	-	-
BVN	-0.231 1.958 (0.051)	0.079 0.493 (0.054)	0.054 0.910 (0.075)	-4.811 5.107 (0.010)	-0.635 0.816 (0.200)	11.402 8.094 (0.003)	-0.651 0.132	295
JOE-N	0.014 1.993 (0.078)	-0.009 0.497 (0.103)	-0.090 0.868 (0.097)	-1.533 3.304 (0.062)	-0.919 0.536 (0.462)	4.880 4.409 (0.073)	-0.249 0.366	370
FRAMED2	-0.107 2.027 (0.071)	0.038 0.500 (0.071)	-0.021 0.848 (0.075)	-1.070 3.578 (0.047)	0.093 1.343 (0.068)	1.576 4.167 (0.102)	-0.191 0.376	295
Size 400	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.047 1.240 (0.068)	0.016 0.309 (0.678)	0.008 0.583 (0.065)	-	-	-	-	-
DOUBLE	1.884 0.663 (0.770)	-0.904 0.149 (1.000)	0.134 0.260 (0.073)	-	-	-	-	-
FRAMED1	0.210 0.910 (0.080)	-0.030 0.258 (0.053)	0.598 0.461 (0.218)	-	-	-	-	-
BVN	-0.067 1.186 (0.056)	0.032 0.300 (0.053)	0.077 0.577 (0.039)	-6.588 5.453 (0.018)	-0.375 0.792 (0.166)	13.738 9.002 (0.018)	-0.666 0.088	337
JOE-N	-0.014 1.236 (0.078)	0.002 0.315 (0.084)	-0.024 0.585 (0.084)	-1.792 2.214 (0.048)	-0.854 0.368 (0.570)	5.146 2.876 (0.638)	-0.385 0.257	395
FRAMED2	-0.055 1.259 (0.061)	0.019 0.314 (0.067)	-0.038 0.571 (0.067)	-1.362 3.800 (0.064)	0.211 1.018 (0.067)	1.612 4.320 (0.096)	-0.107 0.313	342
Size 1000	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.018 0.862 (0.050)	-0.001 0.204 (0.050)	-0.034 0.379 (0.055)	-	-	-	-	-
DOUBLE	1.964 0.458 (0.985)	-0.844 0.096 (1.000)	0.058 0.152 (0.035)	-	-	-	-	-
FRAMED1	0.197 0.600 (0.083)	-0.036 0.155 (0.075)	0.587 0.262 (0.578)	-	-	-	-	-
BVN	-0.120 0.888 (0.052)	0.040 0.209 (0.055)	0.071 0.395 (0.049)	-5.581 4.060 (0.076)	-0.428 0.537 (0.308)	12.395 6.648 (0.598)	-0.670 0.071	328
JOE-N	-0.003 0.858 (0.054)	-0.005 0.206 (0.054)	-0.041 0.382 (0.061)	-1.062 1.229 (0.069)	-0.865 0.222 (0.949)	4.307 1.371 (0.980)	-0.456 0.130	392
FRAMED2	-0.022 0.865 (0.056)	0.001 0.205 (0.065)	-0.030 0.383 (0.050)	-0.609 2.117 (0.080)	0.076 0.547 (0.067)	0.734 2.531 (0.106)	-0.050 0.259	341

***number of replications where convergence is achieved and the Hessian is invertible.**

Table 6. Experiment E: $\alpha = \alpha_2 = 10, \beta = \beta_2 = 3, \sigma = 5, \sigma_2 = 7, \rho = 0.7, \tau = 0.493.$

Size 200	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.313 2.573 (0.060)	0.063 0.535 (0.065)	0.024 0.970 (0.078)	-	-	-	-	-
DOUBLE	0.310 1.244 (0.068)	-0.072 0.272 (0.063)	0.857 0.539 (0.405)	-	-	-	-	-
BVN	-0.524 2.563 (0.054)	0.108 0.526 (0.045)	0.096 0.994 (0.080)	-0.869 2.679 (0.043)	0.185 0.598 (0.045)	0.878 2.537 (0.119)	-0.065 0.255	352
JOE-N	-0.192 2.382 (0.069)	0.043 0.494 (0.053)	0.003 0.947 (0.085)	-1.160 4.278 (0.050)	0.226 0.913 (0.053)	1.090 4.227 (0.140)	0.181 0.299	377
FRAMED2	-0.324 2.578 (0.046)	0.073 0.539 (0.055)	-0.0215 0.979 (0.095)	-1.827 5.696 (0.058)	0.274 1.504 (0.076)	1.756 5.307 (0.076)	-0.176 0.394	328
Size 400	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.189 1.802 (0.040)	0.044 0.375 (0.038)	0.030 0.653 (0.053)	-	-	-	-	-
DOUBLE	0.276 0.876 (0.073)	-0.058 0.193 (0.073)	0.914 0.367 (0.770)	-	-	-	-	-
BVN	-0.280 1.711 (0.041)	0.063 0.356 (0.044)	0.051 0.636 (0.050)	-0.485 1.807 (0.052)	0.104 0.392 (0.057)	0.402 1.644 (0.010)	-0.022 0.211	366
JOE-N	0.065 1.608 (0.074)	-0.004 0.339 (0.072)	-0.021 0.628 (0.080)	-0.405 2.693 (0.090)	0.071 0.555 (0.093)	0.187 2.553 (0.253)	0.265 0.254	376
FRAMED2	-0.160 1.798 (0.053)	0.044 0.373 (0.048)	-0.008 0.660 (0.059)	-1.230 3.955 (0.062)	0.178 1.005 (0.065)	1.255 3.748 (0.070)	-0.130 0.306	356
Size 1000	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.036 1.250 (0.043)	0.004 0.252 (0.040)	-0.010 0.450 (0.053)	-	-	-	-	-
DOUBLE	0.260 0.597 (0.090)	-0.054 0.132 (0.098)	0.936 0.231 (0.990)	-	-	-	-	-
BVN	-0.070 1.199 (0.067)	0.010 0.243 (0.053)	0.009 0.436 (0.050)	-0.231 1.081 (0.044)	0.047 0.234 (0.044)	0.249 0.989 (0.155)	0.007 0.177	342
JOE-N	0.220 1.147 (0.080)	-0.044 0.233 (0.077)	-0.054 0.413 (0.075)	-0.014 1.528 (0.109)	-0.007 0.317 (0.123)	-0.136 1.535 (0.469)	0.296 0.232	375
FRAMED2	0.022 1.216 (0.044)	-0.007 0.245 (0.044)	-0.050 0.434 (0.055)	-0.742 2.282 (0.032)	0.106 0.558 (0.032)	0.730 2.097 (0.050)	-0.065 0.274	343

***number of replications where convergence is achieved and the Hessian is invertible.**

Table 7. Experiment F: $\alpha = \alpha_2 = 10, \beta = \beta_2 = 3, \sigma = 5, \sigma_2 = 7, \tau = 0.499.$

Size 200	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.292 1.757 (0.035)	0.086 0.474 (0.042)	0.069 0.932 (0.070)	-	-	-	-	-
DOUBLE	0.409 1.069 (0.085)	-0.101 0.287 (0.092)	0.759 0.512 (0.315)	-	-	-	-	-
BVN	-0.475 1.722 (0.029)	0.105 0.459 (0.032)	0.071 0.893 (0.076)	-0.385 2.030 (0.044)	0.164 0.596 (0.058)	0.834 2.351 (0.097)	-0.076 0.210	340
JOE-N	-0.349 1.644 (0.028)	0.096 0.444 (0.033)	0.037 0.869 (0.084)	-0.350 1.860 (0.046)	0.129 0.544 (0.043)	0.58 2.088 (0.099)	0.027 0.246	393
FRAMED2	-0.248 1.717 (0.035)	0.069 0.461 (0.050)	-0.021 0.910 (0.085)	-2.145 4.530 (0.043)	0.283 1.257 (0.050)	1.982 5.019 (0.116)	-0.275 0.397	258
Size 400	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.107 1.239 (0.053)	0.028 0.306 (0.055)	0.008 0.589 (0.070)	-	-	-	-	-
DOUBLE	0.371 0.837 (0.088)	-0.098 0.201 (0.110)	0.797 0.366 (0.635)	-	-	-	-	-
BVN	-0.224 1.211 (0.050)	0.038 0.298 (0.055)	0.039 0.558 (0.050)	-0.179 1.391 (0.050)	0.091 0.352 (0.045)	0.470 1.326 (0.077)	-0.049 0.169	379
JOE-N	-0.132 1.205 (0.060)	0.031 0.299 (0.073)	0.003 0.554 (0.060)	-0.259 1.425 (0.055)	0.074 0.362 (0.053)	0.269 1.280 (0.101)	0.026 0.209	397
FRAMED2	-0.055 1.279 (0.066)	0.015 0.316 (0.069)	-0.017 0.596 (0.082)	-3.360 4.573 (0.039)	0.527 1.097 (0.023)	3.305 5.137 (0.069)	-0.358 0.342	304
Size 1000	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.035 0.859 (0.053)	0.010 0.209 (0.050)	0.015 0.371 (0.043)	-	-	-	-	-
DOUBLE	0.377 0.498 (0.113)	-0.084 0.121 (0.113)	0.834 0.221 (0.98)	-	-	-	-	-
BVN	-0.192 0.830 (0.047)	0.026 0.202 (0.055)	0.030 0.362 (0.049)	0.018 0.844 (0.058)	0.043 0.215 (0.047)	0.253 0.924 (0.113)	-0.027 0.138	364
JOE-N	-0.078 0.798 (0.043)	0.018 0.194 (0.053)	0.012 0.355 (0.048)	-0.123 0.822 (0.038)	0.039 0.203 (0.050)	0.169 0.770 (0.090)	0.009 0.155	398
FRAMED2	-0.034 0.864 (0.056)	0.007 0.206 (0.039)	0.015 0.372 (0.052)	-4.327 8.079 (0.056)	0.743 1.808 (0.036)	4.405 9.123 (0.059)	-0.415 0.257	305

***number of replications where convergence is achieved and the Hessian is invertible.**

Table 8. Experiment G: $\alpha = \alpha_2 = 10, \beta = \beta_2 = 3, \sigma = 5, \sigma_2 = 7, \tau = 0.499.$

Size 200	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.192 1.578 (0.035)	0.064 0.410 (0.053)	0.006 0.924 (0.073)	-	-	-	-	-
DOUBLE	0.221 1.028 (0.070)	-0.069 0.268 (0.093)	0.896 0.609 (0.358)	-	-	-	-	-
BVN	-0.588 1.590 (0.027)	0.083 0.420 (0.041)	-0.281 0.834 (0.157)	-0.916 1.745 (0.015)	0.144 0.461 (0.027)	0.051 1.611 (0.183)	0.008 0.209	338
JOE-E	-0.289 1.524 (0.032)	0.072 0.398 (0.053)	-0.077 0.871 (0.111)	-0.271 1.513 (0.076)	0.087 0.423 (0.053)	0.325 1.419 (0.084)	0.034 0.248	380
FRAMED2	-0.181 1.597 (0.030)	0.064 0.418 (0.052)	0.002 0.933 (0.074)	-1.425 3.495 (0.033)	0.368 1.156 (0.044)	0.902 3.128 (0.077)	0.061 0.226	363
Size 400	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	-0.070 1.177 (0.065)	0.019 0.286 (0.055)	-0.027 0.596 (0.065)	-	-	-	-	-
DOUBLE	0.224 0.795 (0.075)	-0.073 0.186 (0.098)	0.937 0.400 (0.715)	-	-	-	-	-
BVN	-0.463 1.183 (0.055)	0.039 0.291 (0.052)	-0.315 0.538 (0.149)	-0.950 1.312 (0.052)	0.154 0.312 (0.044)	-0.100 1.003 (0.195)	0.012 0.166	343
JOE-E	-0.090 1.135 (0.069)	0.017 0.271 (0.056)	-0.059 0.539 (0.053)	-0.223 1.142 (0.056)	0.067 0.295 (0.051)	0.196 0.989 (0.084)	0.039 0.225	394
FRAMED2	-0.008 1.155 (0.070)	0.005 0.274 (0.052)	-0.037 0.588 (0.064)	-1.466 2.950 (0.023)	0.342 0.889 (0.038)	0.804 2.579 (0.049)	0.023 0.220	345
Size 1000	Bias, Standard Deviation (STD) and Empirical Size (Nominal size 5%)						Bias, STD	
	α	β	σ	α_2	β_2	σ_2	τ	reps*
SINGLE	0.017 0.794 (0.065)	-0.003 0.185 (0.050)	0.012 0.382 (0.048)	-	-	-	-	-
DOUBLE	0.212 0.501 (0.063)	-0.058 0.116 (0.080)	0.989 0.240 (0.992)	-	-	-	-	-
BVN	-0.299 0.729 (0.056)	-0.019 0.178 (0.067)	-0.350 0.325 (0.264)	-0.588 0.767 (0.081)	0.067 0.181 (0.039)	-0.398 0.624 (0.419)	0.080 0.164	360
JOE-E	-0.037 0.736 (0.056)	0.004 0.174 (0.051)	-0.009 0.359 (0.059)	-0.068 0.707 (0.038)	0.025 0.181 (0.048)	0.077 0.621 (0.092)	0.041 0.182	393
FRAMED2	0.047 0.781 (0.060)	-0.008 0.184 (0.054)	-0.001 0.381 (0.048)	-0.978 1.799 (0.006)	0.169 0.471 (0.031)	0.466 1.367 (0.028)	-0.007 0.219	352

***number of replications where convergence is achieved and the Hessian is invertible.**

Appendix

The algorithm below can be used to generate pairs of pseudo random numbers with arbitrary marginal distribution functions but with a Joe copula as a joint distribution function: see Embrechts et al (2003). The Joe copula is defined by,

$$C(u,v) = 1 - \left((1-u)^{\theta} + (1-v)^{\theta} - (1-u)^{\theta}(1-v)^{\theta}\right)^{1/\theta}, \quad \theta \in [1,\infty)$$

and has a generator given by $\varphi(t) = -\ln(1-(1-t)^{\theta})$. For any Archimedean copula the function K_c defined as, $K_c(t) = t - \frac{\varphi(t)}{\varphi'(t)}$ is the distribution function of the copula *C*. The algorithm goes through the following steps:

-step 1: simulate two independent U[0,1] random variables *s* and *q*.

-step 2: choose the value of θ and set $t = K_c^{-1}(q)$. Since there is no closed form expression for the inverse of K_c, the equation $\left(t - \frac{\varphi(t)}{\varphi'(t)}\right) - q = 0$ has to be solved numerically using a root finding procedure.

-step 3: set $u = \varphi^{-1}(s\varphi(t))$ and $v = \varphi^{-1}((1-s)\varphi(t))$, where $\varphi^{-1}(t) = 1-(1-e^{-t})^{1/\theta}$ for the Joe copula. The pseudo random numbers *u,v* are uniformly distributed on [0,1] and have a Joe copula as a joint distribution function.

-step 4: for arbitrary distribution functions F_1 and F_2 define,

$$r_1 = F_1^{-1}(u)$$
$$r_2 = F_2^{-1}(v)$$

The pseudo random numbers r_1, r_2 have marginal distributions given by F_1 and F_2 respectively while their joint distribution is given by a Joe copula. If we use the inverse normal transformation in both cases above then we have generated pseudo random numbers with normal marginal distributions and a Joe copula as a joint distribution, if the inverse of the extreme value distribution is used instead then the pseudo random numbers have extreme value marginals with a Joe copula as a joint distribution.