HEDG Working Paper 07/23

Improving precision in cost-effectiveness analysis using copulas.

Casey Quinn

October 2007

ISSN 1751-1976

york.ac.uk/res/herc/hedgwp

Improving precision in cost-effectiveness analysis using copulas.

Casey Quinn

University of York, York YO10 5DD, England

Lehigh University, Bethlehem, PA 18015, USA

October 22, 2007

Abstract

A copula is best described, as in Joe (1997), as a multivariate distribution function that is used to bind each marginal distribution function to form the joint. The copula parameterises the dependence between the margins, while the parameters of each marginal distribution function can be estimated separately.

This is a brief introduction to copulas and multivariate dependence issues within a health economics context. The research presented here will make its own contributions to the development of copulas as a methodology, but more importantly will make deliberate inroads into health economic applications of copulas. To do this, common analytic problems faced by health economists are considered. Some of the differences between the copula methodology and existing alternatives are discussed, and a generalisable, systematic approach to estimation is provided.

JEL classification: C1, C3, C5, I3, I10

1 Introduction

Measuring cost-effectiveness typically involves relating the effectiveness of a treatment (the outcome) to its cost, usually with common health or economic comparators, such as comparing cost with some monetised value of a health gain from treatment. The predominant focus in cost-effectiveness analysis has been the Incremental Cost-Effectiveness Ratio (ICER), an average-in-differences approach to measuring costs and outcomes associated with competing interventions. Concerns over the robustness of ICERs under uncertainty however have led to more informative measures such as Net Monetary Benefit (NMB), which re-scales treatment cost and outcome to the same numeraire (Phelps and Mushlin 1991; Claxton and Posnett 1996; Stinnet and Mullahy 1998).

The use of NMBs provides information about dispersion and includes willingness-to-pay directly, whereas ICERs do not. It also allows regression-based estimation and covariate explanation of individual and environmental effects not related directly to the intervention of interest, but still determining individual-level costs and outcomes. This relaxes the assumption that the only difference is due to the intervention itself: Liu and Zhao 1999; Willan, *et al.* 2004; Willan, *et al.* 2005; Vázquez-Polo, *et al.* 2004). The ordinary simultaneous equations approach to regression-based NMBs however must specify a bivariate distribution for both cost and health outcome that does not contain the best-fitting distribution of each margin, because bivariate distributions are restricted in the marginal distributions available. An example of such an approach is the bivariate Normal distribution.

The normality assumption is generally inherent in analyses of jointly-distributed random variables, rather than for cost-effectiveness specifically, but robustness and efficiency remain a concern. Overcoming this often means estimation of one dependent variable conditional upon the other, or simulation methods when simultaneous estimation is preferred (Lambert and Vandenhende 2002; Lin 2003, respectively). To

the extent that conditional distributions are intractable, simulation will be required using that method also. Attempts to enhance cost-effectiveness analysis through regression can therefore be hampered by distributional mis-specification or functional intractability (Vázquez-Polo and Negrín-Hernández 2004; Briggs 2005). This is common in multivariate data analysis, but the disparate nature of treatment cost and outcome data specifically makes homogenising marginal distributions questionable.

This paper suggests an alternative approach: estimating costs and outcomes simultaneously using copulas. For univariate marginal distribution functions $F_1(x_1)$ and $F_2(x_2)$, a copula is a function that parameterises the dependence between the margins and binds those margins precisely, to form the multivariate distribution function (Smith 2003). The parameters of each marginal distribution function can be estimated separately. Depending on the functional form used, association of quite different types can either be assumed or tested, independently of the functional forms of the marginal distributions used.

This paper demonstrates and discusses regression with copulas using data from a clinical trial for hysterectomy. Procedures for testing goodness-of-fit and comparing different copulas are also discussed.

2 Regression and covariate-adjustment

In practice the true effectiveness of a given treatment is not known, since comparative treatments are submitted to separate samples, or at different times - the so-called 'evaluation problem' (Mullahy and Manning 1995; Hoch, *et al.* 2002). For some new treatment or technology to be compared to an existing one, differences in their mean costs are compared to differences in their mean health outcomes (see for example Phelps and Mushlin 1991; Claxton and Posnett 1996; Gold, *et al.* 1996; Stinnet and Mullahy 1998; Drummond, *et al.* 2005). During trials to obtain cost and outcome data however, imperfect randomisation, attrition and the limits that small samples impose

on reliable inference are common issues.[1] Treatment cost and outcome therefore cannot be assumed to depend only upon the intervention (Vázquez-Polo, *et al.* 2004).

Regression analysis of treatment cost and outcome has been suggested and demonstrated in, for example, Liu and Zhao (1999), Hoch, *et al.* (2002) and Willan, *et al.* (2004). Regression responds to the idea that there is no such thing as an 'average' patient, such that individual covariates are useful in generating cost and outcome estimates that are transferable as well as generalisable (Lessard 2007). Nixon and Thomas (2005) also do this using baseline covariates to identify subgroups in samples, as well as for multi-centre clinical trial data.

2.1 Simultaneous estimation of costs and outcomes

Within the NMB framework, treatment is usually indicated in a single regression model, averaging the effect of the covariates across trial arms, as well as collapsing cost and outcome into one distribution. A more flexible regression method is useful to capture more distributional information, as well as the association between jointly-dependent cost and outcome. Willan, *et al.* (2004) propose Seemingly Unrelated Regression (SUR) techniques for simultaneous estimation of cost and health outcomes, pointing out that this enables more rigorous analysis not only of covariates but also of differences between treatment arms or multiple trial centres using sub-groups. This is an approach favourably reviewed elsewhere (Briggs 2005; Vanness and Mullahy 2005).

Estimation using systems of equations for cost (C) and outcome (or effect, E) relies upon some association between the errors such that, in expectational terms,

$$\begin{pmatrix} \Delta\overline{C} \\ \Delta\overline{E} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_{\Delta C} \\ \mu_{\Delta E} \end{pmatrix}, \begin{pmatrix} \sigma^2_{\Delta\overline{C}} & \sigma_{\Delta\overline{C},\Delta\overline{E}} \\ \sigma_{\Delta\overline{C},\Delta\overline{E}} & \sigma^2_{\Delta\overline{E}} \end{pmatrix}\right) \tag{1}$$

unless independence is assumed for convenience, in which case $\sigma_{\Delta\overline{C},\Delta\overline{E}} = 0$. These five parameters $\mu_{\Delta C}$, $\mu_{\Delta E}$, $\sigma^2_{\Delta\overline{C}}$, $\sigma^2_{\Delta\overline{E}}$ and $\sigma_{\Delta\overline{C},\Delta\overline{E}}$ correspond also to average-effects

approaches to cost-effectiveness analysis, including the ICER. This specification is typically in the interests of practicable estimation. The implication though is that $\sqrt{N}\left(\left(\Delta\overline{C}-\mu_{\Delta C}\right),\left(\Delta\overline{E}-\mu_{\Delta E}\right)\right)' \xrightarrow{d} N\left(0,\Sigma\right)$ for some asymptotic covariance matrix Σ, however this is often not supported even by descriptive information.

A related problem arises with the use instead of conditional distributions to avoid imposing estimates of average effect on two distributions, such as different trial arms or different centres in a multi-centre trial. Nixon and Thomas (2005), for example, introduce correlation by parameterising variation in cost within the equation for the outcome. Conditional distributions are affected by the so-called Borel paradox (Kolmogorov 1950; Newey and Steigerwald 1997; Verhoeven and McAleer 2003). For practical purposes, this holds that the margins in a conditional distribution function are not 'swappable' - i.e. the conditional and conditioning distribution functions are not swappable, unless they are of the same, symmetric, family of distribution functions. Otherwise estimates of the parameters of the conditioning distribution function will not necessarily be consistent.

A simultaneous-equations approach that imposes no distributional restrictions is preferred, but which also measures dependence θ such that, between random variables (X_1, X_2), $\theta_{X_1,X_2} = \theta_{X_2,X_1}$ regardless of the form (or skew, or kurtosis) of the distributions $F_1(X_1)$ and $F_2(X_2)$.

2.2 Copulas

Copulas are functions that parameterise the dependence between univariate marginal distribution functions to form a joint distribution function. They represent an improvement in modelling costs and outcomes simultaneously in two ways: first, by enabling a range of distributions – almost any appropriate parametric or non-parametric distribution – to be given to each margin in a multivariate distribution; and secondly by allowing the association between the random variables in a multivariate distribu-

tion to be specified separately for each bivariate pair of marginal distributions. Thus each margin is precisely defined according to the nearest approximation to the data, as well as invariant to transformations in every other margin, or independent of the choice of every other marginal distribution.[2] Following this, the resulting multivariate distribution may be used to derive densities for estimation via Maximum Likelihood, for example.

A copula is principally a dependence function – each one represents a unique description of the relationship between its margins, while the distribution functions of its margins are assigned separately. Specifically, the distribution function $H(x_1, ..., x_n)$ of some set of random variables $\{X_1, ...X_n\}$ with univariate distribution functions $F_1(x_1), ..., F_n(x_n)$ is given by

$$H(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n); \theta) \tag{2}$$

where $C(.; \theta)$ is the copula, a function of n uniform margins whose association is represented by the parameter of association, θ. Returning to the strictly bivariate framework, Sklar's (1959) theorem holds that, for any bivariate distribution with given margins, there exists a copula that binds these margins to form the joint distribution precisely (Smith 2003). Further, the copula $C(F_1(x_1), F_2(x_2); \theta) = H(x_1, x_2)$ is unique when $F_1(x_1)$ and $F_2(x_2)$ are continuous.[3]

For estimation, the joint probability density function h of distribution H is given by

$$\begin{aligned} h(x_1, x_2) &= \frac{\partial^2 C(F_1(x_1), F_2(x_2); \theta)}{\partial x_1 \partial x_2} \\ &= f_1(x_1) f_2(x_2) C_{12}(F_1(x_1), F_2(x_2); \theta) \end{aligned} \tag{3}$$

and $C_{12}(F_1(x_1), F_2(x_2); \theta)$ is given by

$$C_{12}(u, v, \theta) = \frac{\partial^2 C}{\partial u \partial v} \tag{4}$$

I.e. the twice-differentiated copula with respective to its marginal CDFs, rather than the random variables as in Equation (3). Simplifying notation $F_1(x_1) \rightarrow u$, $F_2(x_2) \rightarrow v$ is used hereon.

In this manner, copulas separate the joint association of two or more random variables from their marginal distributions, since all the information on the dependence structure should be contained within the copula itself, through θ.

For this analysis, estimation is based on the linear-form Farlie-Gumbel-Morgenstern (FGM) copula, as well as the Frank family from the Archimedean class of copulas (Joe 1997; Nelsen 2006).

2.2.1 The Farlie-Gumbel-Morgenstern copula

The FGM is a relatively straightforward copula, easily implemented and suitable for comparison with the bivariate Normal. It is C such that

$$C(u, v; \theta) = uv(1 + \theta(1 - u)(1 - v)) \tag{5}$$

where $-1 \leq \theta \leq 1$, with positive and negative dependence for $\pm\theta$ respectively, and recalling that $F_1(x_1)$, $F_2(x_2)$ are (at least) monotonic.[4] In practical applications this copula has been shown to be a somewhat limited measure of dependence (Prieger 2000). Dependence $\theta \in [-1, 1]$ corresponds approximately to Spearman's correlation $\rho \in \left[-\frac{1}{3}, \frac{1}{3}\right]$, where $\rho = \frac{\theta}{\pi}$, and Kendall's $\tau \in \left[-\frac{2}{9}, \frac{2}{9}\right]$ where that $\tau = \frac{2\theta}{9}$ (Nelsen 2006). Mari and Kotz (2001) provide several extensions of the FGM copula, which expand this range. The FGM density is given by

$$\begin{aligned} C_{12}(u, v, \theta) &= \frac{\partial^2 C}{\partial u \partial v} \\ &= (1 + \theta(1 - 2u)(1 - 2v)) \end{aligned} \tag{6}$$

For estimation, the joint probability density function h of distribution H is derived according to Equations (3) and (4).

The resulting likelihood function for the FGM copula is simply

$$\begin{aligned} L\left(\beta_1, \beta_2, \sigma_1, \sigma_2, \theta\right) &= \prod_{i=1}^{n} f_1\left(x_{i1}; \beta_1, \sigma_1\right) f_2\left(x_{i2}; \beta_2, \sigma_2\right) \\ &\quad \times \left(1 + \theta\left(1 - 2F_1\left(x_{i1}; \beta_1, \sigma_1\right)\right)\left(1 - 2F_2\left(x_{i2}; \beta_2, \sigma_2\right)\right)\right) \end{aligned} \tag{7}$$

for margins characterised in regression by $F_1\left(x_{i1}; \beta_1, \sigma_1\right)$ and $F_2\left(x_{i2}; \beta_2, \sigma_2\right)$, with regression parameters β_1, σ_1 and β_2, σ_2.

2.2.2 The Frank copula

The Frank copula is an Archimedean-class copula. Archimedean copulas are a particular class of copula that includes several popular families, and which are fundamentally different to other families, including the FGM, by virtue of their construction (see Nelsen 2006). The Frank copula is is given by (Frank 1979),

$$C\left(u, v; \theta\right) = -\frac{1}{\theta} \ln\left(1 + \frac{\left(e^{-\theta u} - 1\right)\left(e^{-\theta v} - 1\right)}{e^{-\theta} - 1}\right) \tag{8}$$

where $\theta \in [-\infty, \infty] \setminus \{0\}$. This is a comprehensive family, such that association θ corresponds to $\tau \in [-1, 1] \setminus \{0\}$.[5]

2.2.3 Estimation

FIML estimation follows the same procedure for copulas as for ordinary FIML estimation. Location and scale parameters are estimated in each marginal distribution function (the functional form of each of which is selected separately from the others)

simultaneously with the copula parameters for dependence. Specifically, for some multivariate distribution function $H(X_1, .., X_n; \beta_1, .., \beta_n, \theta)$, consider the corresponding copula $C(F_1(X_1; \beta_1), .., F_n(X_n; \beta_n); \theta)$.

Step 1: Specify the functional forms of each marginal distribution $F_1(X_1; \beta_1), ...,$ $F_n(X_n; \beta_n)$, each with some vector of parameters β_i. This can be done parametrically (by prior FIML estimation of each margin, for example), or non-parametrically (Matlab, for example, has some distribution-fitting tools). Selection can also be made visually, or according to any other prior information.

Step 2: Specify the functional form of the copula, $C(F_1(X_1; \beta_1), .., F_n(X_n; \beta_n); \theta)$. This can be done according to some knowledge of the dependence structure (such as with examination of the variance-covariance matrix) or any characteristics desired of the joint distribution.

Step 3: Construct the copula density $c(F_1(X_1; \beta_1), .., F_n(X_n; \beta_n); \theta)$ according to Equations (**??**)-(**??**), as well as the likelihood and log-likelihood functions.

Step 4: The copula log-likelihood can be estimated according to any maximum-likelihood procedure. If point-estimates are available, for the parameters in either the copula or in the univariate marginal distribution functions, they should be given as precise starting values.[6]

3 Application: the eVALuate hysterectomy trial

The eVALuate hysterectomy trial was a multi-centre randomised trial comparing new laparoscopic procedures for hysterectomy with existing abdominal and vaginal procedures (Garry, *et al.* 2004).[7] The abdominal hysterectomy requires incision, and involves scarring, and more pain, morbidity and likely complications. It is however less technically demanding, and much more commonly undertaken, than the vaginal approach. The laparoscopic procedure uses keyhole surgical techniques, which can ei-

ther replace the abdominal incision or complement the otherwise non-surgical vaginal procedure. In terms of scarring and recovery it is an improvement over the abdominal approach, however it is more costly and time-consuming, as well as requiring technical training.

The eVALuate trial sought to determine the role for laparoscopic techniques before their wide introduction to clinical practice. It was conducted with twin arms: two parallel randomised trials in which patients were first allocated abdominal or vaginal procedures by their surgeon, and then randomly allocated either to that or to the Laparoscopic arm. The result is a trial comparing laparoscopic to abdominal procedures, and laparoscopic to vaginal. For this analysis only the Abdominal arm of the eVALuate trial was used.

3.1 Descriptive analysis

Some description and discussion of the trial data follows. Comprehensive description of the eVALuate trials can be found in the original Health Technology Assessment by Garry *et al.* (2004). The data and results presented here are only those of specific relevance to this analysis.[8]

3.1.1 Patient characteristics

Descriptive statistics concerning patient characteristics and general explanatory variables are contained in Table 3.1.

At the means, patients were sufficiently randomised according to Age and Body Mass Index. Smoking and having had pelvic surgery previously show some differences. These typically occurred because of movement during the trial, or initial allocation, both of which were subject to clinical indicators as well as the need for randomisation. The length of stay results reflect the theory behind the use of laparoscopic techniques, showing lower average length of stay and lower post-operative length of stay.

Table 1: SUMMARY STATISTICS (MEANS AND STANDARD DEVIATIONS, OR PERCENTAGES) OF EVALUATE PATIENT CHARACTERISTICS

	Abdominal	Laparoscopic
Mean (standard deviation)		
Age (years)	41.17	41.68
	(7.58)	(7.15)
Body Mass Index	25.93	26.58
	(5.42)	(5.06)
Length of Stay (days)	5.11	3.95
	(2.72)	(2.38)
Post-operative LOS (days)	4.43	3.4
	(2.49)	(2.57)
Percentage		
Whether current smoker	48.63	41.44
Whether had previous pelvic surgery	0.73	0.54
Sample size	320	573
Proportion of total (i.e. treatment dummy)	35.83	64.17

Table 2: TOTAL COSTS AND QALY GAIN PER PATIENT

	Abdominal	Laparoscopic
QALYs gained over year	0.8616	0.8703
	(0.1356)	(0.1312)
Cost (£)	1518.992	1675.741
	(1329.699)	(1220.504)

Whether or not the individual was a smoker at the time of the trial exhibits some interesting effects in the analysis. Having to stop smoking for the trial confounds, somewhat, the marginal effect of smoking on treatment outcome (for both procedures: the negative effect of being smoker is not distinguishable from the positive effect of having quit smoking, or at least stopped temporarily).

The difference in sample size between the Abdominal and Treatment groups was intentional: the trial was design to over-sample Laparoscopic patients to gather more information, as well as to expose surgeons to the new technology as much as possible.

3.1.2 Cost and Outcomes

Table 3.2 describes cost and outcomes data from the trial. In previous analyses of the eVALuate trial, both QALYs gained and major complications have been used as the health outcome (Garry, *et al.* 2004; Garry, *et al.* 2004b; Sculpher, *et al.* 2004). The outcomes of interest here are the total cost and the QALYs gained per patient.

Results from Kolmogorov-Smirnoff tests reject the hypothesis that Costs are distributed differently in each arm of the trial, but do not reject the same for QALYs gained.[9] This does not imply that Abdominal QALYs gained follow one family of distributions while Laparoscopic QALYs gained follow another; rather that they may both be Beta-distributed according to different parameters. That the result is mixed is not necessarily of concern: while Costs are given one distribution and QALYs gained

another - i.e. one copula, using a treatment indicator, or dummy variable - the trial arms could also be separated into an Abdominal Cost/QALY copula and a Laparoscopic Cost/QALY copula, but with a loss of power due to the sample sizes. Given the size and closeness of the means for QALYs gained in Table 3.2, separating the trial arms is not considered necessary.

Kernel densities and Quantile plots comparing the distributions of Treatment Cost and QALYs gained with the normal distribution are in Figures 3.1 and 3.2.

Table 3: ESTIMATES (STANDARD ERRORS) FROM REGRESSION OF TREATMENT COST AND QALYS GAINED, WITH TWO-SIDED SIGNIFICANCE INDICATED

	Cost (log-Normal)	QALYs gained (Beta)	QALYs gained (Normal)
Treatment	0.0868**	0.1585**	0.0081
	(0.0332)	(0.0748)	(0.0105)
Age at randomisation	0.0040*	0.0037	0.0006
	(0.0022)	(0.0050)	(0.0007)
BMI at randomisation	0.0062**	-0.0116*	-0.0010
	(0.0031)	(0.0069)	(0.0010)
Previous pelvic surgery	0.5753**	-1.0488**	-0.1456**
	(0.2097)	(0.4154)	(0.0666)
Previous smoker	0.0626*	0.3566**	0.0375**
	(0.0326)	(0.0738)	(0.0103)
Constant	7.4256**	0.6805	0.6967**
	(0.2328)	(0.4734)	(0.0739)
Std. Dev	0.3922	0.1376	0.1364
p-value	0.0918	0.2507	0.4739

* Significant at 10%

** Significant at 5%

3.2 Estimation of the margins

Table 3.3 contains results from regression of the Abdominal and Laparoscopic margins.

Tables 3.3 and 3.4 contain bootstrapped standard errors. Cost is Lognormally-distributed; QALYs gained is estimated twice, once according to assumed normality, and assuming Beta-distributed QALYs gained, to account directly for the skewness and bounded responses.[10] Estimates for QALYs gained, assuming normality, have been included with this comparison in mind.

The Beta distribution is, like the Normal, a versatile one, convenient for failures

Figure 1: KERNEL DENSITIES AND QUANTILE PLOTS FOR TREATMENT COST

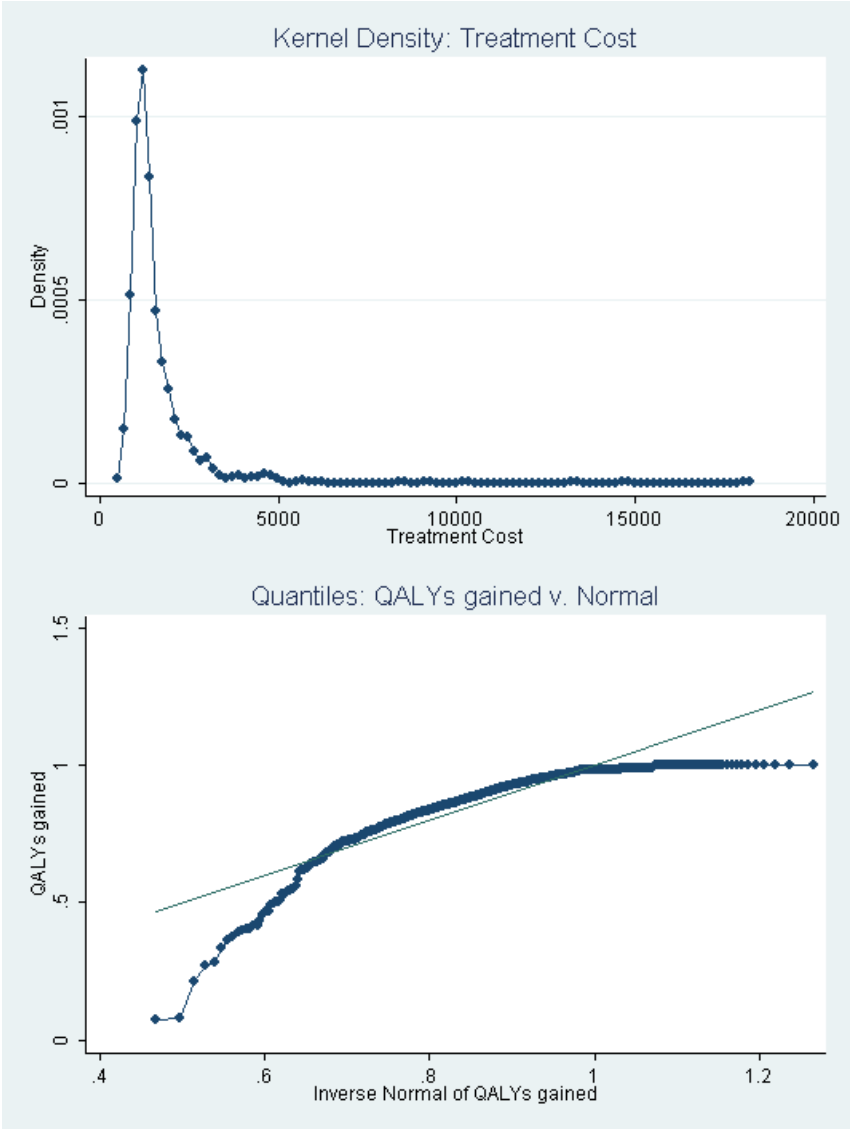
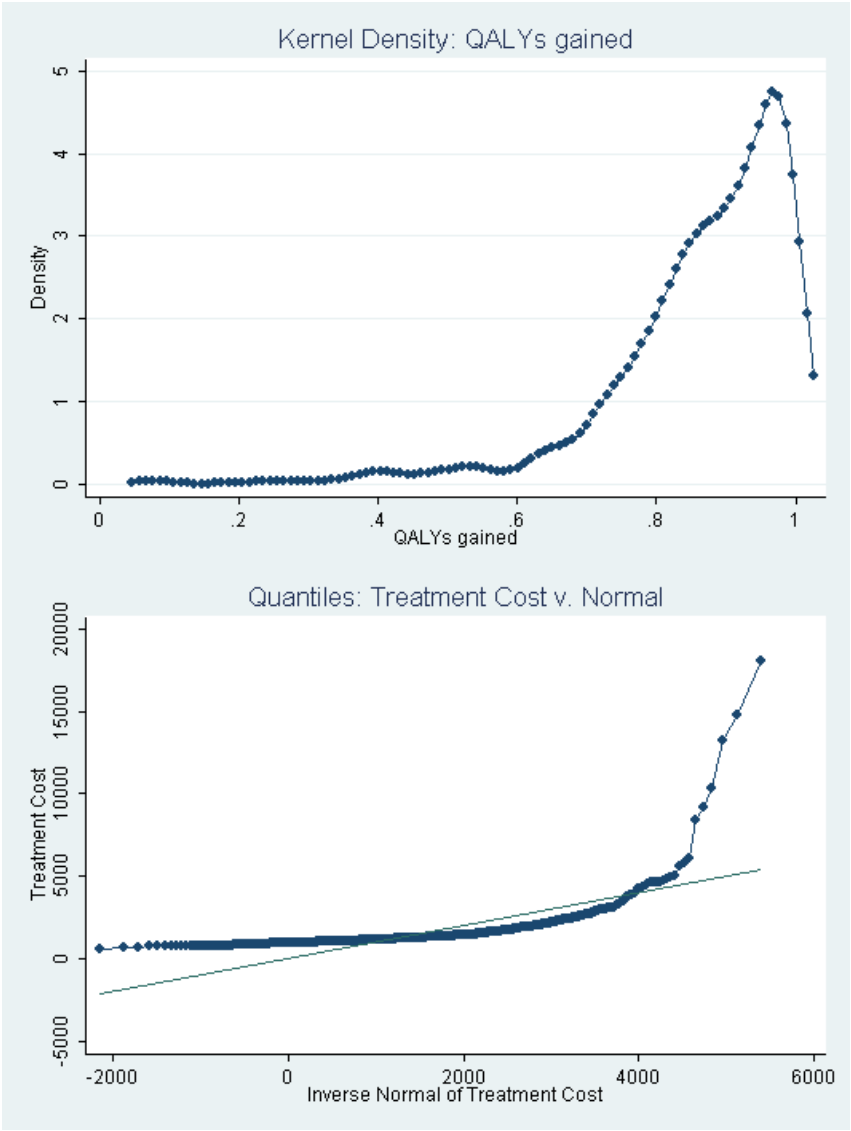


Figure 2: KERNEL DENSITIES AND QUANTILE PLOTS FOR QALYS GAINED



in the normality assumption such as skewness and multimodality. Beta regression however is rare, typically because interpretation is not as straightforward. In Tables 3.3 and 3.4, for example, the coefficients in the Beta regression refer not to the direct effect on QALYs gained, but to a log-odds ratio of the probability $\Pr(Y < y)$ with that covariate rather than without. This is due to the use of a logit link function to transform the mean μ, such that

$$\mu = \frac{\exp(x_i'\beta)}{1 + \exp(x_i'\beta)} \tag{9}$$

which re-scales the mean to the unit plane. Moreover the parameters ω and κ in $Y \sim Beta(\omega, \kappa)$ do not refer to the data directly: ω is not the mean and κ is not variance. The regression here follows the procedure of Smithson and Verkuilen (2006) quite closely, using the transformations

$$E(Y) = \frac{\omega}{\omega + \kappa} \tag{10}$$

and

$$\begin{aligned} Var(Y) &= \frac{\omega\kappa}{(\omega + \kappa)^2 (\omega + \kappa + 1)} \\ &= \frac{\mu(1-\mu)}{(\omega + \kappa + 1)} \end{aligned} \tag{11}$$

where $\mu = E(Y)$. This allows variance and standard deviation to be determined for comparison. Standard deviation of the QALYs gained margin in Table 3.3, for example, has been calculated according to Equation (11).[11]

The independent marginal models suggest the Beta distribution provides a better fit of QALYs gained, compared to the Normal distribution. Covariate explanation of QALYs gained, in particular, is poor across both specifications with the information available. The treatment dummy, representing an incremental increase in QALYs

gained due to treatment, is not statistically significant for the normally-distributed QALYs model.

3.3 Estimation of the joint distribution

Results from regression of the joint distribution are in Table 3.4.

Table 4: ESTIMATES (STANDARD ERRORS) FROM COPULA AND SUR REGRESSION OF TREATMENT COST AND QALYS GAINED, WITH TWO-SIDED SIGNIFICANCE INDICATED

	FGM		Frank		Product		BVN	
	Cost	QALYs gained	Cost	QALYs gained	Cost	QALYs gained	Cost	QALYs gained
Treatment	0.0825**	0.1561**	0.0789**	0.1195	0.0868**	0.1585**	0.0868**	0.0081
	(0.0333)	(0.0748)	(0.0347)	(0.0767)	(0.0332)	(0.0748)	(0.0332)	(0.0105)
Age at randomisation	0.0038*	0.0035	0.0024	0.0043	0.0040*	0.0037	0.0040*	0.0006
	(0.0022)	(0.0050)	(0.0023)	(0.0051)	(0.0022)	(0.0050)	(0.0022)	(0.0007)
BMI at randomisation	0.0064**	-0.0114*	0.0070**	-0.0071	0.0062**	-0.0116*	0.0062**	-0.0010
	(0.0031)	(0.0069)	(0.0033)	(0.0072)	(0.0031)	(0.0069)	(0.0031)	(0.0010)
Previous pelvic surgery	0.6161**	-1.1094**	0.5813**	-0.9183**	0.5753**	-1.0488**	0.5753**	-0.1456**
	(0.2026)	(0.3800)	(0.2150)	(0.3537)	(0.2097)	(0.4154)	(0.2097)	(0.0666)
Previous smoker	0.0555*	0.3764**	0.0559	0.3421**	0.0626*	0.3566**	0.0626*	0.0375**
	(0.0326)	(0.0734)	(0.0344)	(0.0751)	(0.0326)	(0.0738)	(0.0326)	(0.0103)
Constant	7.4708**	0.5815	7.4427**	0.4990**	7.4256**	0.6805	7.4256**	0.6967**
	(0.2270)	(0.4416)	(0.2440)	(0.4198)	(0.2328)	(0.4734)	(0.2328)	(0.0739)
Association parameter	-0.7488**		-6.0078**		n/a		-0.2551**	
	(0.1164)		(0.1995)				(0.0348)	
Log-likelihood	401.648		183.727		384.411		64.958	

* Significant at 10%

** Significant at 5%

The FGM, Frank and Product copulas contain Lognormally-distributed treatment cost and Beta-distributed QALYs gained. The bivariate Normal model (BVN: SUR with Maximum Likelihood) contains Lognormally-distributed treatment cost and normally-distributed QALYs gained. The Product copula is merely a joint distribution assuming independence: the product of each marginal distribution.

Neither the Frank copula nor the bivariate Normal SUR model return a statistically significant effect of treatment on QALYs gained, although all models do for Costs. These are the mean effects of treatment on Costs and QALYs gained, controlling for available individual characteristics. In the BVN model in particular the point-estimate of treatment effect is poor, statistically. Subject to evaluation of the relative performance of the copula models, the evidence of the efficacy of the laparoscopic procedure is mixed, under regression.

An interesting result not presented here is that the standard errors in the copula models tended to shrink asymptotically, while the opposite occurred with those of the SUR model, suggesting greater relative asymptotic efficiency due to the copula method (Joe 2005).

3.4 Copula selection and goodness-of-fit

Following Joe (1997), two other approaches can be taken. The first is to use either the log-Likelihood directly, or information criteria such as the Akaike Information Criterion (AIC), given by $AIC = 2k - 2\ln(L)$ for log-likelihood L and k free parameters, or Bayesian Information Criterion (BIC), given by $k\ln(n) - 2\ln(L)$ and where n is the sample size. Models do not need to be nested for this comparison; with each copula model, as well as the bivariate normal, containing an equal number of free parameters, the punitive approach taken towards parameterisation is also not necessary. The convience of having immediate access to the log of the Likelihood function, post-estimation, is an advantage over other methods.

Table 5: LOG-LIKELIHOODS AND INFORMATION CRITERIA FROM COPULA AND SUR REGRESSION

	FGM	Frank	Product	BVN
Log-Likelihood	401.648	183.727	384.41	64.958
Akaike IC	-779.296	-343.454	-744.820	-105.916
Bayes IC	-724.295	-288.453	-689.819	-50.915

Results from comparing information criteria are in Table 3.5.

Figure 3.3 illustrates this. It shows the bivariate density of each model, superimposed over the observed spread of Costs and QALYs gained.

As Figure 3.2 illustrates, it is the FGM and Product copulas that best approximate the Beta-distributed QALYs gained within the joint distribution, supporting the results from comparison of the log-likelihoods and information criteria. The Frank copula would perform better in cases where there were more extreme values in the distribution of Costs, as its bivariate distribution can be seen to be pulled away by the longer tail in the log-normally distributed Costs.

Finally, the regression-based estimates of treatment effects, along with the coefficients, can be used to estimate cost thresholds for cost-effectiveness (Hoch, *et al.* 2002). This is done by constructing point-estimates of both treatment and non-treatment Costs and QALYs gained, as well as Incremental Cost-Effectiveness Ratios (ICERs), for each individual. The ICER is given by

$$\widehat{ICER}_i = \frac{\widehat{\Delta C}_i}{\widehat{\Delta E}_i} \tag{12}$$

where ΔE_i and ΔC_i represent the individual-level incremental changes in QALYs gained and Costs, respectively. The average of these

$$\overline{ICER} = \frac{\overline{\Delta C}}{\overline{\Delta E}} \tag{13}$$

provides the willingness-to-pay (for a QALY gain) threshold at which, according to a given regression model, a treatment will be cost-effectiveness. Results from this comparison, for all models, are in Table 3.6.

The estimated cost per QALY from the original Health Technology Assessment of Garry, *et al.* (2004) was £26,571, to which the estimate from the bivariate normal SUR model is statistically near.[12] The estimated cost per QALY from the preferred FGM copula is only £6,981, while that of the Frank and Product copulas are nearly as low. The suggestion from these is that, if the assumption that QALYs gained are in fact Beta-distributed, rather than normally distributed, cost-effectiveness is actually achieved at a much lower willingness to pay.

Figure 3: BIVARIATE COPULA DENSITIES, MARGINAL HISTOGRAMS AND BIVARIATE SCATTERPLOTS FOR OBSERVED INDIVIDUAL COSTS AND QALYS GAINED

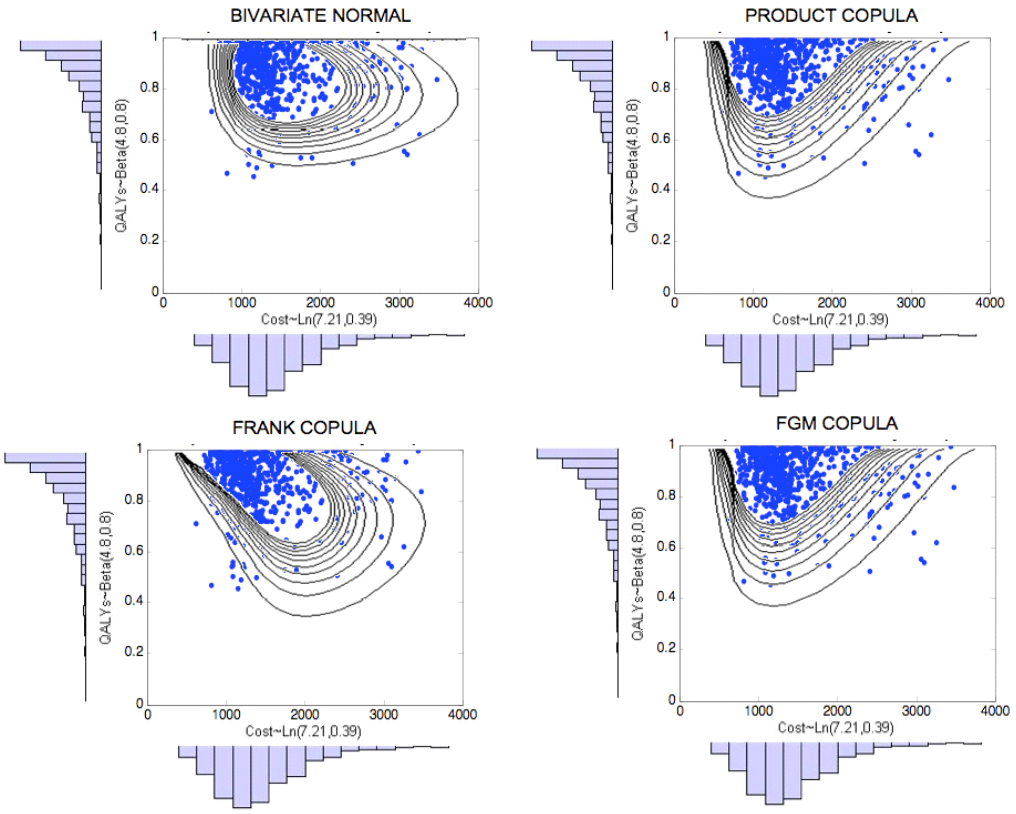


Table 6: ESTIMATED MEAN WILLINGNESS-TO-PAY THRESHOLDS (STANDARD DEVIATIONS) FOR COST-EFFECTIVENESS OF LAPAROSCOPIC HYSTERECTOMY

	Average Cost-Effectiveness Threshold (£)
FGM	6981.2810
	(812.0554)
Frank	8485.4860
	(937.0617)
Product	8120.5680
	(1131.9380)
BVN	29822.9100
	(2507.4270)

The difference between the cost-effectiveness thresholds, while substantial, are explained by the estimated treatment effect on QALYs gained and, to some extent, are an artefact of the trial itself. Because the trial followed patients for only a single year, QALYs gained cannot exceed 1, and the incremental difference due to treatment is very small (only 0.0087, from Table 3.2; the estimated effect in the SUR model also was only 0.0081). As a result the ICER is very sensitive to differences in these predicted treatment effects. On average the copula models predict an incremental difference in QALYs gained of 0.03 - significantly higher than that of the SUR, although their estimates of Cost are not significantly different. The suggestion in this paper, though, is that the copula estimates are more accurate.

4 Discussion

The results from this analysis suggest that laparoscopic hysterectomy is cost-effective at a significantly lower willingness-to-pay per QALY gain than has been found in

previous research. Some caution should be attached to this conclusion, however, as the value of the information gained from the eVALuate trial has not been demonstrated. As well as power issues, the treatment effect on the health outcome was not statistically significant in the bivariate normal SUR model. The explanatory variables included in this model also were not generally statistically significant, and were few in number.

In terms of demonstrating the contribution of the copula method to cost-effectiveness analysis, the illustrating example of the eVALuate trial was sufficient. In terms of demonstrating cost-effectiveness of the laparoscopic procedure itself, however, these results should be considered, at best, incomplete.

5 Conclusion

This paper compares two regression methods for cost-effectiveness analysis with clinical trial data, where individual covariates are used to adjust for differences at baseline. The two methods were Seemingly Unrelated Regression and the copula. The results suggest that the copula can improve estimation of the treatment effect, relative to the Seemingly Unrelated Regression method. The improvement due to the copula stems from the ability to assign any marginal distribution within its bivariate joint distribution, so that the non-normally distributed health outcome in the clinical trial is more accurately estimated. As a result, estimates of the treatment effect are more precise, as well as more reliable.

Like the Seemingly Unrelated Regression method, copulas allow feasible estimation of either Incremental Cost Effectiveness Ratios or Incremental Net Benefit at an individual level, using both the effect of treatment and covariates. A future extension of this paper is to compare regression methods according to estimation of counter-factual trial outcomes, which may provide more reliable estimates of individual Incremental Net Benefit than, for example, subgroup analysis.

Finally, some consideration must be given in future work to the uncertainty of regression estimates. The use of the Cost-Effectiveness Acceptability Curve allows researchers to assign a probability of making the correct decision, at a given willingness-to-pay, when comparing only the means of cost and outcome in a clinical trial. Regression allows estimates of the means to be adjusted for baseline differences, but no natural method for considering the value of the information used to construct those estimates.

Acknowledgements

This paper is a revised version of an earlier HEDG Working Paper (Quinn 2005). The advice of Andrea Manca, Andrew Jones, Nigel Rice, Luigi Siciliani and Stefano Conti was appreciated greatly. I also thank Health, Econometrics and Data Group at the University of York, and the Team for Economic Evaluation and Health Technology Assessment at the Centre for Health Economics, for their guidance. Finally comments by two anonymous reviewers. Financial support by the Centre for Health Economics is acknowledged gratefully .

Notes

[1]Meaning statistical inference for a population, using results from the analysis of sample data.

[2]I.e if some copula is defined $C(u,v)$, then $C(u,v) = C(u,F(v)) = C(u,G(v))$ where $F(v)$ and $G(v)$ could be two different distribution functions of v.

[3]Under discontinuities in $F(x_1), F(x_2)$, the copula C is otherwise determined on $RanF_1 \times RanF_2$, a combination of the range of the margins. Smith (2003) points out that the region outside this is usually not of interest: a multivariate distribution is typically defined within the supports of each margin.

[4]This is trivial: since $u = F_1(x_1), v = F_2(x_2)$ are univariate distribution functions they must be monotonic, at least (Nelsen 2006 contains an explanation of quasi-inverses of non-strictly increasing margins, which can also be used to construct a copula). This property is also necessary to ensure the measure of association, θ, 'obeys' the rules for measures of dependence.

[5]Algebraically, the Frank copula does not nest independence because of the term $\frac{1}{\theta}$. Nelsen (1998), however, demonstrates that $\lim_{\theta \to 0} C_{Frank} = uv$, i.e. the Product Copula.

[6]FIML procedures written for STATA are available online at http://www.york.ac.uk/res/herc/hedg_stata.html. They should be informative, but are specific to this particular analysis.

[7]My use of the eVALuate trial data was done with the kind permission of the team involved with the original Health Technology Assessment, for which I am grateful.

[8]Data were further cleaned using the Grubbs procedure for detecting outliers in Stata.

[9]Tests for QALYs gained did not reject the null of equality with p-values of 0.119 for the Abdominal trial arm. Tests for costs rejected the null with p-values <0.0001, as did all tests for equality between cost and QALYs gained.

[10]Treatment Cost is defined from zero: it is not censored. The Cost data have not been shifted or otherwise adjusted, besides the log-scale transformation.

[11]Code for estimation of the Beta regression model using Stata's ml package is available from the author.

[12]The Garry, *et al.* (2004) estimate of £26,571 is within the 95% confidence interval around the bivariate normal SUR estimate of £29,822.

REFERENCES

Briggs, A., 2005. Statistical methods for cost-effectiveness analysis alongside clinical trials, in: Jones, A. M., ed. Elgar Companion to Health Economics. Edward Elgar Publishing, Cheltenham.

Claxton, K., 1999. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies, Journal of Health Economics, 18: 341–364.

Claxton, K., Posnett, J., 1996. An economic approach to clinical trial design and research priority-setting. Health Economics, 5: 513-524.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., Stoddart, G. L., 2005. Methods for the economic evaluation of health care programmes. 3rd Edition. Oxford University Press, Oxford.

Frank, M.J., 1979. On the simultaneous associativity of F(x,y). and x + y – F(x,y). Aequationes Mathematica, 19: 194–226.

Garry, R., Fountain, J., Brown, J., Manca, A., Mason, S., Sculpher, M., Bridgman, S., Gray, J., Lilford, R., 2004. eVALuate hysterectomy trial: a multi-centre randomised trial comparing abdominal, vaginal and laparoscopic methods of hysterectomy. Health Technology Assessment, 826.: 1–154.

Garry, R., Fountain, J., Mason, S., Napp, V., , Hawe, J., Clayton, R., Abbott, J., Phillips, G., Whittaker, M., Lilford, R., Bridgman, S., Brown, J., 2004b. The eVALuate study: two parallel randomised trials, one comparing laparoscopic with abdominal hysterectomy, the other comparing laparoscopic with vaginal hysterectomy. British Medical Journal, 328: 129.

Gold, M.R., Siegel, J.E., Russell, L.B., Weinstein, M.C., 1996. Cost-effectiveness in Health and Medicine. Oxford University Press, New York.

Hoch, J.S., Briggs, A.H., Willan, A., 2002. Something old, something new, some-

thing borrowed, something BLUE: a framework for the marriage of health econometrics and cost-effectiveness analysis. Health Economics, 11: 415–430.

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman and Hall, London.

Joe, H., 2005. Asymptotic efficiency of the two-stage estimation method for copula-based models, Journal of Multivariate Analysis. 942: 401-419.

Lambert, P., Vandenhende, F., 2002. A copula based model for multivariate non normal longitudinal data: analysis of a dose titration safety study on a new antidepressant. Statistics in Medicine, 21: 3197–3217.

Lessard, C., 2007. Complexity and reflexivity: two important issues for economic evaluation in health care. Social Science and Medicine forthcoming.

Lin, D.Y., 2002. Regression analysis of incomplete medical cost data. Statistics in Medicine, 227.: 1181–200.

Liu, G., Zhao, Z., 1999. Stochastic cost-effectiveness analysis: a simultaneous marginal-effect approach. Value in Health, December: 420-428.

Mari, D.D., Kotz, S., 2001. Correlation and dependence. Imperial College Press, London.

de Matteis, R., 2001. Fitting copulas to data. Diploma thesis. University of Zurich, Zurich.

Mullahy, J., Manning, W., 1995. Statistical issues in cost-effectiveness analyses. In: Valuing Health Care: Costs, Benefits, and Effectiveness of Pharmaceuticals and Other Medical Technologies, Sloan, F., ed. Cambridge University Press, New York.

Nelsen, R. B. 2006. An Introduction To Copulas. 2nd Ed. Springer Verlag, New York.

Newey, W., Steigerwald, D., 1997. Asymptotic bias for quasi maximum likelihood estimators in conditional heteroskedasticity models. Econometrica, 65: 587–599.

Nixon, R.M., Thomas, S.G., 2005. Methods for incorporating covariate adjustment, subgroup analysis and between-centre differences into cost-effectiveness evaluations. Health Economics, 1412:1217-29

Nixon, R.M., Wonderling, D., Grieve, R., 2005. How to estimate cost-effectiveness acceptability curves, confidence ellipses and incremental net benefits alongside randomised controlled trials, technical report of the Medical Research Council Biostaticis Unit, Cambridge.

O'Hagan, A., Stevens, J.W., 2002. The probability of cost-effectiveness. BMX Medical Research Methodology, 2: 5-10.

Phelps, C. E., Mushlin, A., 1991. On the near equivalence of cost-effectiveness and cost-benefit analysis. International Journal of Technology Assessment in Health Care, 71: 12-21.

Prieger, J.E., 2000. A flexible parametric selection model for non-normal data with application to health care usage, Journal of Applied Econometrics 174: 367-392.

Quinn, C., 2005. Generalisable regression methods for cost-effectiveness using copulas. Working Paper 05/13 of the Health, Econometrics and Data Group at the University of York, York.

Resnick, S. I., 1987. Extreme Values, Regular Variation, and Point Processes. Springer Verlag, New York.

Sculpher, M., Manca, A., Abbott, J., Fountain, J., Mason, S., Garry, R., 2004. Cost effectiveness analysis of laparoscopic hysterectomy compared with standard hysterectomy: results from a randomised trial. British Medical Journal, 328:134.

Sklar, A., 1959. Fonctions de répartition à n dimensions et leur marges. Publications of the Institute of Statistics. University of Paris. 8: 229-231.

Smith, M. D., 2003. Modelling sample selection using archimedean copulas. Econometrics Journal, 6: 99–123.

Smithson, M., Verkuilen, J., 2006. A better lemon-squeezer? Maximum-likelihood regression with Beta-distributed dependent variables. Psychological Methods, 111: 54-71.

Stinnet, A., Mullahy, J., 1998. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Medical Decision Making, 18: S68–S80.

van Hout, B.A., Al, M.J., Gordon, G.S., Rutten, F.F., 1994. Costs, effects and C/E-ratios alongside a clinical trial. Health Economics, 35: 309-319.

Vanness, D.J., Mullahy, J., 2005. Perspectives on mean-based evaluation of health care, Jones AM ed. Elgar Companion to Health Economics. Cheltenham, UK: Edward Elgar Publishing, forthcoming.

Vázquez-Polo, F.J., Negrín Hernández, M.A., González López-Valcárcel, B., 2004. Using covariates to reduce uncertainty in the economic evaluation of clinical trial data. Health Economics, 146: 545–557.

Verhoeven, P., McAleer, M., 2003. Fat tails and asymmetry in financial volatility models, Centre for International Research on the Japanese Economy Discussion Paper F-211, Centre for International Research on the Japanese Economy, The University of Tokyo, Tokyo.

Willan, A.R., Briggs, A.H., Hoch, J.S., 2004. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. Health Economics, 13: 461–475.

Willan, A.R., Lin, D.Y., Manca, A., 2005. Regression methods for cost-effectiveness analysis with censored data. Statistics in Medicine, in press.