Work in progress: Please do not quote without permission

Something old, something new, something borrowed, something BLUE: A framework for the marriage of health econometrics and cost-effectiveness analysis

Andrew H. Briggs[1]
Jeffrey S. Hoch[2]
Andrew R. Willan[3]

[1] Health Economics Research Centre, University of Oxford
[2] Department of Epidemiology and Biostatistics, University of Western Ontario
[3] Department of Clinical Epidemiology and Biostatistics, McMaster University

Abstract

Economic evaluation is often seen as a branch of health economics divorced from mainstream econometric techniques with a reliance instead on statistical methods for designed experiments. Furthermore, the statistic of interest in cost-effectiveness analysis, the incremental cost-effectiveness ratio is not amenable to regression-based methods, hence the traditional reliance on aggregating results across the arms of a clinical trial. In this paper, we explore the potential for health economists undertaking cost-effectiveness analysis to exploit the plethora of established econometric techniques through the use of the net-benefit framework—a recently suggested reformulation of the cost-effectiveness problem that avoids the reliance on cost-effectiveness ratios and their associated statistical problems. This allows the formulation of the cost-effectiveness problem within a standard regression type framework. We provide an example with empirical data to illustrate how a regression type framework can enhance the net-benefit method. We go on to suggest that this approach will encourage researchers to address, in a statistically robust way, the underlying economics of cost-effectiveness analysis in the sense of exploring the importance of covariates on the marginal cost-effectiveness of an intervention.

Paper presented to the Health Economists' Study Group Conference, January 3-5 2001, Keble College, Oxford.

1. Introduction

The development of applied health economics has progressed along two broad paths. The traditional path sees applied health economics undertaken in departments of economics, employing the methods of applied econometrics to model health related. The second way in which health economics has devebped has been in the economic evaluation of health care technologies, with health economists undertaking such evaluations often employed in a multidisciplinary setting, working alongside clinicians, statisticians, epidemiologists and trialists in order to produce information on the cost-effectiveness of interventions. It is perhaps of little surprise therefore that the development of economic evaluation alongside clinical trials owes more to medical statistics than to econometrics, the former being characterised by methods surrounding experimental design (Armitage & Berry, 1994), while the latter typically involves information obtained from population surveys (Jones, 2000). Furthermore, cost-effectiveness analysis has traditionally been concerned with the estimation of cost-effectiveness ratios—which are not amenable to regression analysis, the mainstay of methods employed by applied econometricians.

In this paper, we explore the potential for health economists undertaking cost-effectiveness analysis to exploit the plethora of established econometric techniques through the use of the net-benefit framework—a recently suggested reformulation of the cost-effectiveness problem that avoids the reliance on cost-effectiveness ratios and their associated statistical problems. This approach also allows the formulation of the cost-effectiveness problem within a standard regression type framework. We go on to suggest that this formulation will encourage researchers to address, in a statistically robust way, the underlying economics of cost-effectiveness analysis in the sense of exploring the importance of covariates on the marginal cost-effectiveness of an intervention. All too often within the context of a clinical trial, the potential for cost-effectiveness to vary at the margin is obscured by aggregating across the arms of the trial.

For many years, the recommendation for analysts conducting economic evaluation of health care interventions has been to calculate *incremental* cost-effectiveness ratios (ICERs) in order to summarise the value for money of interventions (Weinstein and Stason, 1977; Weinstein and Fineberg, 1981; Drummond et al, 1987). More recently, as economic evaluations have begun to be conducted prospectively alongside clinical trials, the statistical problems associated with ratio statistics have become apparent in the interpretation of sampling uncertainty in the ICER. The non-zero probability that the ICER's denominator will take values very near zero causes instability in estimates of the ICER's value and variance, since the moments of the statistic do not exist. Furthermore, the non-linearity of the ICER means that it is biased (although consistent) estimator, which suggests that bias may be a cause for concern in small samples (Stinnett, 1996).

Ambiguity arises with ICERs, since the value of the ratio itself is not sufficient to give unequivocal treatment recommendations. A negative ICER is consistent with either a more expensive, less effective treatment or a less expensive, more effective treatment. A positive ICER indicates either a more expensive, more effective treatment or a less expensive, less effective treatment. While it is trivial to check which situation has produced the ICER estimate, great care must be taken when constructing its confidence interval so that ratios with similar signs but different interpretations are not grouped together. The analyst who chooses to bootstrap the ICER's confidence interval—a very popular method for handling uncertainty in stochastic cost-effectiveness analysis—may choose to reorder the bootstrap replicates so that ordering reflects the decision-making implications of the replicate. Unfortunately, the negative ICERs do not obey the law of transitivity, such that an unambiguous preference ordering is not possible in the Southeast and Northwest quadrants. Hence, the construction of confidence intervals for cost-effectiveness ratios when uncertainty covers more than one quadrant of the cost-effectiveness plane can be problematic.

Recently, a new framework for cost-effectiveness analysis has been suggested; the net-benefit framework (Stinnett & Mullahy, 1998; Tambour et al, 1998) reformulates the cost-effectiveness problem in linear form to generate a net-benefit statistic with more attractive statistical properties than the ICER making it the preferred choice for handling uncertainty in stochastic cost-effectiveness analysis (O'Brien et al., 1994; Chaudhary and Stearns, 1996; Polsky et al., 1997; Briggs and Fenn, 1998). In this paper, we argue that the linear nature of the net-benefit statistic also allows the analyst to enhance economic evaluation by employing regression methods. Practical advantages include being able to identify important subgroups, adjust for imperfect randomisation and make use of established econometric techniques. The remainder of this paper is organised into four parts. The next section provides a statistical motivation for the use of regression methods in the net-benefit framework. An empirical example of the procedure is then provided in Section 3. The remaining sections offer a discussion and suggest implications of the main results of this paper.

2. Statistical considerations in cost-effectiveness analysis

This section begins by specifying the usual approach of estimating an ICER in cost-effectiveness analysis, together with an outline of the challenges faced when analysing ICERs in a stochastic framework. The net-benefit framework is then introduced as a solution to these problems and a regression framework for net-benefits is outlined in order to demonstrate how cost-effectiveness can be estimated from a single regression equation.

2.1 The incremental cost-effectiveness ratio

In the most common case, an economic analysis involves an evaluation of an intervention treatment (T_1) compared to a standard care treatment (T_0). Denoting the expected values of cost and effect for treatment T_j (for j = 0 or 1) as m_{Cj} and m_{Ej} respectively, the incremental cost-effectiveness ratio comparing T_1 to T_0 is defined as

$$ICER = \frac{m_{C1} - m_{C0}}{m_{E1} - m_{E0}} = \frac{m_{\Delta C}}{m_{\Delta E}} \tag{1}$$

with the implication that the intervention offers good value for money if the ICER is below some maximum willingness to pay for health gain, i.e. that a decision should be made to implement the more costly, but more effective treatment intervention if

$$\frac{m_{\Delta C}}{m_{\Delta E}} < l \ , \tag{2}$$

where l is the maximum acceptable willingness to pay per unit of health gain (or ceiling ratio).

Of course, it is never possible to know the true incremental costs and true incremental effects of an intervention, since it is impossible to simultaneously observe the costs and effects of two different treatments in the same population of patients (Mullahy & Manning, 1995). Using sample data for economic data collected in a clinical trial setting and the analogy principle (Goldberger, 1964; Manski, 1988), it is possible to estimate the true, but unobservable ICER by

$$\widehat{ICER} = \frac{\overline{C}_1 - \overline{C}_0}{\overline{E}_1 - \overline{E}_0} = \frac{\Delta\overline{C}}{\Delta\overline{E}}$$

using the sample mean costs $\overline{C}_j$ and sample mean effects $\overline{E}_j$ for the j=0,1 treatment arms.

Introductory textbooks emphasis the importance of taking an incremental approach (Weinstein and Fineberg, 1981; Drummond et al, 1987) rather than comparing average cost-effectiveness ratios and a more recent contribution to the literature has highlighted the fundamental problem of taking patient-level average ratios: that the mean of ratios is not equal to the ratio of the means (Stinnett & Paltiel, 1996). The consequence is that

$$\frac{\overline{C}_1}{\overline{E}_1} - \frac{\overline{C}_0}{\overline{E}_0} \neq \frac{\overline{C}_1 - \overline{C}_0}{\overline{E}_1 - \overline{E}_0}$$

that is, that the incremental ratio cannot be constructed from the difference between the average cost-effectiveness ratios in each arm of the trial.

The estimated cost and effect differences, $\Delta\overline{C}$ and $\Delta\overline{E}$, can be plotted on the cost effectiveness plane (Anderson & Camm, 1986; Black, 1990). A set of four potential outcomes relating to the signs of $\Delta\overline{C}$ and $\Delta\overline{E}$ can be specified

$$Q \equiv \left\{\left(\Delta\overline{E} > 0, \Delta\overline{C} > 0\right); \quad \left(\Delta\overline{E} > 0, \Delta\overline{C} < 0\right); \quad \left(\Delta\overline{E} < 0, \Delta\overline{C} > 0\right); \quad \left(\Delta\overline{E} < 0, \Delta\overline{C} < 0\right)\right\}$$

which correspond to the four quadrants of the cost-effectiveness plane (see Figure 1).

Under the assumption that $\Delta\overline{C}, \Delta\overline{E} \neq 0$, the ICER maps this set of four potential outcomes to either a positive or a negative number. This mapping can lead to confusion since ICERs with the same sign, but in different quadrants of the CE plane, have diametrically opposed implications for decision-making. Negative ICERs occur when $\Delta\overline{C} < 0$ and $\Delta\overline{E} > 0$ (the southeast quadrant of the plane) or when $\Delta\overline{C} > 0$ and $\Delta\overline{E} < 0$ (the northwest quadrant of the plane); in the former, the new treatment dominates, while in the latter case the existing treatment dominates – thus the meaning of a negative ICER is equivocal. Similarly, the meaning of a positive ICER is also equivocal. Positive ICERs occur if $\Delta\overline{C} > 0$ and $\Delta\overline{E} > 0$ (the northeast quadrant) or if $\Delta\overline{C} < 0$ and $\Delta\overline{E} < 0$ (the southwest quadrant). The decision rule of Equation 2 above implicitly relates to the more expensive, more effective therapy, however, it is important to note in terms of the CE plane that while this decision rule is correct for the northeast quadrant (the new treatment is cost-effective if $\Delta\overline{C}/\Delta\overline{E} < l$) in the southwest quadrant the decision rule is reversed such that the new treatment (which is less costly and less effective than the existing treatment) is considered cost-effective only if $\Delta\overline{C}/\Delta\overline{E} > l$.

This emphasises how the ICER statistic does not give sufficient information for decision-making without knowledge of the quadrant of the CE plane (or the equivalently the sign of the numerator or denominator of the ratio). This is of particular concern when attempting to calculate confidence limits for cost-effectiveness ratios where uncertainty covers more than one quadrant of the CE plane. For example, the bootstrapping method has been widely applied to the problem of ICER confidence interval estimation and this method produces multiple 'replications' of the cost and effect differences. If these replications are used to calculate ICERs unthinkingly without knowledge of where on the plane the replicates fall then negative and positive ICERs from different quadrants could become conflated and the resulting confidence limit estimates could be misleading. Therefore, confidence intervals for ICERs should only be used where uncertainty is restricted to (one of) the positive quadrants of the CE plane. Where uncertainty covers more than one quadrant of the plane, the cost-effectiveness acceptability curve approach to summarising uncertainty should be employed (van Hout et al, 1994; Briggs & Fenn, 1998).

2.2 The net-benefit framework

Very recently, two papers have highlighted the 'net-benefit' approach to handling uncertainty in cost-effectiveness analysis (Tambour et al. 1998; Stinnett and Mullahy, 1998). The decision rule of Equation 2 can be rearranged to give a measure of 'net-benefit' and an associated decision rule that the new therapy should be implemented only if the net-benefits are positive. Two alternative formulations of net-benefit have been suggested based on a simple rearrangement of Equation 1, such that the new therapy should be implemented over the existing treatment if:

$$\text{NMB} = l \cdot m_{\Delta E} - m_{\Delta C} > 0 \tag{3}$$

or, equivalently, if:

$$\text{NHB} = m_{\Delta E} - m_{\Delta C} \big/ l > 0 \tag{4}$$

In Equation 3, the cost scale is employed giving net-monetary-benefit (Tambour et al. 1998), while in Equation 4, the net-benefits are measured in terms of health (Stinnett and Mullahy, 1998).

The same sample analogues are employed to estimate the mean effect and cost differences in order to give the estimated net-benefit statistics

$$\begin{aligned} N\hat{M}B &= l \cdot \Delta\overline{E} - \Delta\overline{C} \\ N\hat{H}B &= \Delta\overline{E} - \Delta\overline{C} \big/ l \end{aligned}$$

However, in contrast to the ICER, where the variance is not defined, the variance of net-benefits estimated from sample mean cost and effects in the trial arms is simply a linear combination of two asymptotically normal variables and can therefore be defined as:

$$\text{var}(N\hat{M}B) = l^{2}\,\text{var}(\Delta\overline{E}) + \text{var}(\Delta\overline{C}) - 2\,l\,\text{cov}(\Delta\overline{E}, \Delta\overline{C})$$

in terms of the monetary net-benefit measure, or:

$$\mathrm{var}(N\hat{H}B) = \mathrm{var}(\Delta\overline{E}) + \frac{1}{l^2}\mathrm{var}(\Delta\overline{C}) - \frac{2}{l}\mathrm{cov}(\Delta\overline{E}, \Delta\overline{C})$$

for the net health benefit measure. Therefore, the advantage of the net-benefits approach is that the $(1-a)\%$ confidence interval for net-benefits can be easily determined in the standard fashion, as $\hat{N}B \pm z_{a/2}\sqrt{s^2_{NB}}$ where $\hat{N}B$, is the estimated net-benefit measure, with variance s^2_{NB}, and $z_{a/2}$ is the critical value from the standard normal distribution.

The linear nature of net benefits makes them preferable to work with relative to the ICER. Stinnett and Mullahy (1998) identify four main advantages of the net-benefit framework: statistical inference is more straightforward; the sample estimate is an unbiased estimate of the true statistic; the mean of a distribution of net-benefit estimates is equal to the value of the statistic evaluated at the mean estimates of effects and costs; and the bootstrap estimate of the variance of the net-benefit statistic is expected to converge more quickly than similar estimates for the ICER. A potential drawback is that the net-benefit statistic is a function of l, a value unknown to the analyst in most cases. Stinnett and Mullahy (1998) consider this attribute a strength as it forces explicit consideration of the value of l, although they emphasise that it remains important to use the framework of net-benefits under different assumptions concerning what the appropriate value of l might be. Indeed, cost-effectiveness acceptability curves estimated using the net-benefit framework will exactly coincide with those calculated using an appropriate analysis on the CE plane since the underlying cost-effectiveness decision rule is the same in each case. Nevertheless, the net-benefit framework provides a much more straightforward method of calculating such acceptability curves.

2.3 A net-benefit regression framework

In this section we exploit the linear nature of the net benefit statistic to show how net-benefits can be used to estimate cost-effectiveness within a regression framework. Without loss of generality, we use net monetary benefits on the cost scale to illustrate the approach – the results could equivalently be presented in terms of net health benefit.

In a their comprehensive examination of the net-benefit statistic, Stinnett and Mullahy (1998) make the point that, in contrast to average cost-effectiveness ratios, the difference in the mean net benefit between arms will give the overall incremental net benefit statistic of Equation 3. This is straightforward to see algebraically

$$\begin{aligned} \overline{NMB}_1 - \overline{NMB}_0 &= \left(l \cdot \overline{E}_1 - \Delta\overline{C}_1 \right) - \left(l \cdot \Delta\overline{E}_0 - \Delta\overline{C}_0 \right) \\ &= l \left(\overline{E}_1 - \overline{E}_0 \right) - \left(\overline{C}_1 - \overline{C}_0 \right) \\ &= l \cdot \Delta\overline{E} - \Delta\overline{C} \\ &= \Delta\overline{NMB} \end{aligned}$$

through simple manipulation of the net-benefit expressions. Therefore the usefulness of average net-benefit is not directly in terms of the average figures themselves, but in the simple linear relationship between average and incremental net-benefit.

This immediately suggests a method by which the linearity of the net-benefit framework could be employed to directly estimate cost-effectiveness within a regression framework. First we formulate a net-benefit value for each patient i as

$$NMB_i = l \cdot E_i - C_i$$

where E_i and C_i are the observed effects and costs for each patient. At the simplest level, we could employ the linear model

$$NMB_i = a + \Delta t_i + e_i \qquad \text{(Model 1)}$$

where a is an intercept term, t a treatment dummy taking the value zero for the standard treatment and the value 1 for the treatment under consideration, and a random error term e. The coefficient Δ on the treatment dummy gives the estimated incremental net-benefit of treatment and will coincide with the usual estimate of incremental net-benefit obtained by aggregating across the treatment arms in a standard cost-effectiveness analysis. Similarly, the standard error of the coefficient is the same as that calculated from the standard approach.

The power of this framework is that it is straightforward to add additional explanatory variables in order to examine their impact on cost-effectiveness directly: for example, we can model the patient-level net-benefit with an alternative model

$$NMB_i = a + \sum_{j=1}^{p} b_j x_{ij} + \Delta t_i + e_i \qquad \text{(Model 2)}$$

where there are p known covariates x. That is, in this model the coefficient Δ on the treatment dummy gives the incremental net-benefit, and therefore the cost-effectiveness, of implementing the new treatment controlling for the known covariates. Of course in the context of an experimental design like a randomised controlled trial (RCT), the randomisation process is expected to ensure a balance of both observed and unobserved potentially confounding factors across the treatment arms.

Although regression methods have been used to correct for unbalanced allocation in observed covariates that has arisen by chance in clinical evaluation, that is not really the main advantage of adopting a regression based approach to cost-effectiveness analysis. Since economics is concerned fundamentally with the margin, the impact of covariates such as age, sex and disease severity on the cost-effectiveness of treatment interventions is of fundamental interest. All too often in RCT based cost-effectiveness analyses, the results are simply aggregated across the two arms to provide the overall ICER without any consideration of how the ICER varies at the margin. The net-benefit regression framework outlined in this section gives an explicit method for examining marginal issues: consider the model

$$NB_i = a + \sum_{j=1}^{p} b_j x_{ij} + \Delta t_i + t_i \sum_{j=1}^{p} g_j x_{ij} + e_i \qquad \text{(Model 3)}$$

where the final term is the interaction between the treatment dummy and the known covariates. The magnitude and significance of the coefficients g_j on the interaction between the covariates of the model and the treatment dummy will give a good indication of how cost-effectiveness of treatment is expected to vary at the margin.

The key advantage of this framework is the ability to use standard regression techniques for model selection and diagnostics to choose an appropriate model for cost-effectiveness instead of the usual approach of aggregating cost and effect differences across arms of the trial. We now illustrate the general approach with an applied example.

3. Example: empirical data from a randomised trial

3.1 Background

The Program in Assertive Community Treatment (PACT) is one of the most studied models of care for persons with severe and persistent mental illnesses (SPMI) (Stein and Test, 1980; Olfson, 1990; Burns and Santos, 1995; Scott and Dixon, 1995). Lehman et al. (1997) found that an assertive community treatment (ACT) program, relative to usual community services, reduced psychiatric inpatient days, emergency room visits, days homeless, and days in jail for *homeless* persons with SPMI in Baltimore, Maryland (USA). The study's rationale was that by providing potentially more expensive but coordinated, community-based care through the ACT program, homeless persons with severe mental illnesses would spend more days in stable community housing with savings realized by shifting the patterns of care from higher cost crisis-oriented inpatient and emergency services to lower cost, ongoing ambulatory services. The results suggest that in the city of Baltimore, ACT was effective in achieving important outcomes warranting an examination of the cost-effect trade-off. Lehman et al. (1999) conducted an economic evaluation of the ACT programme as it was implemented. Their analysis employed ICERs and provides an empirical example of the simplifying and unifying nature of the net-benefit framework.

3.2 Methods and Data

Direct treatment costs across the one year intervention period were examined from the perspective of the state mental health authority. Housing status was chosen as the main effectiveness measure because of its established validity as a primary outcome for homeless persons with SPMI (Newman, 1992). A day of stable housing was defined as living in a non-institutionalised setting not intended to serve the homeless (e.g., independent housing, living with family, etc.) Subjects randomised to the comparison usual care condition had access to services usually available to homeless persons in the city of Baltimore. Lehman, et al. (1999) offer more detail about the study's methodology.

One hundred forty eight persons who were homeless with SPMI were randomised to either the experimental ACT program or to usual community services. Subjects were recruited during a 19-month period in 1991 and 1992 from inner-city psychiatric hospitals, primary health care agencies, shelters, missions and soup kitchens. Baseline data collection included an assessment of overall mental health functioning using the Global Assessment of Functioning (GAF) Scale (American Psychiatric Association, 1994). For this paper, we obtained complete data on 73 participants randomly assigned to the ACT program and 72 randomly assigned to usual care (comparison) services.

3.3 Standard cost-effectiveness results

Baseline group comparisons examined differences between the two intervention groups on demographics, diagnoses and histories of homelessness at baseline (Lehman, 1999). Table 1 presents an abbreviated set of results. The two groups were comparable with ACT subjects being slightly older and higher functioning. In contrast, there was a greater than expected percentage of African Americans randomised to the comparison condition ($p < 0.01$).

Table 2 provides a brief statistical summary of the cost and effect data and provides a conventional cost-effectiveness analysis of the data by looking at the incremental costs and effects between the two groups. ACT subjects had lower costs and more days of stable housing, suggesting this was the dominant treatment. Due to the significant difference in the subjects between treatment arms with respect to ethnic origin, a stratified analysis is also reported in Table 2.

Of course, it is important to take into account the sampling variability of the data. Therefore the results from overall results from Table 2 are presented on the cost-effectiveness plane in Figure 2(a) and a cost-effectiveness acceptability curve (van Hout et al, 1994) is used to summarise the uncertainty in the cost-effectiveness in Figure 2(b). The stratified analyses are presented on the cost-effectiveness plane and as acceptability curves in Figure 3.

3.4 Results using a net-benefit regression approach

An equivalent net-benefit regression approach was employed by estimating the model

$$\mathrm{NMB}_i = a + \Delta \mathrm{ACT}_i + e_i$$

which is Model 1 from the previous section with ACT the treatment dummy variable. The results from this model are presented in Table 3. Net monetary benefits were calculated employing λ values of \$0, \$100, \$500, and \$1000. To estimate the behaviour of the *t*-statistics and p-values as $l \to \infty$, a regression was run with 'days stable housing', the effectiveness measure, as the dependent variable. The coefficients from the 'effect' regression (and not the NMB regression) are reported in Table 3 since as $l \to \infty$, the NMB coefficient estimates and their standard errors tend to $|\infty|$. However, as $l \to \infty$, the *t*-ratios and p-values from the NMB regression tend to those for the 'effect' regression. Therefore, the interest in Table 3 is on the ratio of the standard error to the coefficient estimates rather than on the coefficients themselves since the results in the final column are on a different scale (days stable housing).

Note that the coefficient on the treatment dummy corresponds to the incremental net-benefit (and equivalently incremental cost and effect) and that the regression results are exactly equivalent to the standard approach to cost-effectiveness analysis presented in Table 2. These regression results can also be used to obtain a cost-effectiveness acceptability curve of Figure 2(b) by plotting $1-p/2$ against l where p is the *p*-value from the coefficient on the treatment dummy. (The divisor of two is employed because the acceptability curve is equivalent to a one-sided test.) These values are plotted as points in Figure 2(b) and it is clear that they correspond to points on the acceptability curve calculated in the standard fashion. The dashed line in Figure 2(b) shows where the acceptability curve is tending to as $l \to \infty$.

Having demonstrated that a simple net-benefit regression approach is equivalent to a standard analysis, we consider some richer model specifications. It is clear from Table 1 that despite the randomisation of homeless persons to each arm of the study, many more white people where allocated to the treatment arm. In order to correct for this imbalance in the study design we estimate the following model to adjust for the observed age, race and GAF score of the subjects taking part in the study

$$\mathrm{NMB}_i = a + b_1\mathrm{black}_i + b_2\mathrm{age}_i + b_3\mathrm{gaf}_i + \Delta\mathrm{ACT}_i + e_i$$

which is Model 2 of the previous section. The estimated coefficients from this model are presented in Table 4. It is clear from these results that the covariate adjustment has more impact on the measure of effectiveness than the measure of cost, and therefore has a greater impact on net-benefit regressions based on higher l values, which place greater weight on the effect variable. While none of the individual covariates are significant, there is joint significance in the adjusted vales (for those regressions with greater weight on the effect variable) as measured by the *F*-test.

Again, in this model, it is the coefficient on the treatment dummy that gives the results of interest and these coefficients can be used to plot the adjusted results on the CE plane and to generate a cost-effectiveness acceptability curve. These are presented in Figure 3, where the unadjusted results are presented in light grey to aid comparison with Figure 2.

In Model 2, although coefficients are generated for the covariates, these are not of direct interest since they describe the impact on average net-benefits. Although average net-benefits are useful as a basis from which to obtain incremental net-benefits they are no use for decision making on their own, for the same reason that average cost-effectiveness ratios are of no use for decision-making. The interest in the covariates is on their effect on incremental net-benefit – in other words, their marginal impact on incremental cost-effectiveness. To examine this we employ a model that interacts the treatment dummy with the covariates

$$\mathrm{NMB}_i = a + b_1\mathrm{black}_i + b_2\mathrm{age}_i + b_3\mathrm{gaf}_i + \Delta\mathrm{ACT}_i + g_1\mathrm{ACT}_i\mathrm{black}_i + g_2\mathrm{ACT}_i\mathrm{age}_i + g_3\mathrm{ACT}_i\mathrm{gaf}_i + e_i$$

which corresponds to Model 3 of the previous section. The results from this regression are reproduced in Table 5. These results show that there is a significant interaction between race and treatment with black people achieving lower net-benefits from treatment in comparison to their white counterparts. The results also show that age and GAF score are potentially important covariates on the marginal cost-effectiveness of treatment, although the interaction terms across the regressions are not consistently significant at the 5% level. To see how these results can be used consider Figure 4, which shows uncertainty on the cost-effectiveness plane and cost acceptability curves for the white versus black subjects at average age and GAF score.

4. Discussion

The problems associated with statistical estimation of ICERs are now well documented in the health economic literature. The net-benefit framework, originally suggested as a method for handling uncertainty in economic evaluation, has been extended here to show how the net-benefit statistic can allow cost-effectiveness to be estimated directly in a regression framework. Within the context of simple regression analysis with a single (dummy) variable for the treatment intervention, the results are entirely equivalent to the standard approach to cost-effectiveness analysis. The advantages of the framework outlined in this paper comes from the ability to move beyond simple regression modelling to explore the impact of covariates on marginal cost-effectiveness

As recognised in the original economic evaluation of these data (Lehman et al., 1999), the observed difference in cost and effect between black and white subjects is noteworthy. The authors suggested one interpretation of the data is that the pattern of usual care for homeless persons with SPMI varies according to race in Baltimore. The researchers found that in spite of these differences, ACT tended to reduce excess use of crisis services and to increase the use of under-utilised ambulatory services for both groups. The authors concluded that the overall lower efficiency of ACT for black subjects in producing stable housing suggested that more attention should be given in the programme to differences between races on patterns of homelessness and service utilisation. These important policy findings were discovered by considering race as a covariate in the economic evaluation.

The importance of the net-benefit approach to cost-effectiveness analysis described here is that it is possible to estimate the marginal effect of race on the cost-effectiveness of the programme while controlling for other covariates. This shows the effect of race to be even more important that the original simple stratified analysis would suggest.

In our empirical example, both the negative ICER estimates (i.e. the dominance of the programme in terms of the point estimates) and the small sample sizes if the analysis where stratified underscored the potential of the net-benefit approach. It is now widely accepted that decision-making for cost-effectiveness analysis will depend on λ, the maximum willingness to pay per unit effect. Analysts may also be cautious of normative statements based on λ values in light of the fact that the net-benefit framework has the same prescriptive limitations as an ordinary cost-effectiveness analysis. The net-benefit approach simplifies some of the statistical aspects of cost-effectiveness analysis, but it is still cost-effectiveness analysis and subject to its inherent limitations (Stinnett & Mullahy,1998).

The net-benefit regression framework presented in this paper is largely one of convenience. Regression methods have not been used widely to estimate cost-effectiveness analysis in economic appraisals conducted alongside clinical trials. However, the theory for bivariate regression already exits and the knowledgeable analyst will be able to calculate entirely equivalent analyses without forming a net-benefit regression directly. Nevertheless, the ability to run net-benefit regressions directly on a statistical package is a practical advantage. Although OLS yields consistent estimates of the net-benefit coefficients, it may be possible to improve upon the efficiency of the OLS estimates by adapting the minimum distance estimator framework expressed in Chamberlain (1982) or the generalised method of moments framework described in Hansen (1982). Future econometric work might investigate this promising direction.

5. Summary

This paper has discussed how the net-benefit framework can simplify the statistical work involved in economic evaluation (e.g., avoiding problems associated ratio statistics) and also offer insights (e.g., exploring the importance of covariates on the marginal cost-effectiveness of an intervention). While it is true that the simplifying linearity comes at the cost of conditioning the analysis on values of λ, this may not be such a crucial limitation. We have suggested augmenting the standard net-benefit approach by utilising a regression framework, so that the potential strengths of both approaches may be incorporated into a unified framework. The marriage of econometrics and economic evaluation brings together something old (the regression framework), something new (the net-benefit framework), something borrowed (the decision-maker's λ) to produce something BLUE.

References

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders DSM-III-R,* 3rd revised edition. Washington, DC: American Psychiatric Association (1987).

Anderson, J.P., Bush, J.W., Chen, M. and Dolenc, D. "Policy space areas and properties of benefit-cost/utility analysis." *JAMA* **255**, 794-795, (1986).

Armitage P, Berry G. *Statistical Methods in Medical Research*, 3rd edition. Oxford: Blackwell Scientific Publications, 1994.

Black, W.C. "The CE plane: A graphic representation of cost-effectiveness." *Med.Decis.Making.* **10**, 212-214, (1990).

Briggs A, Fenn P. "Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane." *Health Economics*, 7: 723-740 (1998).

Burns BJ, Santos AB. "Assertive community treatment: an update of randomized trials." *Psychiatric Services* 46: 669-675 (1995).

Chamberlain G. "Multivariate regression models for panel data." *Journal of Econometrics*, 18: 5-46 (1982).

Chaudhary MA, Stearns, SC. "Estimating confidence intervals for cost-effectiveness ratios: An example from a randomized trial." *Statistics in Medicine*; 15: 1447-1458 (1996).

Drummond MF, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press (1987).

Golderberg AS. *Econometric Theory.* New York: J. Wiley (1964).

Hansen L, "Large sample properties of generalized method of moments estimators." *Econometrica* 50: 1029-1054 (1982).

Jones AM. "Health Econometrics" in *Handbook of Health Economics* edited by J.P. Newhouse and A.J. Culyer. Amsterdam: Elsevier Science Ltd, (2000).

Lehman AF, Dixon LB, Kernan E, DeForge BR, Postrado LT. "A randomized trial of assertive community treatment for homeless persons with severe mental illness." *Archives of General Psychiatry* 54: 1038-1043 (1997).

Lehman AF, Dixon LB, Hoch JS, DeForge BR, Kernan E, Frank RG. "Cost-effectiveness of Assertive Community Treatment for Homeless Persons with Severe Mental Illness. " *British Journal of Psychiatry* 174: 346-352 (1999).

Manski CF. *Analog Estimation Methods in Econometrics*. New York: Chapman and Hall (1988).

Mullahy, J. and Manning, W.G. "Statistical issues in cost-effectiveness analyses." In: Sloan, F.A., (Ed.) *Valuing health care*, pp. 149-184. Cambridge: Cambridge University Press (1995).

Newman SJ. "The severely mentally ill homeless: housing needs and housing policy." The Johns Hopkins University Institute for Policy Studies. Occasional paper no. 12 (1992).

O'Brien, BJ, Drummond MF, Labelle RJ, and Willan A. "In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care." *Medical Care* 32: 150-163 (1994).

Olfson M. "Assertive community treatment: an evaluation of the experimental evidence." *Hospital & Community Psychiatry* 41: 634-641 (1990).

Polsky D, Glick HA, Willke R, Schulman K. "Confidence intervals for cost-effectiveness ratios: A comparison of four methods." *Health Economics* 6: 243-252 (1997).

Scott JE, Dixon LB. "Assertive community treatment and case management for schizophrenia." *Schizophrenia Bulletin* 21: 657-668 (1995).

Stein LI, Test MA. "Alternative to mental hospital treatment: I. Conceptual model, treatment program, and clinical evaluation." *Archives of General Psychiatry* 37: 392-397 (1980).

Stinnett AA. "Adjusting for bias in C/E ratio estimates." *Health Economics*, 5: 470-472 (1996).

Stinnett, A. A., Paltiel, A. D. "Estimating CE ratios under second-order uncertainty: the mean ratio versus the ratio of means." *Medical Decision Making*, 17: 483-489 (1997).

Stinnett AA, Mullahy J. "Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis." *Medical Decision Making* 18 (Special Issue on Pharmacoeconomics): S68-S80 (1998).

Tambour M, Zethraeus N, Johannesson M. "A note on confidence intervals in cost-effectiveness analysis." *International Journal of Technology Assessment in Health Care*, 14(3): 467-471 (1998).

van Hout BA, Malwenn JA, Gordon GS, Rutten FFH. "Costs, effects, and C/E ratios alongside a clinical trial." *Health Economics*, 3: 309-319 (1994).

Weinstein MC, Fineburg HV. *Clinical Decision Analysis*. Philadelphia: W. B. Saunders (1980).

Weinstein, M.C. and Stason, W.B. Foundations of cost-effectiveness analysis for health and medical practices. *N.Engl.J.Med.* **296**, 716-721, (1977).

Table 1
Subject characteristics by treatment group[a]

Characteristic	ACT subjects	Comparison subjects
Mean Age (SD)	38.96 years (9.43)	36.00 years (8.30)
Mean GAF[b] Score (SD)	37.90 (9.08)	35.32 (9.06)
African American**	62 %	83 %

[a] ACT indicates Assertive Community Treatment.
[b] GAF is the Global Assessment of Functioning score.
** $p < 0.01$.

Table 2
Sample statistics from the economic evaluation data

Group Variable	Mean	S.D.	S.E.
Overall analysis:			
Comparison Arm (N=72)			
Cost[a]	67,400	76,500	9,020
Effect[b]	159	105	12.4
Correlation = -0.43			
ACT Arm (N=73)			
Cost	51,900	61,100	7160
Effect	212	104	12.2
Correlation = -0.39			
Increments			
Cost difference	-15,500	-	11,500
Effect difference	52.7	-	17.4
Correlation = -0.41			
Partioned analysis:			
Black subjects			
Cost difference	-5,070	-	13,200
Effect difference	35.6	-	21.8
White subjects			
Cost difference	-62,700	-	33,000
Effect difference	98.1	-	39.0

Incremental net benefits:	**Net monetary benefit (SE)**		
Value of ceiling ratio	*Overall*	*Black subjects*	*White subjects*
$l = 0$ [c]	15,500 (11,500)	5,070 (13,200)	62,700 (33,000)
$l = 100$	20,800 (12,300)	8,630 (14,500)	72,600 (35,400)
$l = 500$	41,900 (17,000)	22,900 (21,300)	112,000 (46,800)
$l = 1000$	68,200 (24,500)	40,700 (31,100)	161,000 (63,700)
$l = \infty$ [d]	∞ (∞)	∞ (∞)	∞ (∞)

[a] All costs in US dollars, all results to three significant figures
[b] All effects in ‘days of stable housing’
[c] Equivalent to minus the cost difference
[d] As $l \to \infty$, $N\bar{M}B \to \infty$ and $\mathrm{se}\left(N\bar{M}B\right) \to \infty$

Table 3
Simple net-benefit regression estimates (Model 1)[a]

N = 145 Explanatory Variables	NMB with λ = $0 [b] [se] (p-value)	NMB with λ = $100 [se] (p-value)	NMB with λ = $500 [se] (p-value)	NMB with λ = $1000 [se] (p-value)	Effect [c] [se] (p-value)
Constant Term	67,400 [8,160] (<0.001)	51,500 [8,740] (<0.001)	12,200 [12,000] (0.313)	91,800 [17,400] (<0.001)	159 [12.4] (<0.001)
Treatment dummy:					
ACT	15,500 [11,500] (0.179)	20,800 [12,300] (0.093)	41,900 [17,000] (0.015)	68,200 [24,500] (0.006)	52.7 [17.4] (0.003)
R-squared (adjusted)	0.006	0.018	0.041	0.045	0.054
F(1, 143)	1.82	2.85	6.04	7.75	9.13
Prob > F	0.179	0.094	0.015	0.006	0.003

[a] All monetary measures in U.S. dollars, all results to three significant figures.
[b] When λ= $0, NMB = - Cost.
[c] The coefficients from the Effect regression (and not the NMB regression with $\lambda \to \infty$) are reported since as $\lambda \to \infty$, the *p*-values for the NMB coefficient estimates are equivalent to those obtained when 'days stable housing' is the dependent variable.

Table 4
Covariate adjusted net-benefit regression estimates (Model 2)[a]

N = 145 Explanatory Variables	NMB with λ = $0 [b] [se] (p-value)	NMB with λ = $100 [se] (p-value)	NMB with λ = $500 [se] (p-value)	NMB with λ = $1000 [se] (p-value)	Effect [c] [se] (p-value)
Constant Term	-80,700 [13,900] (<0.001)	-63,800 [14,900] (<0.001)	3,890 [20,600] (0.850)	88,500 [29,500] (0.003)	169 [20.8] (<0.001)
Covariates:					
black (dummy)	16,400 [13,500] (0.224)	15,700 [14,400] (0.276)	13,100 [19,900] (0.531)	9,690 [28,500] (0.734)	-6.73 [20.1] (0.739)
age	-86.0 [660] (0.896)	72 [707] (0.919)	704 [975] (0.471)	1,490 [1,400] (0.287)	1.58 [0.987] (0.112)
gaf	372 [644] (0.565)	527 [690] (0.446)	1150 [953] (0.230)	1,920 [1,370] (0.161)	1.55 [0.964] (0.110)
Treatment dummy:					
ACT	18,400 [12,100] (0.132)	22,600 [13,000] (0.084)	39,600 [17,900] (0.029)	60,900 [25,700] (0.019)	42.51 [18.2] (0.021)
R-squared (adjusted)	<0.001	0.005	0.031	0.074	0.071
F(4, 140)	0.96	1.2	2.15	2.81	3.75
Prob > F	0.431	0.316	0.078	0.028	0.006

[a] All monetary measures in U.S. dollars, all results to three significant figures.
[b] When λ= $0, NMB = - Cost.
[c] The coefficients from the Effect regression (and not the NMB regression with $\lambda \rightarrow \infty$) are reported since as $\lambda \rightarrow \infty$, the *p*-values for the NMB coefficient estimates are equivalent to those obtained when 'days stable housing' is the dependent variable.

Table 5
Covariate adjusted net-benefit regression estimates with treatment interaction (Model 3)[a]

N = 145 Explanatory Variables	NMB with λ = $0 [b] [se] (p-value)	NMB with λ = $100 [se] (p-value)	NMB with λ = $500 [se] (p-value)	NMB with λ = $1000 [se] (p-value)	Effect [c] [se] (p-value)
Constant Term	-109,000 [19,500] (<0.001)	-95,000 [20,800] (<0.001)	-38,900 [28,700] (0.178)	31,200 [41,400] (0.452)	140 [29.9] (<0.001)
Covariates:					
black (dummy)	54,300 [21,400] (0.012)	57,500 [22,800] (0.013)	70,400 [31,500] (0.027)	86,400 [45,400] (0.059)	32.1 [32.7] (0.329)
age	1,300 [1,000] (0.189)	1,580 [1,070] (0.141)	2,620 [1,470] (0.077)	3,930 [2,120] (0.067)	2.61 [1.53] (0.091)
gaf	1,260 [915] (0.170)	1,560 [690] (0.112)	2,760 [1,350] (0.043)	4,260 [1,940] (0.030)	2.99 [1.40] (0.034)
Treatment dummy:					
ACT	64,400 [23,400] (0.007)	73,400 [25,000] (0.004)	109,000 [34,400] (0.002)	154,000 [49,700] (0.002)	89.8 [35.8] (0.013)
Treatment-covariate interactions:					
ACT· black	-60,200 [26,900] (0.027)	-66,400 [28,700] (0.022)	-91,000 [39,600] (0.023)	-122,000 [57,100] (0.035)	-61.5 [41.2] (0.137)
ACT· age	-2,860 [1,310] (0.031)	-3,100 [1,400] (0.029)	-4,050 [1,930] (0.038)	-5,230 [2,780] (0.063)	-2.37 [2.01] (0.240)
ACT· gaf	-2,280 [1,270] (0.075)	-2,600 [1,350] (0.057)	-3,900 [1,870] (0.038)	-5,540 [2,690] (0.042)	-3.26 [1.94] (0.095)
R-squared (adjusted)	0.069	0.082	0.150	0.112	0.097
F(7, 137)	2.53	2.83	3.46	3.60	3.2
Prob > F	0.018	0.009	0.002	0.001	0.004

[a] All monetary measures in U.S. dollars, all results to three significant figures.
[b] When λ= $0, NMB = - Cost.
[c] The coefficients from the Effect regression (and not the NMB regression with $\lambda \rightarrow \infty$) are reported since as $\lambda \rightarrow \infty$, the *p*-values for the NMB coefficient estimates are equivalent to those obtained when 'days stable housing' is the dependent variable.

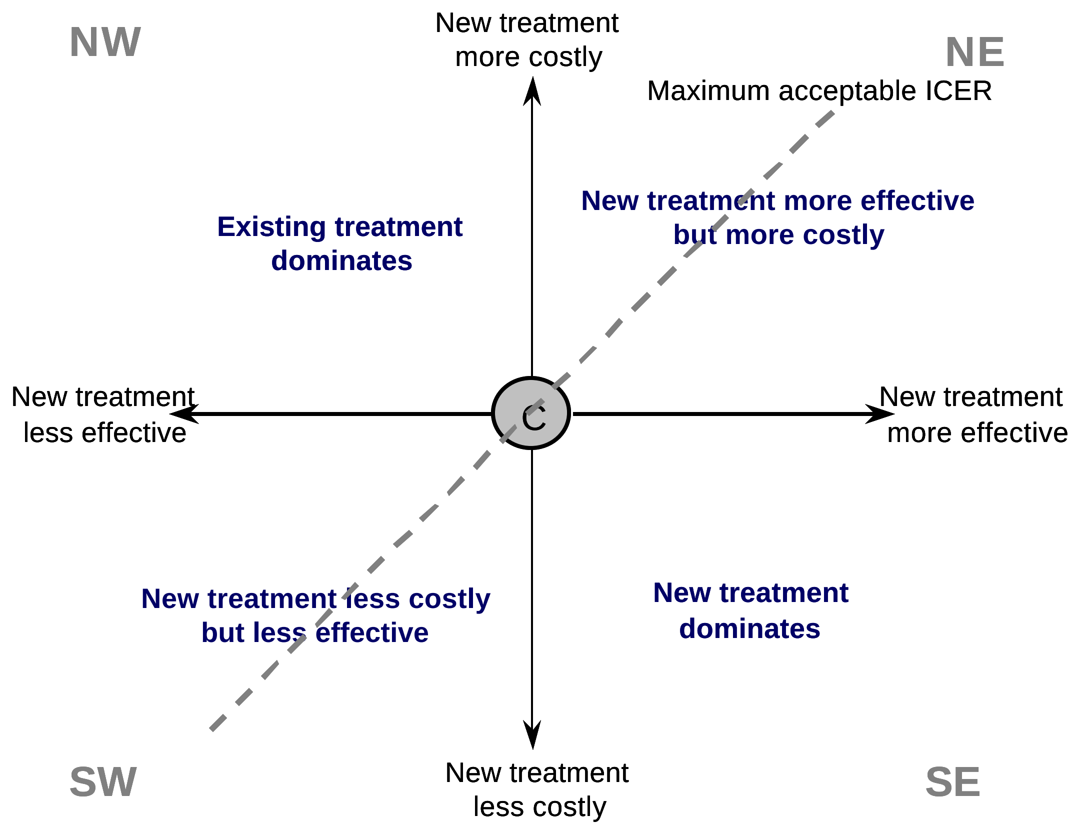
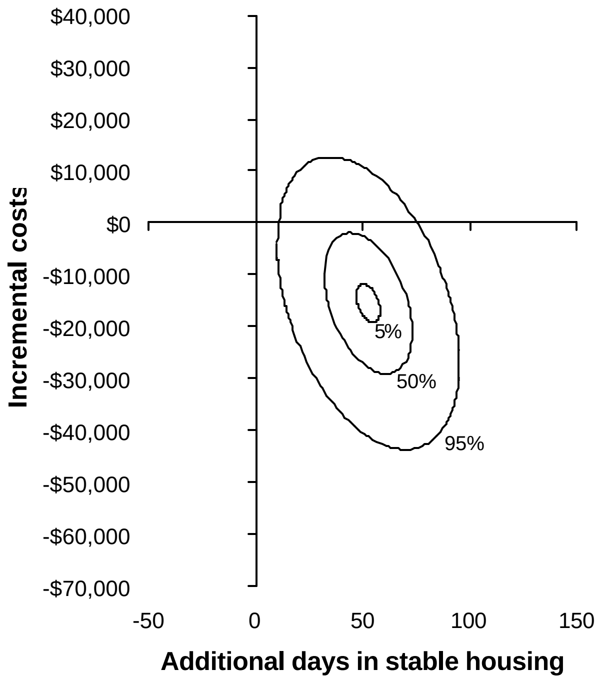
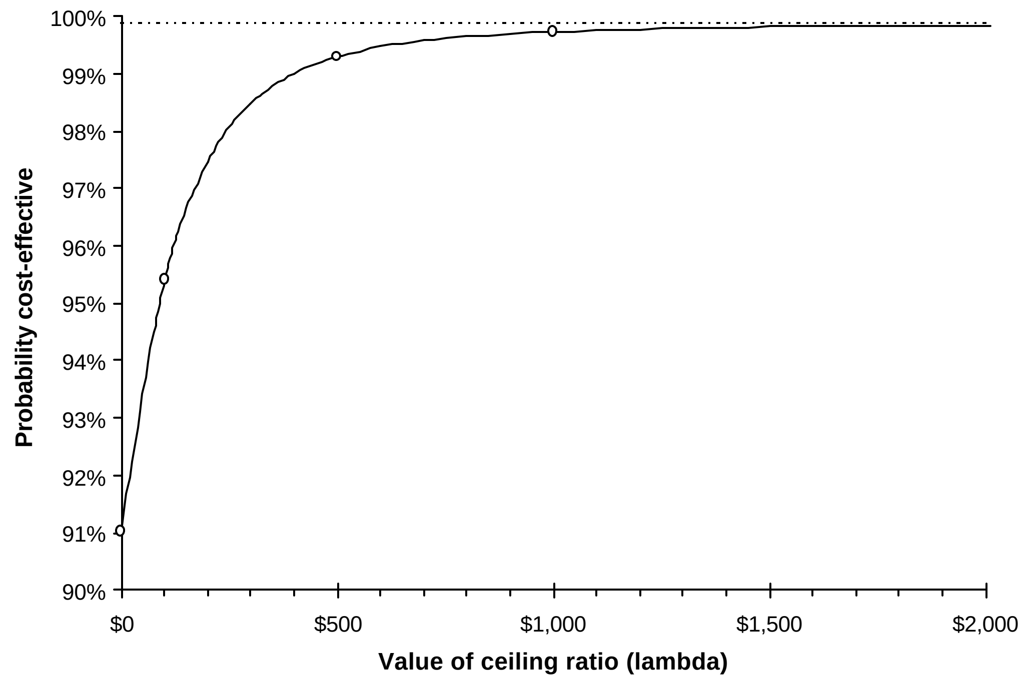


Figure 1
The cost-effectiveness plane



(a) The cost-effectiveness plane

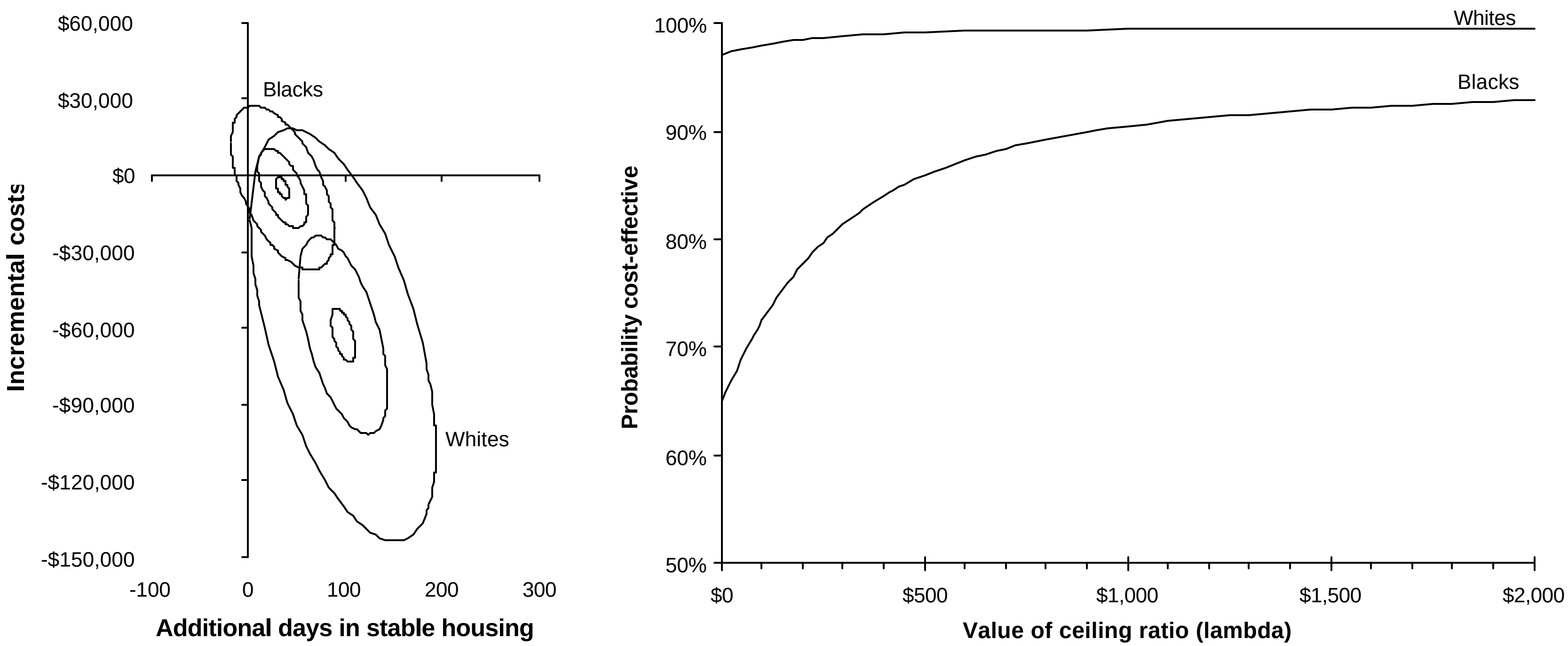


(b) The cost-effectiveness acceptability curve

Figure 2

The cost effectiveness plane and acceptability curve for the standard cost-effectiveness analysis of the ACT programme

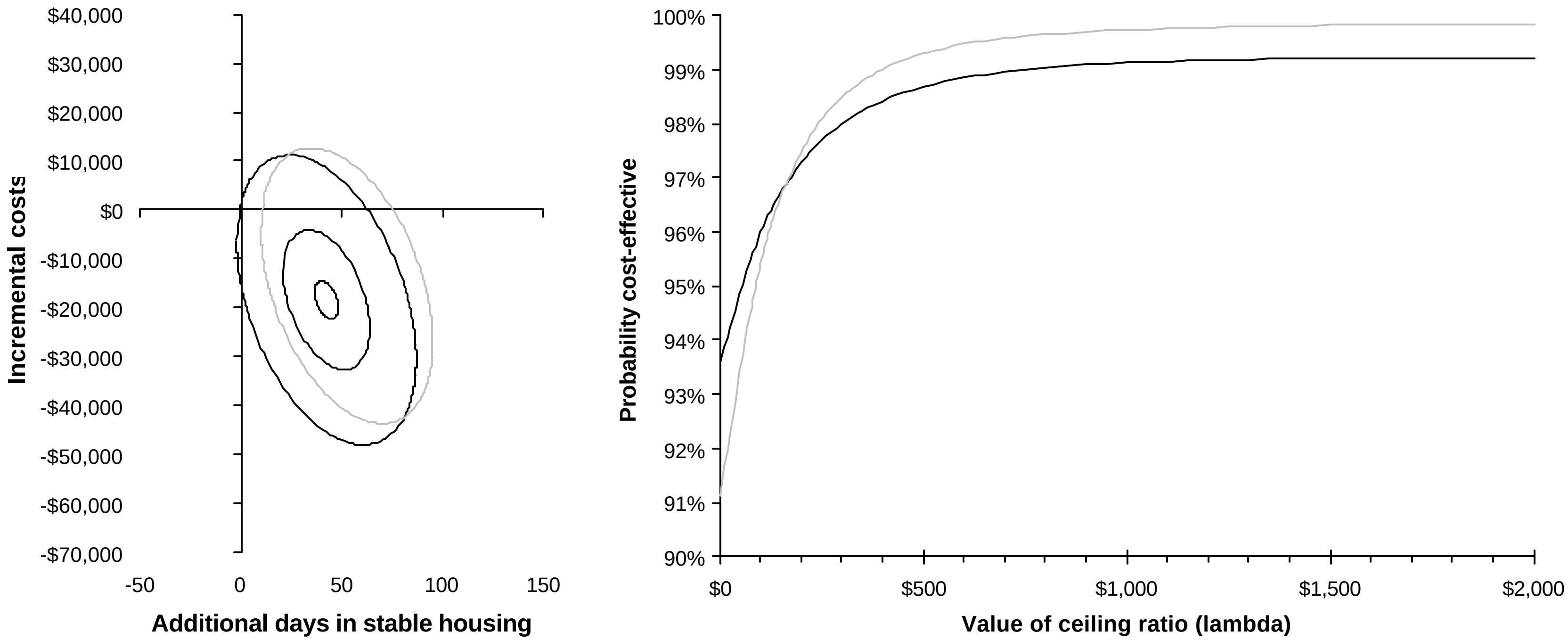
Uncertainty on the CE plane is represented by elliptical contours covering 5%, 50% and 95% of the joint density of cost and effect differences



(a) The cost-effectiveness plane

(b) The cost-effectiveness acceptability curve

Figure 3
The cost effectiveness plane and acceptability curve for the standard partitioned cost-effectiveness analysis of the ACT programme



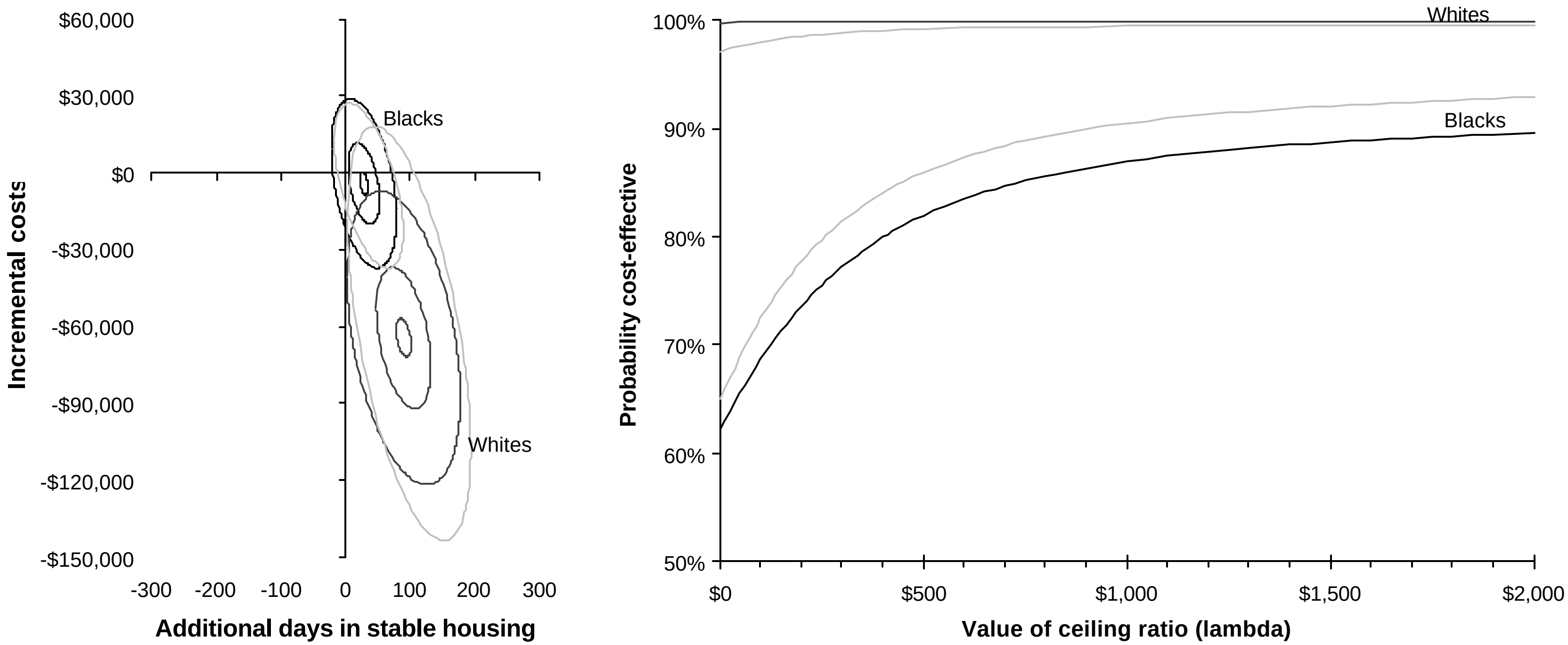
(a) The cost-effectiveness plane

(b) The cost-effectiveness acceptability curve

Figure 4

The cost effectiveness plane and acceptability curve for the covariate adjusted cost-effectiveness analysis of the ACT programme

The light grey ellipse and curve in each panel shows the position of the unadjusted estimates from Figure 2



(a) The cost-effectiveness plane

(b) The cost-effectiveness acceptability curve

Figure 5
The cost effectiveness plane and acceptability curve for regression-based cost-effectiveness analysis of the ACT programme by race
The light grey ellipses and curves in each panel shows the position of the unadjusted estimates from the standard partitioned analysis in Figure 3