

# ICEinfer in R:

## **Quantification and Visualization of Statistical Uncertainty and Economic Preference In Incremental Cost-Effectiveness Inference**

ICEinfer R-Package [28], Version 0.1-1, July 2007

Robert L. Obenchain, Ph.D.  
Risk Benefit Statistics, LLC  
[softrx@iquest.net](mailto:softrx@iquest.net)

Senior Research Advisor, Retired  
Eli Lilly and Company, US Medical, Outcomes Research

### **Table of Contents**

<b>Section</b>	<b>Topic</b>	<b>Page</b>
<b>1</b>	<b>Introduction, History and Outline</b>	<b>2</b>
<b>2</b>	<b>Basic ICE Notation and Terminology</b>	<b>4</b>
2.1	ICE Outcomes ( $\Delta E$ , $\Delta C$ )	4
2.2	ICE Linear Subgroup of Transformations	4
<b>3</b>	<b>ICE Economic Preference Maps</b>	<b>5</b>
3.1	Basic Preference Assumptions and Fundamental Axioms	5
3.2	Functional Form of Two-Parameter “Signed Power” Maps	6
3.3	ICE Returns-to-Scale Power Parameter, $\beta$	6
3.4	Willingness-to-Pay, Willingness-to-Accept and Bob O’s “link” function, $\lambda = \sqrt{WTP \times WTA}$	7
3.5	Visualizing Indifference Curves using ICEepmap() & ICEomega()	9
3.6	Implications of Using a “Wrong” Numerical Value for $\lambda$ .	10
3.7	Laupacis’ visualization of ICE preferences corresponds to the limit as $\beta$ approaches 0.	12
3.8	The Relationship between the Slope of the ICE Ray through a Point and the Willingness Rate at that Point	13

# Table of Contents, Concluded

Section	Topic	Page
4	<b>Transformations of the Bootstrap ICE Uncertainty Distribution within the Linear Subgroup</b>	14
5	<b>A Wedge-Shaped, Equivariant Statistical Confidence Region</b>	16
6	<b>Acceptability: VAGR versus ALICE Curves</b>	18
7	<b>Additional Uncertainty due to dependence of Economic Preferences on the Shadow Price of Health</b>	23
	<b>Final Remarks / Acknowledgement</b>	26
	<b>References / Tables / Figures</b>	28

## 1. Introduction, History and Outline

The primary purpose of this PDF Documentation file is provide users of my ICEinfer R-package with technical background information about basic statistical and econometric concepts and graphical visualizations useful in Incremental Cost-Effectiveness (ICE) analyses. The tutorial of Briggs and Fenn [2] provides a good introduction to and review of the extremely wide variety of methodologies that have been proposed to quantify ICE uncertainty.

ICE statistical inference is a complex 2-sample, 2-variable (bivariate) problem. Specifically, ICE inference usually uses data from patients who receive only one of the two treatments that are to be compared. (In other words, a patient does not contribute data to both samples as he/she would in a treatment “cross-over” experiment.) Furthermore, possibly relevant differences between treatment groups involve two types of outcomes (cost and effectiveness) that, although possibly correlated, are typically considered quite distinct by various health care stake holders. Payers are rightfully most interested in the cost dimension, and patients subject to insurance and out-of-pocket charges clearly also share some of this concern. On the other hand, health care providers and patients (plus their families) are quite justifiably most concerned about treatment differences in effectiveness.

**Some History:** As an aid to visualization in 2-dimensional, real Euclidean space, Black [1] proposed that the incremental difference (new treatment minus standard treatment) in mean effectiveness,  $\Delta E$ , be plotted horizontally while the corresponding difference in mean cost,  $\Delta C$ , is plotted vertically. Technically, since the effectiveness difference is displayed along the horizontal, X-axis, confusion might be reduced of this display were known as the “effectiveness versus cost difference” plane. Traditionally, this particular setup is called the “ICE Plane.”

To depict uncertainty in a  $(\Delta E, \Delta C)$  point estimate (a bivariate statistic), its bootstrap distribution under resampling of patient outcome pairs both with replacement and within treatment groups is

displayed as a scatter of points on the ICE Plane. See Briggs and Fenn [2, pp 731-734] for a summary of ICE bootstrapping methodology.

A (univariate) confidence interval for the unknown true expected value of any scalar valued ICE summary statistic can be computed from the bivariate bootstrap distribution of uncertainty in  $(\Delta E, \Delta C)$  estimates. Unfortunately, substantial challenges to the use and interpretation of confidence intervals for the ICE Ratio,  $ICER = \Delta C / \Delta E$ , have been noted [16,11,12]. Although closely related to the ICE Ratio, the Net Benefit (NB) approach [4,27,3,15,26] may have avoided similar criticism because it attempts to quantify overall preference or utility. In NB, constant preference contours (indifference curves) are straight lines on the ICE plane with positive slope,  $\lambda$ . Within the North East (NE) quadrant, this slope can be interpreted as the willingness-to-pay (WTP) a higher cost in return for increased effectiveness; within the South West (SW) quadrant, this same slope is interpreted as the willingness-to-accept (WTA) a less effective treatment in return for lower cost. In other words, NB collects outcomes on the ICE plane into linear equivalence classes (straight lines of iso-preference) ordered by a single index, the NB.

**Section 2** introduces basic notation and demonstrates that all commonly considered transformations of the ICE plane, including both treatment re-labelings (new versus standard) and axis rescalings (changes in the shadow price of health), are simple linear transformations with a fixed point at the ICE origin. With  $\lambda$  held fixed, treatment differences in outcome are first standardized by expressing both  $(\Delta E, \Delta C)$  differences in identical units, either both in cost units or else both in effectiveness units.

**Section 3** covers the theory and application of (nonlinear) ICE Preference Maps. The assumption of uniform slope for iso-preference curves, as in traditional NB, is unrealistic in the sense that linear utility has been consistently contradicted in empirical studies [18,30,17]. O'Brien et al. [17] provide a good introduction to WTP and WTA concepts as well as a highly readable summary of relevant literature.

Obenchain [23,24] argues that coherent ICE preferences satisfy four intuitive axioms and proposes a 2-parameter family of maps that satisfy these axioms and provide highly realistic generalizations of NB. For example, nonlinear maps do not require that Returns-to-Scale be linear (constant) or that willingness-to-pay (WTP) and willingness-to-accept (WTA) be equal to the Shadow Price of Health,  $\lambda$ .

**Section 4** discusses alternative ways to visualize the effects of transformations within the Linear Subgroup on the bootstrap distribution of ICE uncertainty.

**Section 5** describes how to use ICE Angle order statistics (around a circle) to form a wedge-shaped ICE Confidence Region that is “equivariant” (commutative) under the Linear Subgroup.

**Section 6** then discusses important distinctions between “ALICE” curves and the traditional, linear-frontier VAGR measure of acceptability.

**Section 7** illustrates the consequences of equivalence of treatments along the straight line through the ICE origin of slope  $\lambda$  within all Economic Preference Maps. Because such maps

tend to reduce dimensionality in ICE inference from 2 dimensions (cost and effectiveness) to only 1 “composite” preference dimension (a range of scalar values), some authors have claimed that inference is thereby simplified. We demonstrate the exact opposite! Economic uncertainty about  $\lambda$  can not only swamp the statistical uncertainty in patient level data about the true location of  $(\Delta E, \Delta C)$  but also can introduce inconsistency.

## 2. Basic ICE Notation and Terminology

### 2.1 ICE outcomes

An ICE outcome is represented by a pair of expected treatment differences, usually expressed in Cartesian coordinates as  $(\Delta E, \Delta C)$ . Here,  $\Delta E$  is a difference in average treatment effectiveness of the form “new” treatment minus “standard” treatment. The underlying effectiveness measurement needs to be defined in such a way that larger (more positive) values of  $\Delta E$  are unambiguously more favorable to the new treatment. The corresponding difference in average per-patient cost,  $\Delta C$ , must be such that smaller (more negative) values are unambiguously more favorable to the new treatment.

### 2.2 ICE linear subgroup of transformations

Transformations of ICE outcome coordinates occur quite naturally. For example, interchanging the labels (new and standard) on the two treatments being compared would multiply both  $\Delta E$  and  $\Delta C$  by minus one.

Let  $\lambda$  denote society’s fixed “shadow price” for one unit of health care effectiveness. In other words,  $\lambda$  is a strictly positive *substitution rate* expressed in units of cost per unit of effectiveness.

For any specified value of  $\lambda$ , a cost difference of  $y = \Delta C$  is re-expressed in effectiveness units by dividing it by  $\lambda$ .

Alternatively, the corresponding effectiveness difference of  $x = \Delta E$  would be expressed in cost units by multiplying it by  $\lambda$ .

An arbitrary ICE outcome  $(\Delta E, \Delta C)$  thus gets transformed into either  $(x, y) = (\lambda \Delta E, \Delta C)$  in cost units or into  $(x, y) = (\Delta E, \Delta C / \lambda)$  in effectiveness units. Either choice represents a ***standardized, canonical form*** for expected, overall treatment differences that clearly depends upon choice of a fixed, numerical value for  $\lambda$ .

The transformations of interest can be represented using 2×2, diagonal matrices of the form

$$\begin{bmatrix} \alpha \lambda & 0 \\ 0 & \alpha \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \alpha & 0 \\ 0 & \alpha / \lambda \end{bmatrix},$$

where the  $\alpha$  values are either both +1 or else both -1, and  $\lambda$  is a strictly positive and finite. Specifically, visualize a  $2 \times 1$  (column) vector of Cartesian coordinates being multiplied on the left by a  $2 \times 2$  diagonal matrix of the above form, producing a transformed  $2 \times 1$  vector of standardized coordinates,  $(x, y)$ . The implied subgroup of transformations is clearly linear, and the straight line through the origin defined by  $\Delta E / \Delta C = \lambda$  has been transformed into the  $x = y$  line with slope plus one.

The above transformations form a linear subgroup much more restrictive than the full group of “affine” transformations commonly considered in multivariate analysis (bivariate, in our case.) In particular, this linear subgroup not only preserves the ICE origin,  $(0, 0)$ , as a fixed point by excluding additive “translation” vectors but also excludes “rotations” and/or “shears” between axes that result from  $2 \times 2$  transformation matrices that are neither diagonal nor symmetric (or possibly have diagonal entries that differ in numerical sign.)

### 3. ICE Economic Preference Maps

Let  $P(x, y)$  denote a real-valued function, called an “ICE preference map,” that determines, as explained below, not only which of two treatments is preferred but also the strength of that preference. Specifically,  $P(x, y)$  can be visualized as a surface defined over the entire 2-dimensional Euclidean plane,  $(x, y)$ , of *standardized* ICE outcomes.

#### 3.1 Basic Preference Assumptions and Fundamental Axioms

Our three primary interpretation conventions (assumptions) for  $P(x, y)$  will be as follows:

$P(x, y) = 0$  means that the  $(x, y)$  pair of treatment differences correspond to no preference whatsoever, either for the new treatment over the standard treatment or vice-versa.

$P(x, y) > 0$  means that the treatment currently called new is preferred over the treatment currently called standard. Strictly positive  $P(x, y)$  values are at least ordinal measures of strength of preference for the new treatment over the standard treatment.

$P(x, y) < 0$  means that the treatment currently called standard is preferred over the treatment currently called new. The *absolute values* of negative  $P(x, y)$  values are at least ordinal measures of strength of preference for the standard treatment over the new treatment.

Table 1 lists four axiomatic properties of ICE preference maps. When first examining this table, it may be helpful to note that the *linear* preference map, [27], is  $NB(x, y) = x - y$ , which clearly satisfies all four of these axioms.

**Table 1. Four Axioms of ICE Preference**

Indifference and Direction of Preference:	$P(x, y) = 0$ when $x = y$ , $P(x, y) > 0$ when $x > y$ , and $P(x, y) < 0$ when $x < y$ .
Monotonicity:	$P(x, y) \geq P(x_0, y_0)$ for all $x \geq x_0$ and $y \leq y_0$ .
Re-labeling:	$P(x, y) = -P(-x, -y)$
Symmetry and Anti-Symmetry:	$P(x, y) = P(-y, -x)$ $= -P(y, x)$

Note that the re-labeling, symmetry and anti-symmetry axioms represent additional restrictions on ICE preferences only when  $x \neq y$ . After all, the  $P(x, x) = 0$  property of the first axiom renders the implications of all other axioms moot for all outcomes with  $x = y$ .

### 3.2 Two-parameter ICE preference maps.

To generalize the linear preference map  $NB(x, y) = x - y$ , let us now consider the family of ICE preference maps of the form

$$P(x, y) \propto (x^2 + y^2)^{(\beta-\gamma)/2} \{x - y\}^\gamma, \quad (1)$$

where  $\propto$  means “is proportional to,”  $\beta$  and  $\gamma$  are strictly positive “power” parameters, and the special notation  $\{z\}^\gamma$  denotes a “signed-power.” Specifically,  $\{z\}^\gamma$  denotes the product of  $\text{sign}(z)$  [which is +1, 0 or -1] times the absolute value of  $z$  raised to the power  $\gamma$ . Special care has been taken here because non-integer powers of negative real numbers are generally imaginary; ICE preferences need to be expressible as real numbers even though they may provide only ordinal measures of preference strength.

It is straight-forward to verify that all ICE maps of form (1) satisfy axioms 1, 3 and 4 of Table 1. Obenchain [23,24] showed that the following range restriction on the ratio of the  $\beta$  and  $\gamma$  power parameters,

$$1/\Omega \leq \gamma/\beta \leq \Omega \quad \text{for } \Omega = (1 + \sqrt{2})^2 \approx 5.828, \quad (2)$$

is necessary and sufficient for maps defined by equation (1) to also satisfy axiom 2, ICE monotonicity. The class of two-parameter ICE preference maps satisfying equations (1) and (2) constitutes the “signed-power” family.

### 3.3 ICE Returns-to-Scale Parameter, $\beta$

Suppose now that the observed treatment differences in cost,  $y$ , and effectiveness,  $x$ , are somehow both multiplied by a strictly positive and finite real valued factor  $f$ . In other words, the

observed effectiveness difference of  $x$  becomes  $f$  times  $x$ , while the observed cost difference of  $y$  becomes  $f$  times  $y$ . The resulting new value of preference in equation (1) is then

$$P(fx, fy) \propto f^{(\beta-\gamma)+\gamma} [x^2+y^2]^{(\beta-\gamma)/2} \{x-y\}^\gamma \propto f^\beta P(x, y). \quad (3)$$

In other words, for every map in the 2-parameter family, (1), returns-to-scale depend solely upon the  $\beta$  power parameter attached to the ICE radius factor. Specifically, returns-to-scale will be:

$$\begin{array}{ll} \text{decreasing} & \text{if } 0 < \beta < 1, \\ \text{constant (linear)} & \text{if } \beta = 1, \text{ and} \\ \text{increasing} & \text{if } 1 < \beta < +\infty. \end{array}$$

### 3.4 Willingness-to-Pay, Willingness-to-Accept and Bob O's "link" function.

WTP or WTA at any point on the ICE plane is assumed here to be determined by the iso-preference contour that passes through that given point. In fact, we define a *standardized* "willingness" rate (of  $\text{WTP}/\lambda$  within the NE quadrant or  $\text{WTA}/\lambda$  within the SW quadrant) as being equal to the  $dy/dx$  slope of the tangent to the iso-preference contour at the point of interest. This is fully consistent with NB analysis in which iso-preference contours are straight lines of slope  $\text{WTP} = \text{WTA} = \lambda$ ; note that the standardized value for all three of these quantities is +1 at all points in the lower-left panel of Figure 1.

Obenchain[24] showed that the *standardized* willingness rate at  $(x, y)$  for all signed-power ICE preference maps of form (1) is

$$\begin{aligned} w(x, y) &= [\beta x^2 + (\gamma-\beta)xy + \gamma y^2] / [\gamma x^2 + (\gamma-\beta)xy + \beta y^2] \\ &= [1 + (\eta-1)s + \eta s^2] / [\eta + (\eta-1)s + s^2], \end{aligned} \quad (4)$$

where  $s = y/x$  is the standardized ICE ratio and  $\eta = \gamma/\beta$  is the map "power-parameter ratio." Remembering that  $(x, y)$  denotes either  $(\lambda\Delta E, \Delta C)$  in cost units or  $(\Delta E, \Delta C/\lambda)$  in effectiveness units, it follows that  $w(x, y)$  represents either

- [a] a non-negative value of  $\text{WTP}/\lambda$  when  $\Delta E, x, \Delta C$  and  $y$  are all positive
- or else
- [b] a non-negative value of  $\text{WTA}/\lambda$  when  $\Delta E, x, \Delta C$  and  $y$  are all negative.

Since  $\beta$  and  $\gamma$  are unitless parameters and  $x$  and  $y$  are both measured here in the same units, it follows that  $\eta = \beta/\gamma$ ,  $s = y/x$  and  $w(x, y)$  are all unitless quantities.

Within generalized-linear maps ( $\eta = 1$ ), note that  $w(x, y) \equiv 1$  is a fixed value at all points  $(x, y)$  and for all directions  $s = y/x$ .

For any fixed value of  $\eta$  different from 1, the standardized willingness rate (4) varies only with  $s$ . In other words, standardized willingness is then constant everywhere along each straight-line trajectory,  $s$ , passing through the origin of the ICE plane except at the origin itself. After all,

neither  $w(0, 0)$  nor  $s = y/x$  are well defined at the ICE origin. Unlike ICE preferences,  $P(x, y)$ , standardized willingness does not vary with ICE radius,  $r = \sqrt{x^2 + y^2}$ , within the 2-parameter maps of equation (1).

The (un-standardized) ICE willingness rate takes on 3 different, simple forms in the following special cases:

**Table 3. Some Willingness values when  $\eta \neq 1$ .**

$W(\Delta E, \Delta C)$	$= \lambda$	when $s = \pm 1$ ( $x = y \neq 0$ or $x = -y \neq 0$ ),
	$= \lambda/\eta$	when $s = 0$ ( $y = 0$ and $x \neq 0$ ),
	$= \lambda\eta$	as $s$ approaches $\pm\infty$ ( $x = 0$ and $y \neq 0$ .)

The single, most important implication of the ICE symmetry axiom appears to be that every well matched pair of values for a WTP within the North East ICE Quadrant with a WTA within the South West ICE Quadrant must have the property that WTP times WTA equals  $\lambda^2$ . In other words, for our signed-power ICE preference maps (or any differentiable ICE preference map satisfying the symmetry axiom), the following key relationship will always hold:

$$\lambda = \sqrt{\text{WTP} \times \text{WTA}} \quad (5)$$

Equation (5) states that the shadow price of health is the geometric mean of all well-matched pairs of strictly positive WTP and WTA values. In other words, equation (5) shows that WTP and WTA can both vary simultaneously within a *fixed*, nonlinear ICE preference map corresponding to a single *fixed* value of  $\lambda$ . [Figure 4 on page 12 illustrates the analytical geometry behind equation (5).] Anyway, relative to choice of  $\lambda$ , choices for the values of the  $\beta$  and  $\eta$  (or  $\gamma$ ) parameter are *clearly less important*. After all, equation (5) holds for all choices of  $\beta$  and  $\eta$ . This “link” function, Obenchain[24], has profound theoretical (and practical) implications.

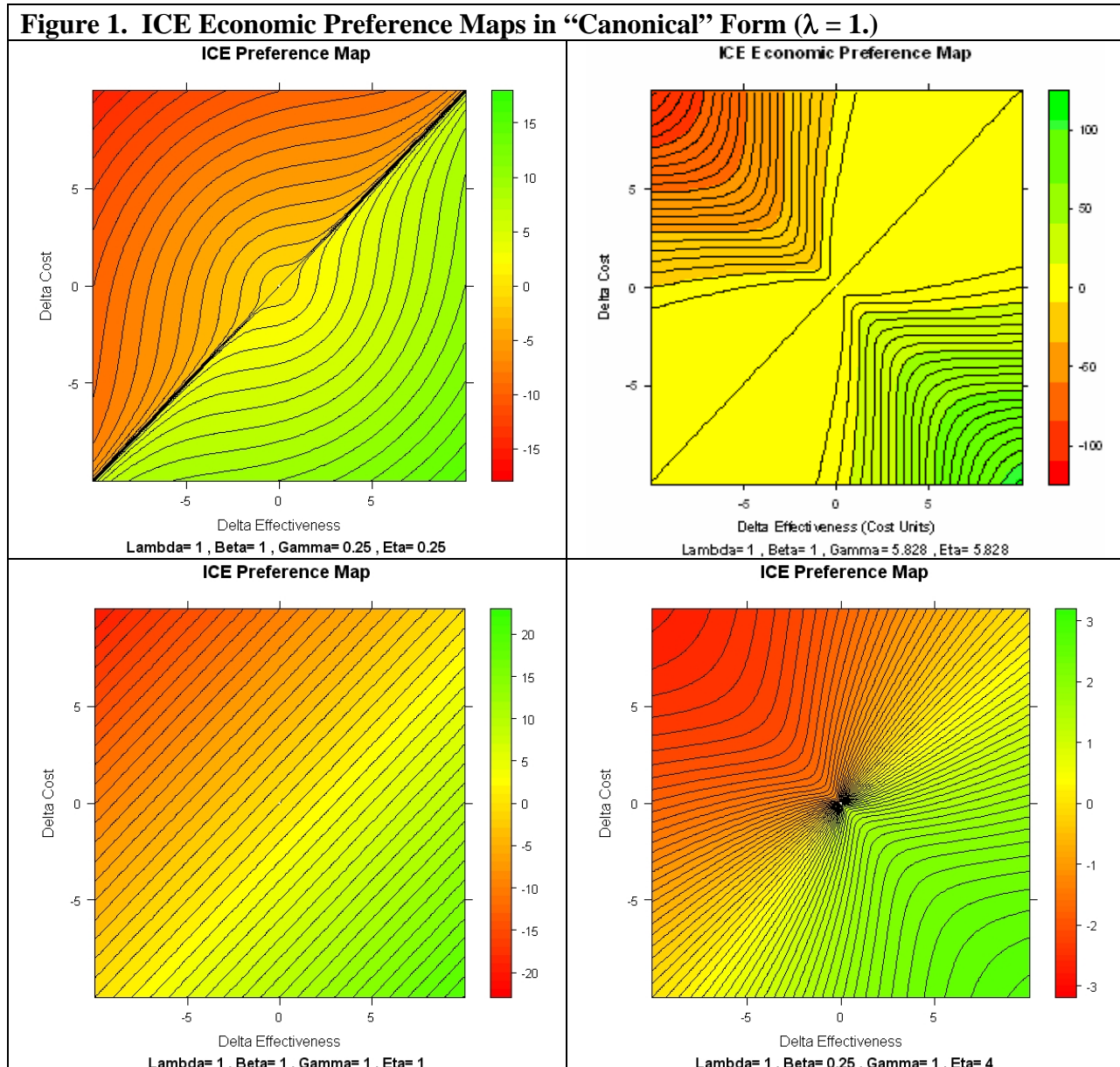
In summary then, only the nonlinear ICE preference maps of form (1) with power parameter ratio,  $\eta = \gamma/\beta$ , confined to the finite interval of  $1 < \eta \leq \Omega = 3 + 2\sqrt{2}$  can be fully realistic. And only the maps with  $\eta = \Omega$  allow willingness (standardized or un-standardized) to vary all of the way from 0 to  $+\infty$ .

Preliminary standardization, in which  $(x, y)$  represents either  $(\lambda\Delta E, \Delta C)$  in cost units or  $(\Delta E, \Delta C/\lambda)$  in effectiveness units for a fixed numerical value of  $\lambda$ , is essential to be able to express not only standardized directions,  $s = y/x$ , but also standardized willingnesses,  $w$ , as unitless quantities. In turn, unitless quantities become absolutely essential when *reciprocals* are to be compared. After all, if the quantities being compared were not unitless, the statistic and its reciprocal would be expressed in different units, such as  $\$/\text{QALY}$  and  $\text{QALY}/\text{\$}$ , and clearly could not then be directly compared!



### 3.5 Visualizing Indifference Curves using ICEepmap() and ICEomega()

The four subplots of Figure 1, below, depict ICE preference maps using equally spaced “indifference curves” (alternatively, called “iso-preference contours” or “level curves”). Whatever was the specified value of  $\lambda$ , what we now wish to visualize is the standardized form that results from using that specified  $\lambda$  to either scale both axes in “cost” units or else both in “effectiveness” units. As a result, the true  $\lambda$  must now be visualized as having been transformed to  $\lambda = +1$ . These R-graphics use the `contourplot()` function of the “lattice” package, [28].



Rather “round” maps like the one at the top-left of Figure 1 result when  $\gamma < \beta$ . Unfortunately, the slopes of the indifference curves below the lower-left to upper-right diagonal of the ICE plane,  $x$

= y, show that WTA < WTP in these rather “round” maps. This is quite unrealistic in the sense that this ordering has apparently never been observed in empirical studies!

The Linear map with constant Returns-to-Scale,  $NB(x, y) = x - y$ , results when  $\gamma = \beta = 1$ , which is the case displayed in the lower-left panel of Figure 1.

The “highly directional” maps on the right-hand side of Figure 1 result when  $\gamma > \beta$  and tend to be much more realistic than those on the left-hand side. Note that the  $\gamma/\beta$  ratio of 0.25 for the top-left map in Figure 1 is well above the  $1/\Omega \approx 0.1716$  lower limit allowed under restriction (2), while  $\gamma/\beta = 4$  at the bottom-right of Figure 1 is well below the  $\Omega \approx 5.828$  upper limit for maps possessing ICE monotonicity. The upper-right map of Figure 1 is the ICE- $\Omega$  map with constant Returns to Scale,  $\beta = 1$  and  $\eta = \gamma/\beta = \Omega$ .

### 3.6 Implications of Using a “Wrong” Numerical Value for $\lambda$ .

The ICEepmap( ) and ICEomega( ) functions can also be used to visualize the consequences of using a “wrong” numerical value for the Shadow Price of Health,  $\lambda$ . Specifically, consider the representation:

$$\lambda = \lambda_f \times \lambda_o \quad (6)$$

where  $\lambda_o$  represents the true value of  $\lambda$  in cost/effectiveness units and  $\lambda_f$  is a scalar (unitless) factor the one can pretend has a known value. In other words, one is using the “right” numerical value for  $\lambda$  when  $\lambda_f = 1$  and a “wrong” value when  $\lambda_f > 1$  (i.e.  $\lambda$  “too large”) or  $\lambda_f < 1$  (i.e.  $\lambda$  “too small.”) Thus, although the “standardization” process supposedly transforms  $\lambda$  to become 1, the consequences of using a  $\lambda_f \neq 1$  are easily visualized using the ICEepmap( ) or ICEomega( ) functions.

Curiously, all generalized-linear ICE maps (i.e. maps with  $\gamma = \beta$  or  $\eta = 1$ ) are sufficiently naïve and simplistic that nothing “appears” to go wrong when a factor of  $\lambda_f \neq 1$  is used! Specifically, as depicted in Figure 2, the direction of steepest-ascent and steepest-descent in generalized-linear preferences simply rotates to correspond to the direction orthogonal to  $\lambda_f$ , with slope  $-(\lambda_f)^{-1}$ .

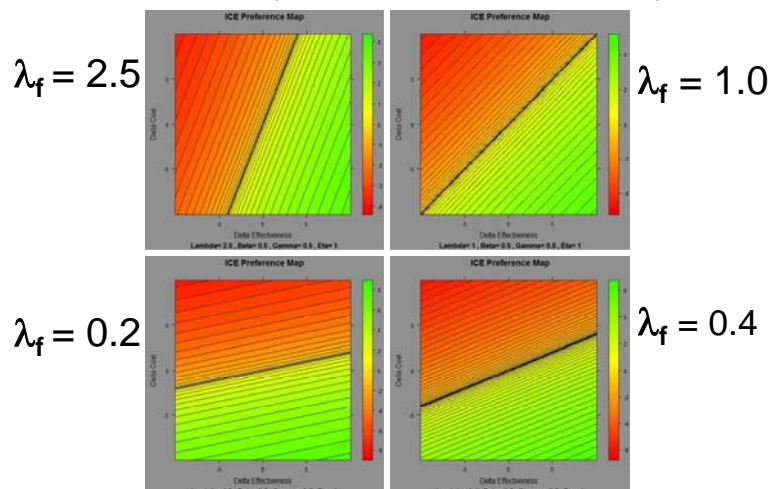
In other words, generalized-linear maps always remain symmetric (Fourth Axiom) relative to the  $-(\lambda_f)^{-1}$  direction for all values of  $\lambda_f$ .

In sharp contrast, as seen in Figure 3, realistic, highly-directional maps ( $\gamma > \beta$  or  $\eta > 1$ ) get “twisted” when  $\lambda_f \neq 1$ . Specifically, the direction of steepest-ascent and steepest-descent in realistic non-linear preferences tends towards having slope  $-\lambda_f$  ...instead of slope  $-(\lambda_f)^{-1}$ , orthogonal to  $\lambda_f$ .

The main implication here is that ICE Symmetry (Fourth Axiom) is assured only if  $\lambda_f = 1$  when using any **truly realistic, non-linear** ICE Preference Map. In fact, this objective property can be viewed as the very “definition” for the unknown, true Shadow Price of Health,  $\lambda = \lambda_o$ .

**Figure 2.**

**Generalized Linear ( $\eta = 1.0$ ) but Diminishing Returns ( $\beta = 0.5$ )**

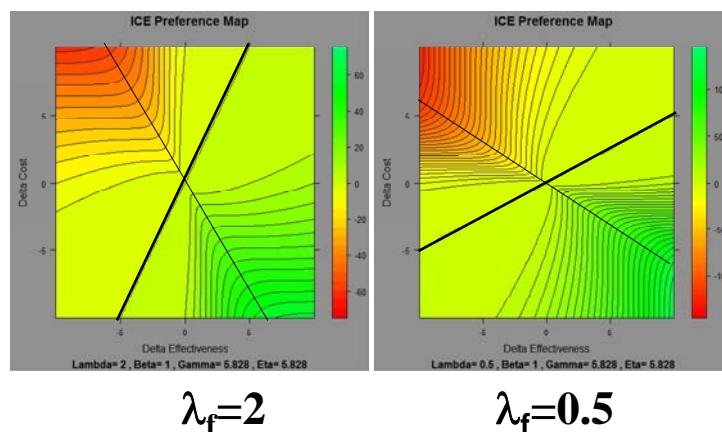


Note that the directions of steepest ascent and descent remain orthogonal to  $\lambda$ .

**Figure 3.**

**Highly Non-Linear:  $\beta = 1$ ,  $\gamma = \eta = \Omega$**

**“Wrong”  $\lambda \Rightarrow$  No Symmetry!**



Furthermore, the “link” function,  $\lambda = \sqrt{\text{WTP} \times \text{WTA}}$ , (5), provides a fully objective way to estimate this “symmetrizing” Shadow Price via existing empirical methods of eliciting paired “kinked” values of  $\text{WTP} < \text{WTA}$ , [18,30,17].

For example, assume that  $\lambda$  really is the often quoted value of \$50,000/QALY but that government authorities or local payers in some region insist that \$10,000/QALY is the maximum additional cost that they can possibly agree to pay. This is a simple budget constraint that *does nothing to change the full, fair shadow price of health*, but it does reduce the local WTP to  $\lambda/5$ . It is unfair and arbitrary to assume that a budget (extra cost) restriction like this also reduces WTA to  $\lambda/5$ . Instead,  $\lambda = \sqrt{\text{WTP} \times \text{WTA}}$  implies that the corresponding “fair” value of WTA

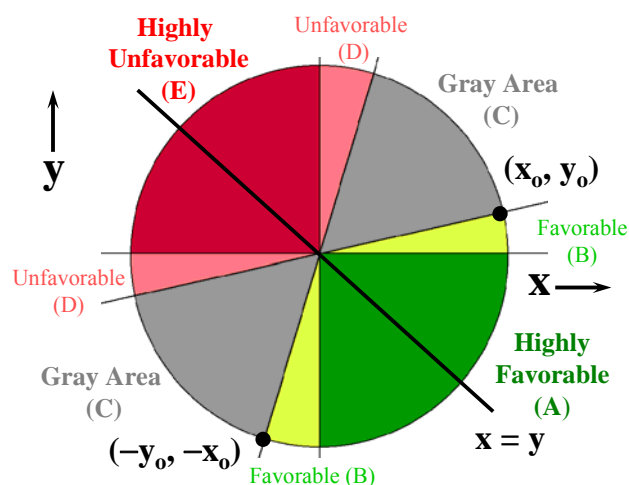
then becomes  $5\lambda = \$250,000/\text{QALY}$ . Only treatments that reduce cost (by reducing effectiveness) by at least  $\$250,000/\text{QALY}$  have preferences to society as high as the clearly desirable treatments that increase cost (and increase effectiveness) by less than  $\$10,000/\text{QALY}$ .

### 3.7 Laupacis' visualization of ICE preferences corresponds to the limit as $\beta$ approaches 0.

Figure 4 depicts the limit of our signed-power family of ICE preference maps of equation (1) as the returns-to-scale parameter,  $\beta$ , approaches *zero* (while the  $\gamma$  parameter is held fixed at any finite value.) In other words,  $\eta$  then approaches  $+\infty$  and the standardized willingness of equation (4) becomes  $w = (s + s^2)/(1 + s) = s$  in this limit.

While failing to possess ICE monotonicity and allowing negative values for  $w$  in equation (4) within the SE and NW quadrants, this limiting map still has [i] iso-preference curves that correspond to pairs of dual rays with reciprocal slopes, and [ii] *orders preferences on ICE polar angle in the exact same way* that they are ordered on all of our  $\beta > 0$  maps for outcomes at the *same ICE radius*.

Figure 4. Colored “Pie Chart” graphic depicting the ICE Plane as being divided into wedge-shaped regions by Laupacis et al. [16].



Note that Figure 4 also illustrates the symmetry axiom,  $P(x_0, y_0) = P(-y_0, -x_0)$ . Furthermore, one sees that the pair of dual ICE Rays separating the Yellow and Gray wedges have reciprocal standardized ICE slopes:  $y_0/x_0$  in the NE quadrant and  $(-x_0)/(-y_0)$  in the SW quadrant.

On the other hand, this limiting ( $\beta = 0$ ,  $\gamma > 0$ ) map is not very realistic precisely because it implies zero Returns to Scale. It ignores ICE radius as a potential, partial determinant of preference (especially within the **SE** and **NW** quadrants!)

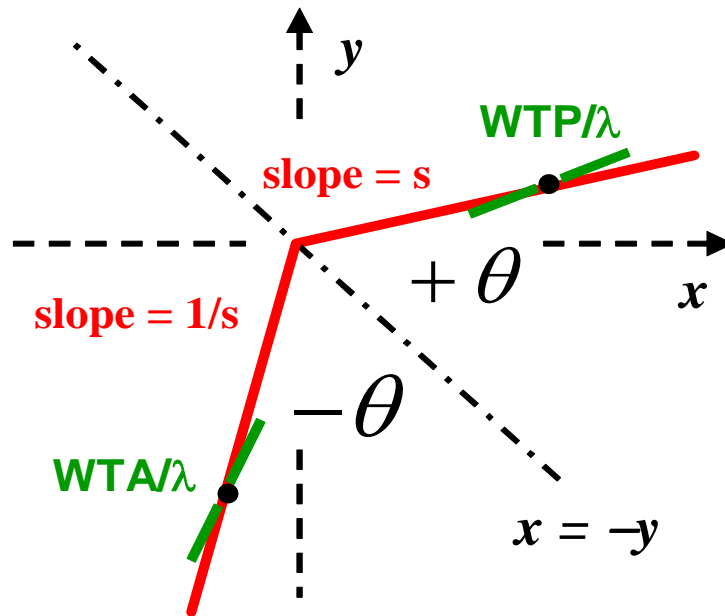
### 3.8 The Relationship between the Slope of the ICE Ray through a Point and the Willingness Rate at that Point

In realistic, nonlinear ICE preference maps, the willingness rate [slope of the indifference curve] through a point  $(x, y)$  is generally different from the standardized ICE ratio,  $s = y/x$ , of the ICE ray passing through that same point. The analytic geometry of this situation is as illustrated in Figure 5, which is a generalization of Figure 4 that allows for realistic, nonlinear preferences.

Figure 5 depicts a pair of standardized “dual rays” in **red**. Each ray subtends the same absolute ICE polar angle,  $|\theta|$ , relative to the  $x = -y$  diagonal. Each also contains the same distribution of ICE preference strengths (as a function of ICE radius) in the same direction (here “new” over “std” because  $|\theta| < 90^\circ$ .) And the rays have standardized slopes ( $s = y/x$ ) and standardized willingnesses ( $w = \text{WTP}/\lambda$  or  $\text{WTA}/\lambda$ ) that are numerical reciprocals.

Note that the “link” function, (5), does not actually establish numerical values for WTP, WTA,  $\eta$ ,  $w$  or  $s$  but only a relationship between WTP, WTA and the shadow price of health,  $\lambda$ . In other words,  $\text{WTP} = \text{WTA} = \lambda$  is always one possibility. Most importantly,  $\lambda = \sqrt{\text{WTP} \times \text{WTA}}$  illustrates that WTP and WTA can both vary within a map where  $\lambda$  is held fixed.

Figure 5 The case depicted here corresponds to  $0 < s < w < 1 < 1/w < 1/s$  because  $\eta = \gamma/\beta > 1$ .



#### 4. Transformations of the Bootstrap ICE Uncertainty Distribution within the Linear Subgroup

The material in Section §2.2, pages 4 and 5, describes the subgroup of linear transformations that arises naturally in ICE statistical inference. Here we address the question: “What happens to individual bivariate bootstrap ICE outcomes (resampled points on the ICE plane) under these transformations?”

From the so-called “alias” perspective commonly used in physics and most statistical data graphing software, each point is visualized as remaining at a fixed position in space as linear transformations occur. One of the resulting outcome coordinates does change because  $\lambda$  changes the scaling along either the x-axis (cost rescaling) or else along the y-axis (effectiveness rescaling.) All points are thus fixed but (partially) renamed (given aliases.)

Mathematicians (and most teachers) view transformations from the perspective in which all points literally move (are given alibis) relative to fixed axes with fixed scales (coordinate tick marks.) For example, each point then moves either left or right when  $\lambda$  changes and both ICE axes are measured in “cost” units. Similarly, each point moves either up or down when  $\lambda$  changes and both ICE axes are measured in “effectiveness” units.

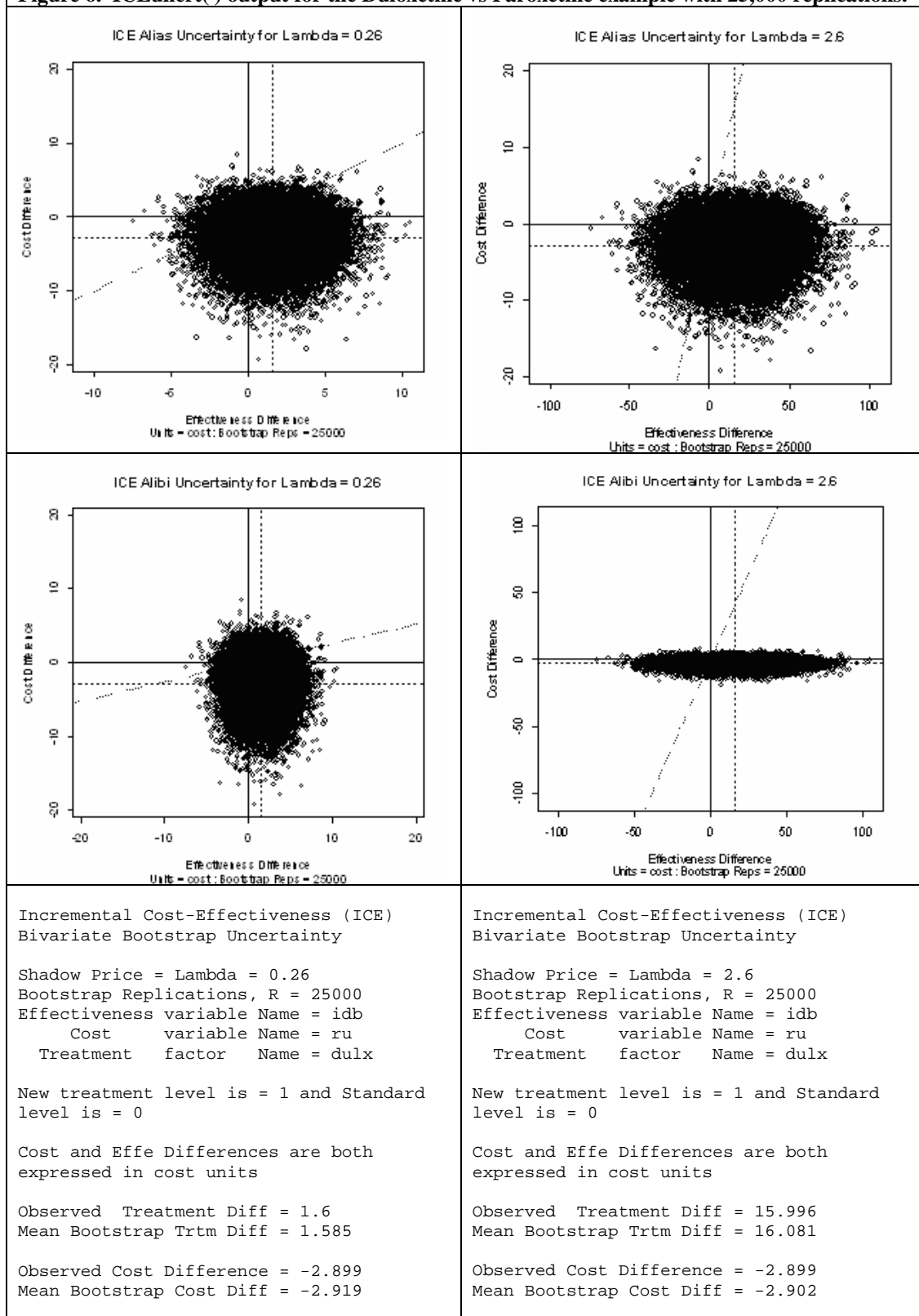
When alias and alibi perspectives and/or terminology are carelessly intermixed, it is well known that confusion and chaos can result.

The two plots displayed in the top row of Figure 6 use an alias perspective to illustrate that the Bootstrap distribution of ICE Uncertainty is “equivariant” (commutative) under the ICE linear subgroup of transformations. As  $\lambda$  increases from 0.026 (left side) to 0.26 (right side) in Figure 6, the bootstrap scatter is depicted as being essentially fixed in space or “unchanged.” Note that the x-axis scaling (tick mark spacing) also increases by (roughly) a factor of 10 from the left-hand panel to the right-hand panel of Figure 6. Again, this is the default ICEuncrt( ) “alias” plotting perspective (alibi = FALSE).

The correspondence between the left-hand and right-hand plots within Figure 6 would be similar (but not “exact” in the above weak sense) if the two scatters resulted from separate (independent) invocations of ICEuncrt( ) with different random number seed values. If either [1] a second invocation of ICEuncrt( ) uses the numerical “seed” value stored in the output list generated by the first invocation but a different value of “lambda =” or if [2] print.ICEuncrt( ) or plot.ICEuncrt( ) is invoked with its “lfact” parameter different from 1 (to simply transform the initial  $R \times 2$  matrix, **t**, of bootstrap (x, y) coordinates), the two resulting ICEuncrt( ) scatters will appear “identical” except for the scaling along one axis and the slope of the dotted line that represents the current choice for  $\lambda$ .

The plots displayed in the second row of Figure 6 illustrate, instead, the “alibi=TRUE” perspective. As  $\lambda$  again increases from 0.026 (left side) to 0.26 (right side), the points in the bootstrap scatter really spread out horizontally in these (default) “cost” unit graphics. There really is no vertical movement here in the sense that the vertical range of outcomes within the

**Figure 6. ICEuncrt() output for the Duloxetine vs Paroxetine example with 25,000 replications.**





scatter is roughly  $-20$  to  $+10$  in both second row plots. The vertical plotting range increased in the right-hand plot simply because its horizontal range was forced to increase ...in order to keep the scatter within alibi plotting limits. It is possible, of course, to use an alibi perspective to argue that the Bootstrap distribution of ICE uncertainty is “equivariant” (commutative) under the linear subgroup. However, those particular arguments strike me, at least, as being much less clear or “obvious” than the alias motivation.

## 5. A Wedge-Shaped, Equivariant ICE Statistical Confidence Region

The ICEwedge( ) function uses my methodology [19-21] for forming **statistical ICE confidence regions** and is sufficiently robust and reliable for routine use by health services researchers. The ICE confidence (or tolerance) regions of particular interest use bootstrap ICE angle order statistics around a circle to form a wedge-shaped region that has the ICE origin,  $(\Delta E, \Delta C) = (0, 0)$ , as at most a limit point and use a pair of ICE Rays as their clockwise and counter-clockwise limits. Specifically, my so-called “central” method “counts outwards” the same number of ICE Angle Order Statistics,  $\text{floor}(\text{reps} \times \text{conf}/2)$ , both Counter-Clockwise and Clockwise from the “center” Order Statistic (nearest the Observed ICE Ratio) to define a pair of ICE Ray endpoints at order statistics positions “jlo” and “kup”. The resulting wedge-shaped region is guaranteed to be **equivariant** under changes in  $\lambda$  ...in either its original (singly) unbounded form or when truncated at the minimum and maximum ICE radii observed within the confidence wedge.

The key practical implication of equivariance is that, when patient level  $(x, y)$  outcomes are transformed using alternative values for  $\lambda$  [but the same random number seed is used for bootstrap resampling], then every  $(x, y)$  point in the ICE uncertainty scatter always remains inside, on or outside of the resulting wedge-shaped ICE ray confidence limits. It is relatively easy to see this sort of “invariance” property in default ICEwedge( ) plots as  $\lambda$  changes because the plots use the “alias” perspective under which points appear to remain at fixed positions on the ICE plane. The concept still holds from the “alibi” perspective, but the resampled outcomes and the confidence wedge then BOTH move as  $\lambda$  changes ...making it somewhat more difficult to literally “see” that the claim is true.

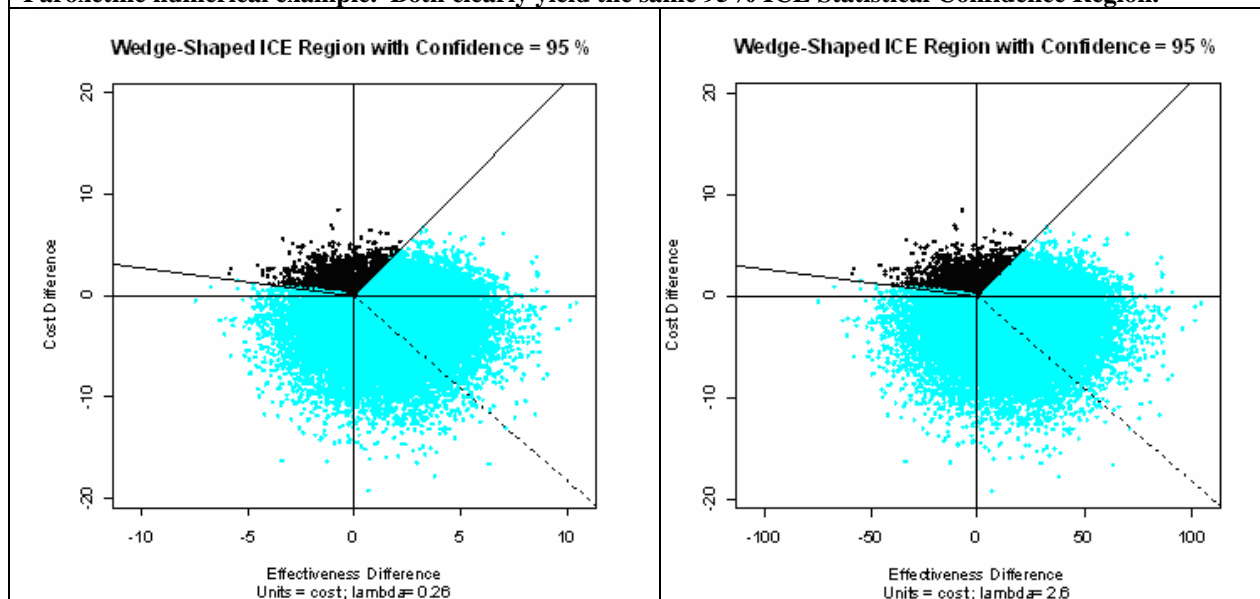
**KEY POINT:** Because bootstrap “count outwards” confidence regions have this equivariance property under changes in  $\lambda$ , I contend that they quantify only the **statistical** uncertainty within the two samples of patient level cost and effectiveness data about where the unknown true  $(\Delta E, \Delta C)$  outcome falls on the ICE plane. No uncertainty about choice of  $\lambda$  can possibly be captured by this **equivariant region**.

Technical Note: ICE Angles are computed by ICEwedge( ) from the “alibi” perspective in which  $(x, y)$  standardized differences are both measured in the **same** units for the current value of  $\lambda$ . Thus, for “cost” units,  $(x = \lambda \times \Delta E, y = \Delta C)$ . These numerical values for ICE Angles usually look “wrong” on (default) “alias” displays because the x-axis and the y-axis are then typically displayed using **different** scales. In other words, the 95% confidence regions for  $\lambda = 0.26$  and  $\lambda = 2.6$  displayed in Figure 7, below, compute ICE angles quite differently but clearly yield the SAME “count outwards” confidence region.



My experience with all numerical examples where a Fieller's theorem confidence interval for the ICE Ratio (at a stated confidence level) actually does exist (i.e. has limits that are real rather than imaginary), my “count outwards” wedge-shaped region is in quite good agreement with the “correct” half of the corresponding Fieller's bow-tie region.

**Figure 7. Alias visualizations for 2 different  $\lambda$ -scalings of the “high uncertainty” Duloxetine versus Paroxetine numerical example. Both clearly yield the same 95% ICE Statistical Confidence Region.**



For the “high uncertainty” Duloxetine vs Paroxetine example, bootstrap ICE outcomes within the 95% confidence ICE statistical wedge subtend an ICE Polar Angle greater than 180°! Both the bootstrap scatter of ICE uncertainty and the statistical wedge will always appear “unchanged” in a pair of “alias” plots for different values of  $\lambda$ , like those above. Note that the x-axis scaling (tick mark spacing) always changes by an appropriate factor from the left-hand panel to the right-hand panel.

The `print.ICEwedge( )` function reports this polar angle to be 260° when  $\lambda = 0.26$  (“alibi” perspective, 72.2% of 360°) and 193° when  $\lambda = 2.6$  (“alibi” perspective, 53.6% of 360°.) Unfortunately, neither of these values for ICE polar angle agrees with what it actually “appears” to be for the two “alias”  $\lambda$ -scalings depicted in Figure 7. An invocation of `ICEwedge( )` with `lfact = 0` causes calculation of the (approximate) value of  $\lambda$  that causes the “alias” and “alibi” perspectives to be in agreement (as well as making time consuming calculations and sorting of the implied ICE Angles.) For the `dulxparx` data, this value is  $\lambda = 0.454$ , and the implied ICE Polar Angle is then 240.7° (66.9% of 360°.)

## 6. Acceptability: VAGR and ALICE Curves

The concept of an “acceptability curve” graphic was proposed by Van Hout, Al, Gordon and Rutten (VAGR) [29] in 1994 to portray ICE uncertainty. Given either [i] a parametric, bivariate distribution (normal, say) with mean ( $\Delta E$ ,  $\Delta C$ ) that has been fitted to some observed patient-level data or else [ii] a bootstrap resampling distribution of ICE uncertainty, the VAGR curve depicts the estimated “confidence level” associated with the region to the right or below a rotating *straight line through the ICE origin* that starts out horizontal (representing  $WTP = 0$ ) and rotates counter-clockwise by  $90^\circ$ , ending up being vertical (representing  $WTP = +\infty$ .) In 2004, Fenwick, O’Brien and Briggs (FOB) [7] cataloged 13 somewhat ill-defined “special cases” yielding VAGR curves with quite different shapes, ranging from rather flat, to increasing, to decreasing, to distinctly non-monotone.

An acceptability curve that is always *monotone non-decreasing* results from the unpublished alternative definition of acceptability independently proposed by me in 2001 and by Professor Ken Buckingham of Otago University, New Zealand, in 2003. My “ICEplane” Windows software [20] and my ICEinfer R-package both use Buckingham’s terminology: *Acceptability Levels In Cost Effectiveness* (ALICE) curves. For any *given and fixed* positive value of  $\lambda$ , the ALICE frontier is defined using a *pair of “linked” dual ICE rays* (i.e. rays that remain symmetric relative to the  $x = -y$  diagonal while rotating so that their absolute ICE polar angle,  $|\theta|$ , is constantly increasing, as in Figure 5, page 13.) Table 3 compares the VAGR and ALICE definitions within all four quadrants of the ICE plane.

**Table 3. VAGR and ALICE definitions of Acceptability  
when  $\lambda$  is held fixed while the scalar  $s$  varies from 0 to  $+\infty$ .**

ICE quadrant	VAGR Definition	ALICE Definition
$\Delta C > 0, \Delta E > 0$ NE quadrant	Acceptable if $\Delta C / \Delta E < \lambda s$	Acceptable if $\Delta C / \Delta E < \lambda s = WTP$
$\Delta C < 0, \Delta E > 0$ SE quadrant	All outcomes are Acceptable	All outcomes are Acceptable
$\Delta C < 0, \Delta E < 0$ SW quadrant	Acceptable if $\Delta C / \Delta E > \lambda s$ (see Note)	Acceptable if $\Delta C / \Delta E > \lambda/s = WTA$
$\Delta C > 0, \Delta E < 0$ NW quadrant	No outcomes are Acceptable	No outcomes are Acceptable

Note: Decision Rule (4), page 176, of O’Brien et al. [17] contains a typographical error; the SW quadrant rule is not identical to the NE quadrant rule,  $\Delta C / \Delta E < \lambda s$ .

In the notation of Table 3, the standardized ICE slope,  $s$ , is a unitless scalar that increases from 0 towards  $+\infty$ , while  $\lambda$  denotes a *given, fixed* value for the shadow price of health. Within the VAGR column of definitions, the product,  $\lambda$  times  $s$ , thus denotes a *variable* quantity corresponding to different *common values* for shadow price = WTP = WTA defining different *linear* VAGR thresholds for acceptability. Within the ALICE column of definitions, WTP =  $\lambda s$  increases with  $s$  within the NE quadrant while WTA =  $\lambda/s$  simultaneously decreases with  $s$  within the SW quadrant, defining a range of *kinked* ALICE thresholds for acceptability, clearly satisfying the *link* equation, (5).

Note that his ALICE definition of acceptability agrees also with the sum of double integrals in Willan et al [30], page 3255.

Technically, interest could even be restricted to the finite range  $0 \leq s \leq 1$  for ALICE curves because  $s > 1$  corresponds to  $WTA < \lambda < WTP$ , an ordering that has apparently never been observed empirically.

Note in Table 3 that the VAGR and ALICE definitions of acceptability *differ only within the SW quadrant*. VAGR and ALICE curves thus contain the same basic information (displayed using different horizontal axes) whenever the ICE uncertainty distribution attributes zero credibility to the SW quadrant. At the other extreme, where 100% credibility is attributed to the SW quadrant, the VAGR and ALICE curves are again equivalent, but the VAGR curve would then be decreasing while the ALICE curve is increasing, as usual. In other cases, the VAGR curve will usually be non-monotone and biased. In fact, cases where the ICE uncertainty distribution attributes credibility not only to the SE quadrant but also to the most desirable parts of both the SW and NE quadrants are actually quite important!

Let us now consider the “dulxparx” numerical example of the above “high uncertainty” type. For simplicity, we have assumed that  $\lambda = \$0.26/\text{Week}$  per unit of idb = Integrated Decrease from Baseline in HAMD-17 Score. Note in the top-right panel of Figure 8 that this example illustrates a key situation. Relative to the standard treatment (paroxetine), the new treatment (duloxetine) here could represent what is commonly known (somewhat derisively) as a “me too” treatment for MDD. Specifically, the bootstrap distribution of uncertainty here covers the ICE origin and lends considerable credibility to at least 3 of the 4 ICE quadrants, at least when the \$/Week difference in medication *acquisition cost* is zero, as assumed here and in [25].

Our objective in exploring this particular case-study example is to convince you that traditional VAGR acceptability curves are biased towards their average value in these critical “high uncertainty” cases. In particular, the all-important lower values of VAGR acceptability are then biased upwards.

The top-left panel of Figure 8 displays the non-monotone VAGR acceptability curve for our high uncertainty example that corresponds to a relatively wide range (from 0 to 6) for the unknown common value of WTP =  $\lambda s$  = WTA. Only one numerical value within this wide range of alternative values for  $\lambda s$  can correspond to the “true” shadow price of health.

It is not clear how VAGR or FOB themselves would interpret the information provided by a plot like Figure 8. Because acceptability is always rather high ( $> 0.80$ ) over the finite range displayed here and the curve is also rather flat ( $\max - \min < 0.09$ ), outcomes researchers might conclude that [i] choice of  $\lambda$  is relatively unimportant here and/or that [ii] the odds that the new treatment is more cost-effective than standard are at least 4:1 because  $0.80/0.20 = 4$ .

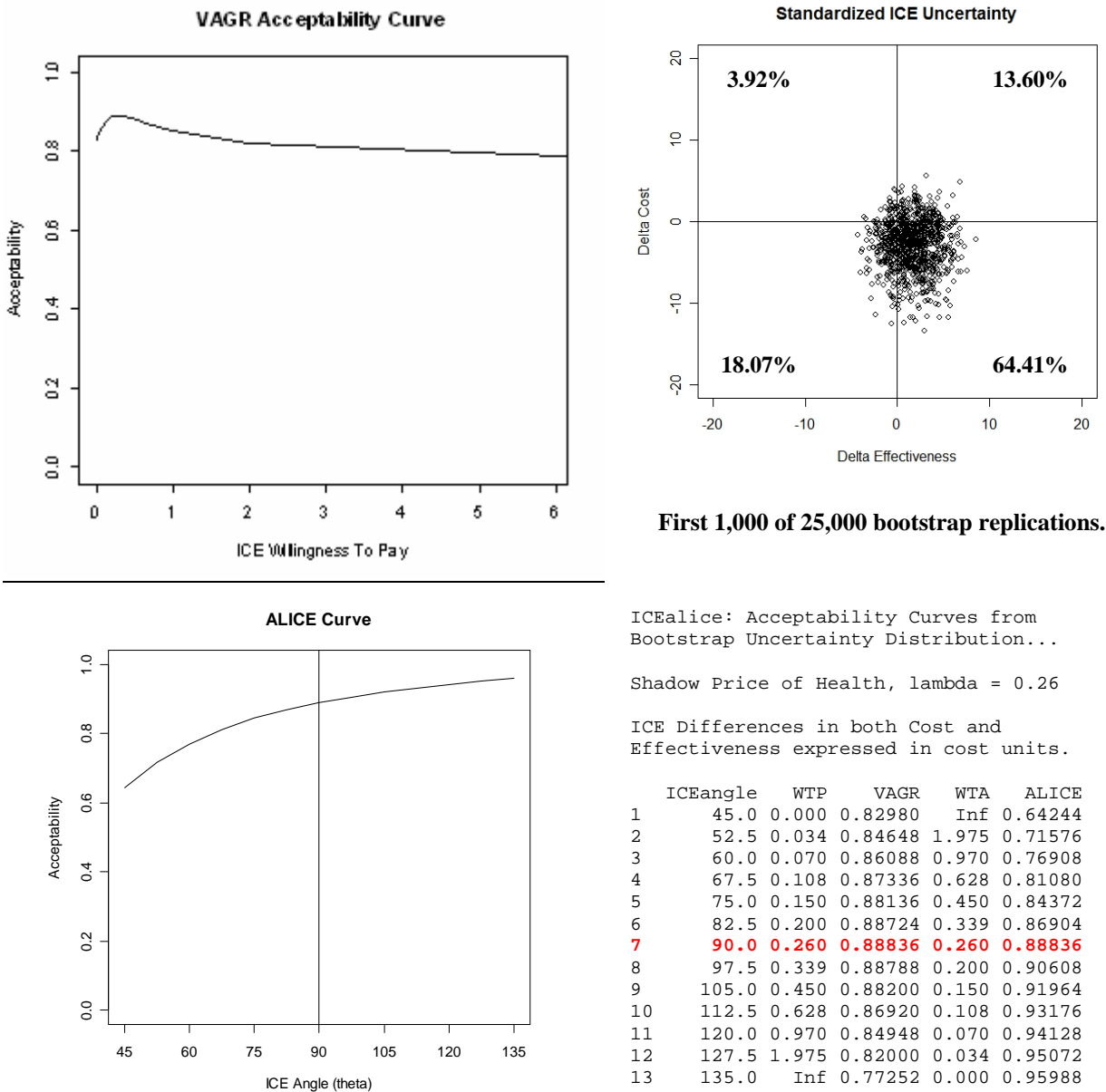
The bottom-left panel of Figure 8 displays the corresponding, monotone ALICE curve for this high uncertainty example. By covering the *finite range* for absolute ICE angles of  $45^\circ \leq |\theta| \leq 135^\circ$ , the full *infinite range* of  $0 \leq s \leq +\infty$  is easily visualized here. While the plot assumes that  $\lambda = 0.26$  \$/Week/idb is the fixed, most relevant value for the true shadow price of health, it also allows an overall acceptability level to be determined for all possible budget constraints of the form  $WTP = \lambda s$  in the NE quadrant with  $s < 1$  plus, by symmetry,  $WTA = \lambda/s$  within the SW quadrant.

Using different choices for  $\lambda$  will result in different ALICE curves!!! On the other hand, all alternative ALICE curves for a given set of data have the same starting and ending points at  $|\theta| = 45^\circ$  ( $s = 0$ ) and  $|\theta| = 135^\circ$  ( $s = +\infty$ .) Namely, the smallest ALICE value (0.6441 here) will always be the estimated confidence that the new treatment is both “less costly AND more effective” than standard, while the largest ALICE value (0.9608 here) will always be the estimated confidence that the new treatment is either “less costly OR more effective” than standard. These limits correspond to the two key ICE quadrant confidence levels used to quantify “statistical dominance” levels [24].

Note that  $WTP = 0.26$  \$/Week/idb yields the maximum VAGR acceptability (of 0.8870) in the left-top panel of Figure 8. Thus, it follows that no ALICE curve for any alternative value of  $\lambda$  (different from 0.26 \$/Week/idb) can yield a larger acceptability level than the value (of 0.8870) displayed in the upper-right panel of Figure 8 at  $|\theta| = 90^\circ$  ( $s = 1$ ). We are definitely not recommending that the numerical value of  $\lambda$  used to define ALICE levels be routinely chosen in this particular way. After all, this specific choice of  $\lambda$  is, in a weak sense, always “most favorable” to the new treatment!

Rather, our point here is simply that VARG acceptability curves, by using only alternative *linear frontiers* ( $WTP = \lambda = WTA$ ), are badly biased in all high uncertainty cases where the VARG curve ends up being rather flat and/or non-monotone. ALICE curves are then much less biased (upwards or downwards) because they use realistic *linked frontiers*. Even the ALICE curve that is biased upwards *as much as possible*, as in Figure 8, still suggests that administrative budget constraints (that reduce WTP and, when fair, also increase WTA) can *drastically decrease the overall acceptability level* of new over standard. This reduction is from 0.8870 at  $|\theta| = 90^\circ$  ( $s = 1$ ) to 0.6441 at  $|\theta| = 45^\circ$  ( $s = 0$ ) in Figure 8, which is a reduction in confidence of 0.243. After all, for two exactly *equivalent treatments*, the VAGR acceptability is expected to always be 0.50 for all values of WTP. The corresponding ALICE level would then also be expected to equal 0.50 at  $s = 1$  ( $|\theta| = 90^\circ$ ), but it would be *expected to drop* to 0.25 at  $s = 0$  ( $|\theta| = 45^\circ$ ) as well as *to rise* to 0.75 at  $s = +\infty$  ( $|\theta| = 135^\circ$ ), at least when cost and effectiveness differences are uncorrelated.

Figure 8. Acceptability Curves for the “high uncertainty” Duloxetine versus Paroxetine numerical example when  $\lambda$  either varies (VAGR formulation) or the Shadow Price of Health is held @  $\lambda = 0.26$  (ALICE formulation.)

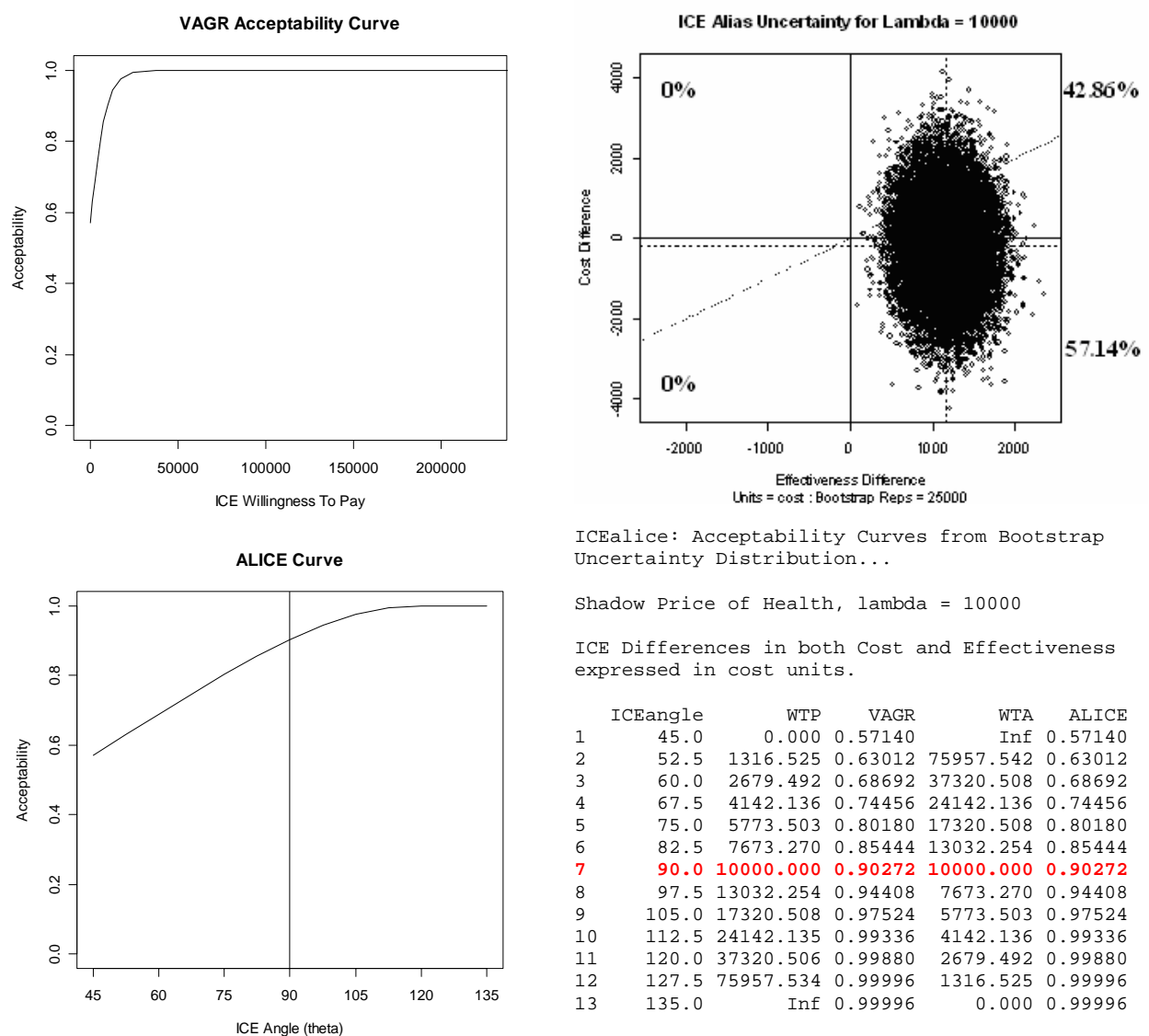


In all cases where the ICE uncertainty distribution lends credibility to only one quadrant or to at most two adjacent quadrants, the information contained in VAGR and ALICE curves will really be equivalent (even when the VAGR curve is decreasing due to increases in WTA.) In these relatively simple (lower uncertainty) cases, VAGR acceptability is not really biased relative to the corresponding ALICE level.

ALICE curves always concentrate attention upon the *uncertainty within the available data* supporting an ICE policy decision rather than upon any uncertainty about  $\lambda$  itself. Whenever a

VAGR curve is non-montone, it is actually also depicting additional uncertainty about  $\lambda$ . When a VAGR curve is monotone, Table 3 shows that it can be reinterpreted as corresponding to a fixed value of  $\lambda$ . For example, when a VAGR curve is non-decreasing, it can be reinterpreted as displaying the uncertainty associated with values of WTP less than any value of ICE Ratio =  $\lambda$  within the plotting range. When a VAGR curve is non-increasing, it can be reinterpreted as displaying the uncertainty associated with values of WTA larger than any value of ICE Ratio =  $\lambda$  within the plotting range.

**Figure 9.** The VAGR Acceptability Curve is actually equivalent to the ALICE curve for the “fluoxtea” numerical example. Actually, one must examine the lower-right numerical listing to see this; this is really not that clear from the plots with very different horizontal axes!



## 7. Additional Uncertainty due to dependence of Economic Preferences on the Shadow Price of Health

To this point, we have concentrated upon the “positive” and “desirable” features of (linear and nonlinear) maps that quantify economic preferences. Again, those properties hold within a **fixed  $\lambda$  context**.

In current ICE practice,  $\lambda$  is given several different names and descriptions but is always said to be an **unknown** constant. Health services researchers quite rightly interpret use of the “u”-word as justification for playing “what-if” scenarios where  $\lambda$  is **varied over a relatively wide numerical range**. Unfortunately, economic preference maps (linear and nonlinear) have undesirable characteristics in this **varying  $\lambda$  context**.

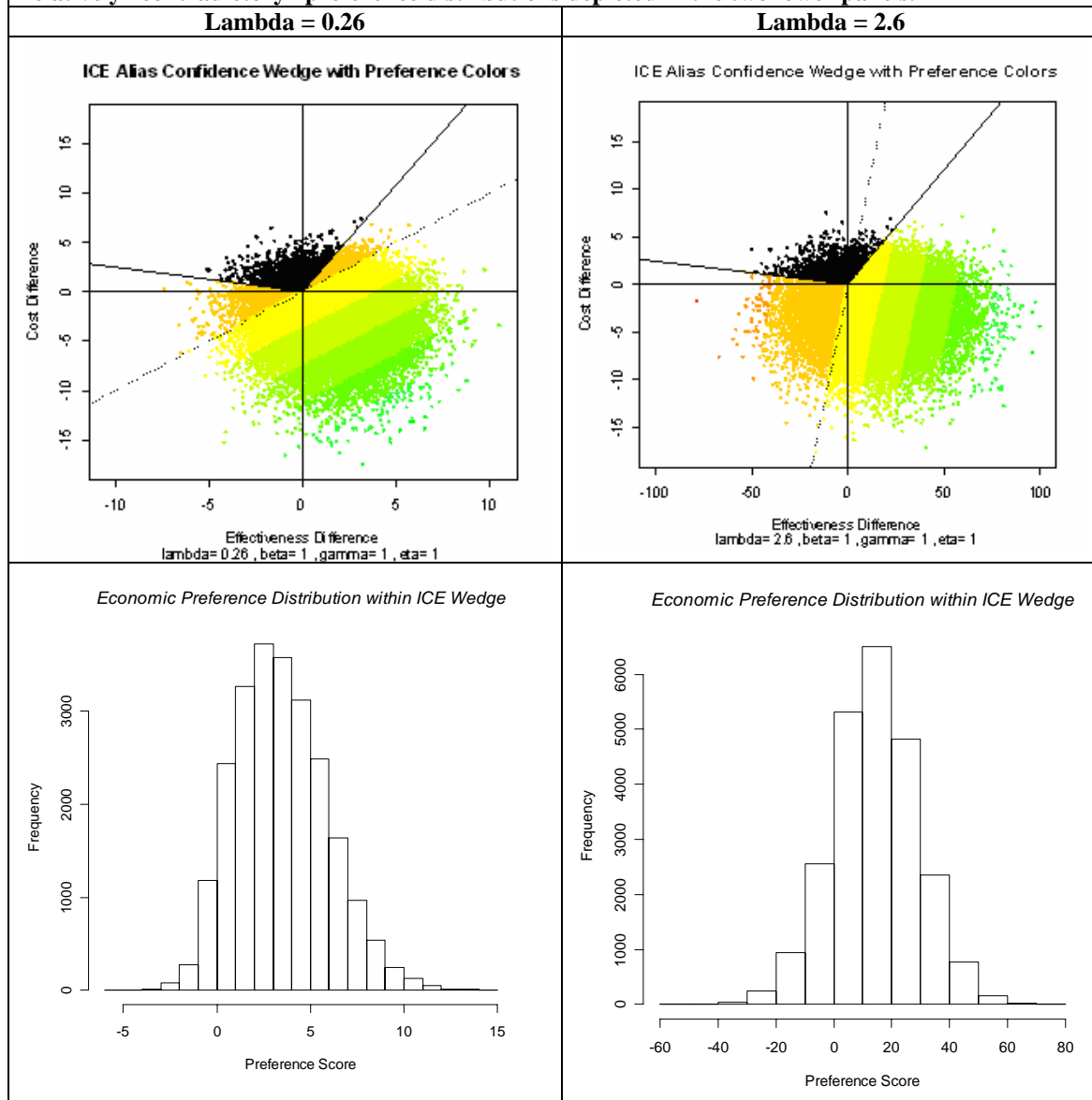
Here, we show that **economic** uncertainty [about choice of  $\lambda$ ] can not only (i) **swamp statistical** uncertainty [about where the unknown true ( $\Delta E$ ,  $\Delta C$ ) outcome might fall on the ICE plane] but usually also (ii) encompasses mutually **contradictory** possibilities (injects **inconsistency**.) While I am definitely not saying that routine **sensitivity analyses** should not be performed, I do think that health services researchers need to be much more aware of the extent of the trauma thereby injected into ICE **statistical** inference ...i.e. the conflicting assumptions implied by **economic** uncertainty about  $\lambda$  goes well beyond the uncertainty within observed patient level outcome data.

We will explore the dramatic extent of trauma introduced into an equivariant **statistical** confidence region by superimposing alternative **economic** preference maps corresponding to very different numerical values of  $\lambda$  using the ICEcolor( ) function.

Researchers wishing to quantify the **overall statistical plus economic** uncertainty (including choice of  $\lambda$  as well as, possibly, also choice of  $\beta$  and  $\gamma$  or  $\eta$ ) should compute a numerical **economic preference score distribution** for only the (x, y) points in the ICE uncertainty scatter that fall **inside or on** the **statistical** wedge ICE confidence limit rays. These economic preference distributions are displayed as histograms by plot(ICEcolor), as illustrated below.

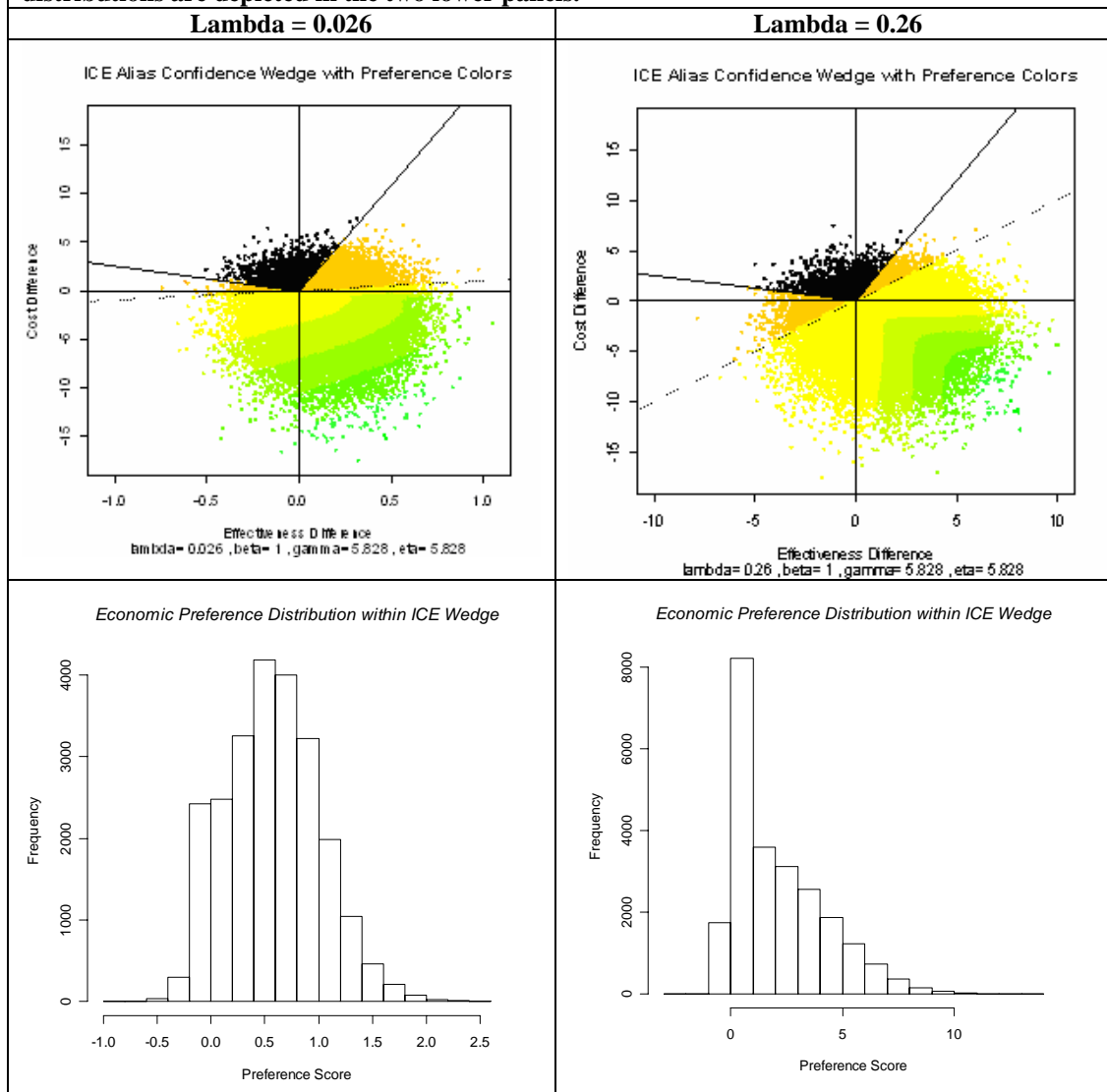
[The 95% confidence Fieller’s theorem bow-tie region is imaginary in the “high uncertainty” example illustrated here.]

Figure 10. Here, the statistical wedges are overlaid with purely Linear economic preference maps (Beta = Gamma = Eta = 1) that add variation which cannot be ignored. Note the resulting pair of relatively “contradictory” preference distributions depicted in the two lower panels.



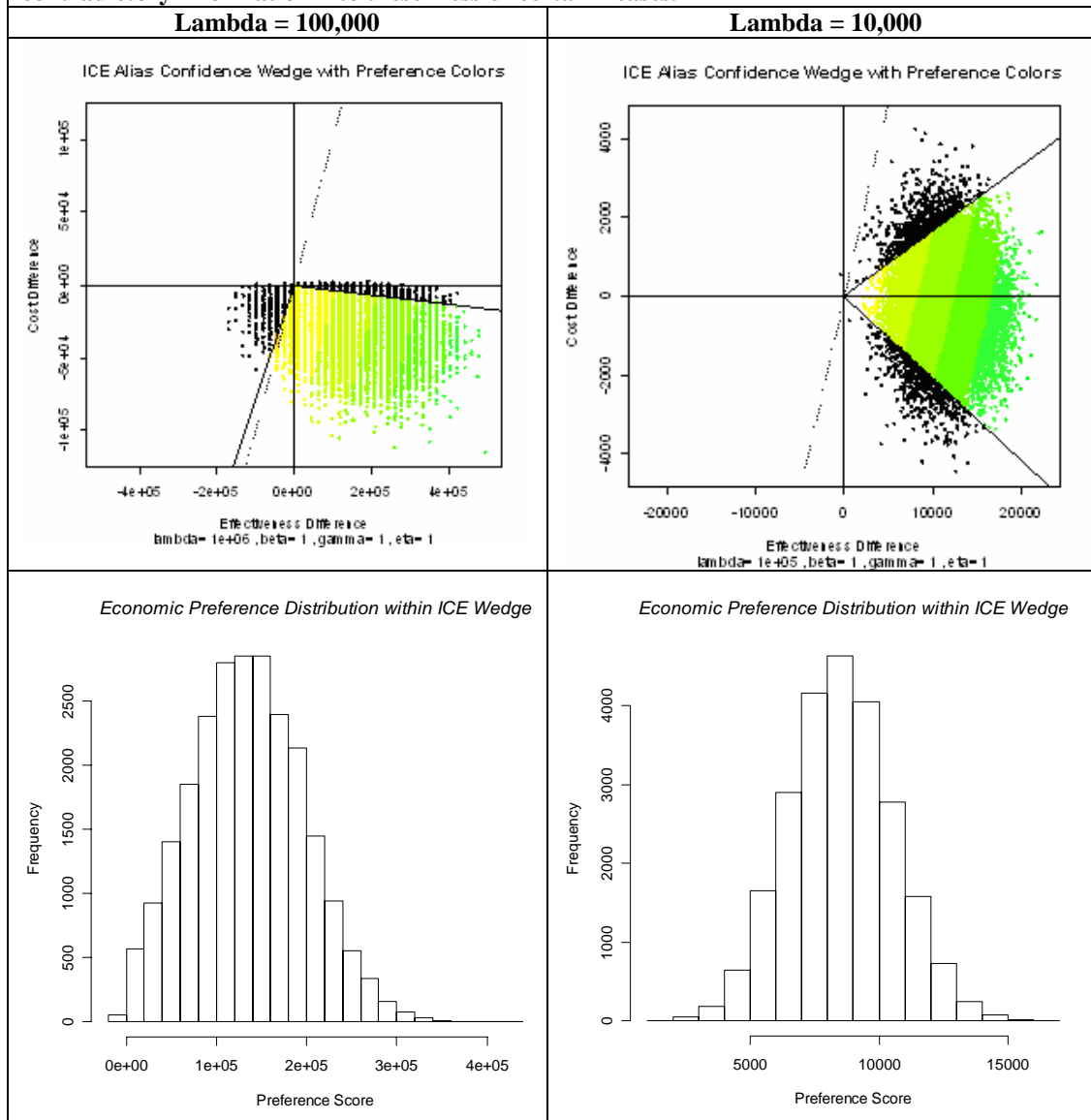


**Figure 11.** Again for the Duloxetine vs Paroxetine example and the bootstrap ICE outcome 95% confidence “Count Outwards” statistical wedge, the top pair of plots illustrate the “equivariance” perspective on increasing  $\lambda$  from 0.026 (left side) to 0.26 (right side.) Again, the bootstrap scatter of ICE uncertainty and the statistical wedge appear “unchanged” because the x-axis scaling (tick mark spacing) again increases by a factor of 10 from the top left panel to the top right panel. Here, the statistical wedges are overlaid with highly-directional nonlinear economic preference maps (Beta = 1, Gamma = Eta = 5.828) that again add variation which cannot be ignored. The corresponding pair of relatively “contradictory” preference distributions are depicted in the two lower panels.



The above figures clearly illustrate that combining **economic** uncertainty (about choice of  $\lambda$ ) with **statistical** uncertainty [about the true location of  $(\Delta E, \Delta C)$  within the ICE plane] can not only increase overall uncertainty but also embrace contradictions ...all due to imposition of systematic biases.

**Figure 12.** Here are the plot(ICEcolor) graphics for the “fluoxpin” and “fluoxtea” datasets. The statistical wedges are overlaid with linear economic preference maps (Beta = Gamma = Eta = 1), and the corresponding pair of preference distributions are depicted in the two lower panels. As in Figures 10 and 11, drastic changes in Lambda would also inject non-ignorable and self-contradictory information into these “less-uncertain” cases.



**Some Final Remarks:** No two really different values of  $\lambda$  can possibly both produce symmetric, nonlinear standardized outcomes. Two rather close values may possibly be both approximately correct. But any two really distant values are mutually exclusive and contradictory.

Health economists have apparently believed for years that “preferences are nonlinear” [13], and our nonlinear maps illustrate that WTP and WTA can both vary with  $\lambda$  held fixed via the

relationship  $\lambda = \sqrt{\text{WTP} \times \text{WTA}}$ . Furthermore, no single nonlinear map (i.e. specific choices for our  $\beta < \gamma$  parameters) needs to be singled out because all *symmetric, differentiable maps* necessarily satisfy the same geometric-mean relationship,  $\lambda = \sqrt{\text{WTP} \times \text{WTA}}$ .

Studies which elicit only WTP, rather than both WTP and WTA, have the potential to seriously bias  $\lambda$  downwards by assuming that  $\lambda = \text{WTP}$ . While unrealistically low values of  $\lambda$  are unduly restrictive within the NE quadrant, *really serious damage* can then result within the SW quadrant. After all, a decision maker might thereby assume that  $\text{WTA} = \text{WTP}$  and accept an unduly large decrease in effectiveness in exchange for a relatively small cost reduction.

Again, our maps *encompass the entire ICE plane* rather than focus on any particular sub-region. We have argued that our ICE preference maps with  $1 < \eta \leq \Omega = 3 + 2\sqrt{2}$  are both realistic and coherent. That property alone makes them invaluable to cost-effectiveness practitioners. It can be quite confusing and counter-productive to, instead, use different basic terminology [14] within the NE and SW quadrants. Similarly, the FOB suggestion [7] to divide the ICE plane up into many sub-regions is tedious and distracting.

In the “intervals or surfaces?” terminology of Briggs and Fenn[2], we are firmly of the opinion that ICE inference needs to be based upon 2-dimensional confidence regions (surfaces) rather than upon an infinite family of confidence intervals for any scalar value of preference as  $\lambda$  varies [27,3,26,15]. There is no current consensus about how geometrically simple (easy to define in written text) the boundary of an “ideal” ICE confidence region should be. Further, does this region need to be constrained to be finite in overall measure? Unlike the self-contradictory confidence bands generated in linear NB analysis as  $\lambda$  varies, wedge shaped confidence regions [1,19-22,5,23] have the potential to focus attention upon meaningful ICE sub-regions and to suggest clear preference-based actions. Most importantly, the “count outwards” wedge-shaped ICE confidence regions are defined [20-22] using ICE polar angle order statistics [5] in a way that makes them equivariant (commutative) relative to choice of  $\lambda$ . In other words, one would always end up with the very same confidence region for every choice of  $\lambda$ .

## ACKNOWLEDGEMENT

I wish to thank my Lilly colleagues, particularly Gerhardt Pohl and Joe Johnston, for their invaluable comments on numerous versions of these materials. I also wish to thank Ken Buckingham for personal communications about his give/get motivation for ICE symmetry.

## REFERENCES

1. Black WC. The CE plane: a graphic representation of cost-effectiveness. *Medical Decision Making* 1990; 10: 212-214.
2. Briggs AH, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane (Student Corner.) *Health Economics* 1998; 7: 723-740.
3. Chaudhary MA, Stearns SC. "Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. *Statistics in Medicine* 1996; 15: 1447-1458.
4. Cook JR, Heyse JF. Use of an angular transformation for ratio estimation in cost-effectiveness analysis. *Statistics in Medicine* 2000; 19: 2989-3003.
5. Copley-Merriman C, Egbuonu-Davis L, Kotsanos JG, Conforti P, Franson T, Gordon G. Clinical economics: a method for prospective health resource data collection. *Pharmacoeconomics* 1992; 1(5): 370-376.
6. Fan MY, Zhou XH. "A simulation study to compare methods for constructing confidence intervals for the incremental cost-effectiveness ratio." *Health Serv Outcomes Res Method* 2007; DOI 10.1007/s10742-006-0017-9.
7. Fenwick E, O'Brien BJ, Briggs AH. Cost-effectiveness acceptability curves – facts, fallacies and frequently asked questions. *Health Economics* 2004; 13: 405-415.
8. Gafni A, Birch S. Incremental cost-effectiveness ratios (ICERs): The silence of the lambda. *Social Science & Medicine* 2006; 62: 2091-2100.
9. Goldstein DJ, Lu Y, Detke MJ, Wiltse C, Mallinckrodt C, Demitrack MA. Duloxetine in the treatment of depression - A double-blind, placebo-controlled comparison with paroxetine. *J Clin Psychopharmacol* 2004; 24: 389-399.
10. Hamilton M. Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology* 1967; 6: 278-296.
11. Heitjan DF, Moskowitz AJ, Whang W. Problems with interval estimates of the incremental cost-effectiveness ratio. *Medical Decision Making* 1999; 19: 9-15.
12. Heitjan DF, Moskowitz AJ, Whang W. Bayesian estimation of cost-effectiveness ratios from clinical trials. *Health Economics* 1999; 8: 191-201.
13. Johnson FR. Einstein on willingness to pay per QALY: is there a better way? [editorial.] *Medical Decision Making* 2005; 25: 607-608.
14. Kent DM, Fendrick AM, Langa KM. New and dis-improved: on the evaluation and use of less effective, less expensive medical interventions. *Medical Decision Making* 2004; 24: 281-286.
15. Laska EM, Meisner M, Siegel C, Stinnett AA. Ratio-based and net benefit-based approaches to health care resource allocation: proofs of optimality and equivalence. *Health Economics* 1999; 8: 171-174.
16. Laupacis A, Feeny D, Detsky AS, Tugwell PX. How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. *Can Med Assoc J* 1992; 146 (4): 473-81.
17. O'Brien BJ, Gertsen K, Willan AR, Faulkner LA. Is there a kink in consumers' threshold value for cost-effectiveness in health care? *Health Economics* 2002; 11(2): 175-80.
18. O'Brien BJ, Sculpher MJ. Building uncertainty into cost-effectiveness rankings: portfolio risk-return tradeoffs and implications for decision rules. *Medical Care* 2000; 38: 460-468.
19. Obenchain RL. Issues and algorithms in cost-effectiveness inference. *Biopharmaceutical Reports* 1997; 5(2): 1-7. Washington, DC: American Statistical Association.

20. Obenchain RL. ICEplane: a Microsoft Windows application for calculation and graphical display of bootstrap ICE confidence and tolerance regions. Copyright © Pharmaceutical Research and Manufacturers of America (PhRMA.) Version 2005.11. Downloads at [www.math.iupui.edu/~indyasa](http://www.math.iupui.edu/~indyasa) 1997-2005.
21. Obenchain RL. Resampling and multiplicity in cost-effectiveness inference. *J Biopharm Stat* 1999; 9(4): 563-582.
22. Obenchain RL. The key role of “symmetry” in cost-effectiveness analyses. *Proc. Biopharmaceutical Section, Joint Statistical Meetings*. Washington, DC: American Statistical Association. 2000, pages 15-17.
23. Obenchain RL. Incremental Cost-Effectiveness (ICE) Preference Maps. *Proc. Biopharmaceutical Section*, Washington, DC: American Statistical Association. 2001; CD-ROM only: 10 pages.
24. Obenchain RL. ICE Preference Maps: Nonlinear Generalizations of Net Benefit cost-effectiveness inferences from bootstrap quadrant confidence levels: three degrees of dominance. *J Biopharm Stat* 2005; 15(3): 419-436.
25. Obenchain RL, Robinson RL, Swindle RW. Cost-effectiveness inferences from bootstrap quadrant confidence levels: three degrees of dominance. *J Biopharm Stat* 2005; 15(3): 419-436.
26. Stinnett AA. Is it really so bad to be unambiguously inefficient? The role of dominance in stochastic cost-effectiveness analysis [editorial.] *Medical Decision Making* 1999; 19: 102-103.
27. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Medical Decision Making*, Special Issue on Pharmacoeconomics 1998; 18: S68-S80.
28. The R Project for Statistical Computing. <http://www.r-project.org>
29. Van Hout BA, Al MJ, Gordon GS, Rutten FFH. Costs, effects and C/E ratios alongside a clinical trial. *Health Economics* 1994; 3: 309-319.
30. Willan AR, O’Brien BJ, Leyva RA. Cost-effectiveness analysis when the WTA is greater than the WTP. *Statistics in Medicine* 2001; 20(21): 3251–3259.
31. Wolfram Research, Inc., Mathematica®, Version 4.1, Champaign, IL, 2000.

# Tables

	<b>Title</b>	<b>Page</b>
<b>1</b>	<b>Four Axioms of ICE Preference</b>	<b>6</b>
<b>2</b>	<b>Some Willingness values when <math>\eta \neq 1</math>.</b>	<b>8</b>
<b>3</b>	<b>VAGR and ALICE definitions by Quadrant</b>	<b>18</b>

# Figures

	<b>Topic</b>	<b>Page</b>
<b>1</b>	<b>ICE Economic Preference Maps in “Canonical” Form (<math>\lambda = 1</math>.)</b>	<b>9</b>
<b>2</b>	<b>Generalized Linear (<math>\eta=1</math>) Map, Diminishing Returns (<math>\beta=0.5</math>)</b>	<b>11</b>
<b>3</b>	<b>“Wrong” <math>\lambda \rightarrow</math> “No Symmetry” in Highly Nonlinear Maps</b>	<b>11</b>
<b>4</b>	<b>Laupacis “Pie Chart” graphic with Colored “Wedges”</b>	<b>12</b>
<b>5</b>	<b>Graphic with <math>0 &lt; s &lt; w &lt; 1 &lt; 1/w &lt; 1/s</math> because <math>\eta = \gamma/\beta &gt; 1</math></b>	<b>13</b>
<b>6</b>	<b>Alias and Alibi displays of plot(ICEuncrt) for 2 values of <math>\lambda</math></b>	<b>15</b>
<b>7</b>	<b>Alias plot(ICEwedge) region for “dulxparx” at 2 different <math>\lambda</math>s</b>	<b>17</b>
<b>8</b>	<b>VAGR Acceptability and ALICE Curves for “dulxparx”</b>	<b>21</b>
<b>9</b>	<b>VAGR Acceptability and ALICE Curves for “flouxtca”</b>	<b>22</b>
<b>10</b>	<b>plot(ICEcolor) with Linear Net Benefit for “dulxparx”</b>	<b>24</b>
<b>11</b>	<b>plot(ICEcolor) with Nonlinear Preferences for “dulxparx”</b>	<b>25</b>
<b>12</b>	<b>plot(ICEcolor) with Linear NB for “fluoxpin” and “fluoxtca”</b>	<b>26</b>