Are Recreation Values Systematically Underestimated? Reducing Publication Selection Bias for Benefit Transfer

T. D. Stanley
Department of Economics
Hendrix College
Conway, AR, USA
Stanley@Hendrix.edu

Randall S. Rosenberger
Department of Forest Ecosystems and Society
Oregon State University
321 Richardson Hall
Corvallis, OR 97333-5703
541-737-4425
R.Rosenberger@oregonstate.edu

Abstract: This paper shows how reported recreation values can be systematically underestimated when they are derived from price coefficients. Simulations show that these publication selection biases can be very large. Because selected price coefficients are transformed into estimates of value, conventional meta-analytic methods used to deal with publication selection will often make the bias worse. Simulations also show that an alternative, meta-regression estimator, Root-n MRA, can greatly reduce this potential bias and has lower MSE than alternative meta-analytic methods. This method uses the square root of a study's sample size as a proxy for the standard error of welfare measures, thus avoiding simultaneity bias associated with welfare measures and standard errors of price coefficients. Methods for detecting publication selection are illustrated by applying alternative estimators to the outdoor recreation valuation literature, in general, and freshwater fishing, in particular.

Keywords: Benefit transfer, Meta-analysis, Non-market values, Publication selection bias, Recreation values, Simulations

Acknowledgments:
We thank Klaus Moeltner and John Loomis for their helpful comments on an earlier draft of this paper. This research was supported, in part, by funds from the U.S. EPA STAR Grant #RD-832-421-01 to Oregon State University. Any errors are sole responsibility of the authors.

Are Recreation Values Systematically Underestimated? Reducing Publication Selection Bias for Benefit Transfer

1. Introduction

There is great interest in the estimation of environmental values and in applying estimated values from studied sites to sites for which there are no data (Bateman and Jones 2003; Bergstrom and Taylor 2006; Iovanna and Griffiths 2006; Johnston et al. 2005, 2006; Nelson and Kennedy 2009; Rosenberger and Stanley 2006; Smith and Pattanayak 2002). The benefits of such a transfer are obvious. Yet, benefit transfer often entails large errors and substantial biases (Rosenberger and Stanley 2006). In recent years, meta-regression analysis (MRA) is gaining acceptance as one method to generate estimates of value suitable for benefit transfer (Bateman and Jones 2003; Bergstrom and Taylor 2006 Johnston et al. 2005; Moeltner et al. 2007; Moeltner and Rosenberger 2008; Rosenberger and Phipps 2007; Rosenberger and Loomis 2003), although not without questions about its validity (Nelson and Kennedy 2009; Poe et al. 2001; Smith and Pattanayak 2002; US EPA 2007). Nonetheless, MRA, as well as all other benefit transfer approaches to environmental valuation, are vulnerable to publication selection bias (Nelson and Kennedy 2009; Rosenberger and Johnston 2009; Rosenberger and Stanley 2006).

It is widely recognized in economics, medical research and the social sciences that researchers, reviewers and editors have a preference for statistically significant estimates. Statistical significance is often regarded as a criterion for scientific meaningfulness. Insignificant empirical estimates are thought to contain no information and are thus not considered sufficiently important to be published or even reported. As Leamer and Leonard (1983) recognized a quarter century ago,

> Empirical results reported in economic journals are selected from a large set of estimated models. Journals, through their editorial policies, engage in some selection, which in turn stimulates extensive model searching and prescreening by prospective authors. Since this process is well known to professional readers, the reported results are widely regarded to overstate the precision of the estimates, and probably to distort them as well. As a consequence, statistical analyses are either greatly discounted or completely ignored. (p. 306)

This tendency to report only statistically significant results is greatly bolstered when there is also professional consensus regarding the existence and direction of an effect—such as the 'Law' of demand. When environmental survey data are used to estimate the price coefficient of a demand relation, the first estimated coefficient produced will not necessarily be the one that is reported. Rather, analysts will wish to be sure that the estimated demand relation is 'valid.' Validity will require, at a minimum, that the price coefficient be negative and in many cases that it be statistically significant as well.[1] Thus, the sample of reported estimates of environmental values may not be random, and, if not, any summary of values will be biased. "Publication bias (aka 'file-drawer problem') is a form of sample selection bias that arises if primary studies with statistically weak, insignificant, or unusual results tend not to be submitted for publication or are less likely to be published" (Nelson and Kennedy 2009, p. 347).

In contrast to other economic applications, the publication bias in environmental non-market valuation is likely to cause reported values to be *underestimated*, often by a lot.[2] Nelson and Kennedy (2009) argue that 'best practice' in environmental meta-analyses is to "(a)ssess the quality of the final results in light of formal tests for publication bias. . ., including funnel plots and more advanced tests such as the funnel asymmetry test (or FAT)" (p. 372). Unfortunately, these meta-analytic methods, found useful in other areas of economic research, will likely cause the publication bias of non-market environmental valuations to be much larger—see Section 5.3 below.

The purpose of this paper is to:

1. Document how publication selection bias may seriously underestimate recreation values and how several conventional meta-analytic methods can make this bias worse;
2. Offer meta-analytic methods that can reduce publication selection bias;
3. Validate and compare alternative estimators through Monte Carlo simulations; and
4. Apply these methods to all reported values of freshwater fishing and thereby to estimate its consumer surplus, corrected for publication selection bias.

The principal objective of this paper is to offer meta-analytic methods that minimize publication selection biases routinely found among recreation valuations. Our main finding is that all methods, whether conventional or meta-analytic, are vulnerable to potentially large publication selection bias. Nonetheless, we offer novel MRA methods that can be shown to reduce potential publication selection bias.

2. Publication Selection and Meta-Regression Analysis

Publication selection bias results from a literature of reported estimates that are not an unbiased sample of the actual empirical evidence. Researchers and reviewers often have a preference for statistically significant results or for results that conform to prior theoretical expectations, or both. Publication selection has long been recognized as an important problem in economics (e.g. Card and Krueger 1995; DeLong and Lang 1992; Feige 1975; Leamer 1983; Leamer and Leonard 1983; Lovell 1983; Roberts and Stanley 2005; Tullock 1959, to cite but a few). This problem is receiving heightened scrutiny in medical research as a consequence of the Vioxx and Paxil scandals.[3] A recent systematic review of publication bias of clinical medical trials found that "trials with positive findings . . . had nearly four times the odds of being published compared to findings that were not statistically significant" (Hopewell et al. 2009, p. Summary). To counter the intentional suppression of unfavorable research findings, leading medical journals now require the prior registration of clinical trials as a condition for the later publication of any of their findings (Fairman and Curtiss 2009; Krakovsky 2004).

In economics, it is unlikely that the consequences of publication selection will be so dire; although mistaken economic research can have lasting effects on policy. At one level, publication selection can be benign, even appropriate. Many editors and researchers regard statistical significance as a necessary condition for scientific importance. Regardless, publication selection can bias the entire distribution of reported empirical values along with any summary of these values. For example, our simulations show that the simple average may *underestimate* the actual recreation value many fold.

Worse still, conventional meta-analytic summary statistics, fixed-effects and random-effects weighted averages, which have been recently recommended by Nelson and Kennedy (2009) for routine use in environmental applications, are likely to make this selection bias much larger (see Table 1). Any project or policy that hinges on an environmental value estimate may, therefore, be grossly misinformed.

Wide application of meta-regression analysis in economics suggests that publication biases are often as large as or larger than the underlying parameter being estimated (Doucouliagos and Stanley 2009; Krassoi Peach and Stanley 2009; Stanley 2005a, 2008). For example, price elasticities of water demand have been found to be exaggerated four-fold through publication selection bias (Dalhuisen et al. 2003; Stanley 2005a). The negative sign of own-price elasticity is often required to validate the researcher's estimated demand relation. Should a positive coefficient be produced, researchers feel obligated to re-specify the demand relation, find a different econometric estimation technique, identify and omit outliers, or somehow expand the dataset. As a result of such publication selection, reported price elasticities of water demand are exaggerated by a factor of nearly four (Stanley 2005a). Needless to say, the water board of a drought-stricken area will be greatly disappointed to find that a doubling of residential water rates reduces usage by a mere 10% and not the expected 40%. Because the 'Law' of demand is so widely accepted, demand studies will ironically exhibit the greatest publication bias. Recreation valuation is often based on an estimated demand relation. As a result, recreation values may contain large publication biases.

We need some method to mitigate publication selection if benefit transfer is to be reliable. Over the past decade, meta-analysis has become routinely employed to identify and correct publication selection (Ashenfelter et al. 1999; Card and Krueger 1995; Coric and Pugh 2008; Doucouliagos 2005; Doucouliagos and Stanley 2009; Egger et al. 1997; Gemmill et al. 2007; Görg and Strobl 2001; Knell and Stix 2005; Krassoi Peach and Stanley 2009; Longhi et al. 2005; Mookerjee 2006; Roberts and Stanley 2005; Rose and Stanley 2005; Stanley 2005a, b, 2008). It has become standard practice to include the standard errors (or their inverse, precision) in a meta-regression analysis (MRA) to identify and correct for publication selection bias

$$effect_i = \beta_0 + \alpha_0 SE_i + \mathbf{Z_i}\gamma + \varepsilon_i \tag{1}$$

(Card and Krueger 1995; Doucouliagos 2005; Doucouliagos and Stanley 2009; Egger et al. 1997; Gemmill et al. 2007; Rose and Stanley 2005; Stanley 2005a, 2008). Where ε_i is a random error, $\mathbf{Z_i}$ is a matrix of moderator variables that reflect key dimensions in the variation of the 'true' empirical effect (heterogeneity) or identify large-sample biases that arise from model misspecification, and SE_i are the reported standard errors of the estimated effects.

Simulations have shown that meta-regression model (1) provides a valid test for publication bias (H_0:α_0=0), called 'funnel-asymmetry test' (FAT), and a powerful test of genuine empirical effect beyond publication selection (H_0:β_0=0), called a 'precision-effect test' or PET) (Stanley 2008).[4] The reason why this approach works is that the standard error serves as a proxy for the amount of selection required. Studies that have large standard errors are at a disadvantage in finding statistically significant effect sizes. Effect sizes need to be proportionally larger than their standard errors, because statistical significance is typically determined by a calculated t-value where the standard error is in the denominator. Such imprecise estimates will likely require further re-estimation, model specification, and/or data adjustments to become statistically significant. Thus, we expect to see greater publication selection in estimates with larger *SE*, *ceteris paribus*. This correlation between reported effects and their standard errors has been observed in dozens of different areas of economics research.

However, for MRA model (1) to correctly identify publication bias, estimates must be symmetrically distributed and independent of their standard errors in the absence of publication selection. Unfortunately, environmental values derived from price coefficients will not be independent of their standard errors, even in the absence of publication selection. Therefore the FAT-PET-MRA, equation (1), is not applicable.

3. Why Conventional Meta-Analytic Methods to Correct Publication Bias Fail

Typically recreation values are transformed from estimated demand coefficients. Linear and semilog relations are the most commonly employed demand models; where the latter uses the logarithm of quantity. This is especially true of outdoor recreation valuation literature, where we have identified 463 estimates from linear and 584 estimates from semilog demand functions.[5] When a linear model of daily demand is estimated,

$$CS = -.5\overline{Q} / \hat{\beta} \tag{2}$$

(Adamowicz et al. 1989, p. 416). Where CS is the estimated consumer surplus, $\overline{Q}$ is the average daily quantity demanded, and $\hat{\beta}$ is the estimated price coefficient. When the semilog model is used, consumer surplus simplifies to:

$$CS = -1/\hat{\beta} \tag{3}$$

(Adamowicz et al. 1989, p. 416).[6] It is this transformation of a statistical estimate, $\hat{\beta}$, that causes consumer surplus and its standard error to be simultaneously determined, thereby invalidating MRA model (1).[7]

Note how these formulas force researchers to report only negative values for $\hat{\beta}$; otherwise, consumer surplus is undefined or, worse, a meaningless negative value. Because the purpose of these recreation studies is to estimate value, this sign constraint will always be strictly binding.[8] It is the combination of the above consumer surplus transformations and this sign constraint that gives reported recreation values their unique distribution.

To further explore this distribution, we need to understand the relationship of consumer surplus and its standard error. Because *CS* is a derived estimate from (2) or (3), its standard error is proscribed by this transformation. To calculate the standard error, the delta method is generally used.[9]

$$S_i = - CS_i \cdot S_\beta \,/\, \hat{\beta} = - CS_i/t_i \;=\; CS_i^2\, S_\beta \tag{4}$$

Where S_i is the standard error of a consumer surplus estimate, S_β is the standard error of $\hat{\beta}$, and t_i is the t-value of $\hat{\beta}$ (Davidson and MacKinnon 2004, p.205).[10] Note how the consumer surplus and its standard error are dependent upon one another. Any factor that might influence the estimated price coefficient, including sampling error, will simultaneously affect both CS_i and its standard error, S_i. Simple MRA models of publication selection and its correction are invalid, because they assume that an estimate and its standard error are independent in the absence of publication selection (Stanley 2005a, 2008).[11]

Worse still, larger standard errors will no longer signal greater publication selection. Because value is a function of the inverse of the estimated price coefficient, which is likely to be selected for statistical significance, selected consumer surplus will often have a *smaller* standard error than those not reported. To see this, recall equations (2), (3) and (4). Imprecisely estimated demand relations (*i.e.*, those with small samples) will need to report larger price coefficients (in magnitude) in order to be statistically significant. Such selected price coefficients will produce *smaller* CS_i and S_i values. This follows directly from equations (2) and (3). Furthermore, the standard error of CS_i will also tend to be smaller because consumer surplus is smaller—equation (4). Again, it is this transformation of price coefficients to values that obscures the usual trace of publication selection that can be found in the relationship between standard errors and the associated reported estimates. Rather, this transformation imparts its own unique distributions and relationships upon CS_i and S_i, regardless of publication selection. As a result, methods designed to mitigate publication selection bias found useful in other areas of research will only make recreation value estimates more biased.

[INSERT FIGURE 1 HERE]

To see how this transformation of price coefficients distorts estimates of value, consider the funnel graph, Figure 1. A funnel graph plots precision, the inverse of the standard error (1/S), against estimates of empirical effect (consumer surplus). Funnel graphs are widely used in economics and medical research to identify publication selection or its absence (Doucouliagos and Stanley 2009; Stanley 2005a; Sutton et al. 2000). In the absence of publication selection, the plot should resemble an inverted funnel, peaked at the top and symmetric (see Figure 2). Here, the transformation of price coefficients to value and the selection of negative, but not necessarily statistically significant, price coefficients generates extreme asymmetry unlike anything seen in other areas of economics research—see Figure 1. As a result, conventional meta-analytic methods will not be valid for recreation valuation.

[INSERT FIGURE 2 HERE]

4. Using the Square Root of the Sample Size to Identify Publication Selection

As discussed above, the transformation of price coefficients to consumer surplus values causes consumer surplus and its standard error to be simultaneously determined, and S_i will be a poor indicator of the severity of publication selection. Thus, an obvious approach to circumvent this difficulty would be to substitute a measure of precision that is not tainted through this process of transformation. The inverse square root of the sample size is just such a proxy for the standard error of consumer surplus, $S_{i,}$ and it has been used elsewhere for this very reason (Stanley 2005a, b). Begg and Berlin (1988) show that publication bias is proportional to the inverse of the square root of sample size. "(T)heoretical models predict that publication bias is strongly and inversely related to sample size" (Berlin et al. 1989, p. 383).

The connection of a study's standard error and the square root of its sample size, $\sqrt{n_i}$ (or degrees of freedom), has been made previously by meta-analysts. For example, Card and Krueger (1995) use the inverse of the square root of the degrees of freedom as a proxy for the standard error and its associated statistical power in their meta-regression analysis of minimum-wage elasticities. Stanley (2005a) uses $\sqrt{n_i}$ as an instrumental variable for precision to correct for potential error-in-variables bias that arises when a sample estimate of $1/S_i$ is required as an independent variable in the WLS version of equation (1). And, Stanley (2005b) employs $\sqrt{n_i}$ as a proxy for the precision of statistical tests that possess no standard error. Here, the use of $1/\sqrt{n_i}$ allows us to avoid the potential simultaneity bias of employing S_i in MRA model (1) when CS_i is the dependent variable.

Ceteris paribus, larger samples lead to smaller standard errors of the price coefficient (S_β). With larger samples, demand can be estimated more precisely and price coefficients are more likely to be statistically significant with little or no extra specification searching. In fact, the standard error of the price coefficient can be shown to be proportional to the inverse of the square root of the sample size ($1/\sqrt{n_i}$), under the conventional statistical assumptions. The advantage of the square root of the sample size is that it cannot be affected by the transformation of demand to values. Nonetheless, its inverse should remain correlated to the severity of publication selectivity.

Substituting $1/\sqrt{n_i}$ for SE_i in MRA model (1) and consumer surplus estimates for $effect_i$ gives:

$$CS_i = \beta_0 + \alpha_0 (1/\sqrt{n_i}) + Z_i\gamma + \varepsilon_i \tag{5}$$

In the absence of publication selection, estimates of the price coefficient will be independent of the number of observations used in their estimation. Because consumer surplus is the inverse of the estimated price coefficient, it too will be independent of the sample size. However, with publication selection for statistical significance, this independence ends. Small-sample studies are less likely to be statistically significant on their own; thus, researchers will tend to engage in some additional specification searching to obtain the required significance. $1/\sqrt{n_i}$ will be correlated to the expected publication bias, should it be present.

As before (recall footnote 2), MRA model (5) will have heteroscedastic errors; thus some version of weighted least squares should always be used to estimate (5). Correcting for heteroscedasticity may be accomplished either by multiplying equation (5) through by $\sqrt{n_i}$ or by using some pre-programmed WLS routine with n_i as the weights. In the next section, we report Monte Carlo simulations of this 'Root-n' MRA, equation (5), and compare its properties to alternative estimates of recreation value.

5. Simulations

5.1 Simulation design

Our simulations are grounded firmly on the observed distribution of outdoor recreation values. Observed statistics from the actual distribution of reported nonmarket values of freshwater fishing define key parameters of the data generating process upon which these simulations are built.[12] For the purposes of these simulations, we assume that value estimates are derived from price coefficients through transformations such as those found in equations (2) and (3). Publication selection occurs at two levels. First, consumer surplus must be positive to be meaningful, which implies that the price coefficient must be negative.[13] Secondly, there will be selection for statistical significance; e.g., 93% of the reported price coefficients in the outdoor recreation literature are statistically significant. Different combinations of these two publication selection mechanisms are explicitly simulated.

The basic structure of our meta-analysis simulations may be roughly sketched as:

1. Generate the price-quantity data randomly.[14]
2. Use OLS to estimate the price coefficient, β, and to test H_0:β=0. Generate various mixes of negative and significantly negative price coefficients.
3. Use these selected price coefficients and their standard errors to generate estimates of consumer surplus and their standard errors.
4. Simulate meta-regression model (5) and several alternative estimators of consumer surplus by repeating the previous steps either 80 or 240 times. At

this stage, meta-regression model (5) and alternative estimators are calculated to provide summary estimates of consumer surplus.

5. Repeat all of the above steps 10,000 times while tracking each estimator's expected value (hence, bias) and MSE (mean squared error).

The first step, above, defines the data-generating process. The price variable (P_i) for each study is simulated by a random normal variable (μ=20, σ=5). The logarithm of quantity, Y_i is then generated from:

$$Y_i = 5.5 + \beta P_i + \beta_2 X_{2i} + e_i \qquad\qquad i=1, 2, \ldots \text{n} \qquad\qquad (6)$$

e_i ~NID(0,.5). The 'true' price coefficient, β, is assumed to be either -.02 or -.01.[15] In the valuation of freshwater fishing the average reported price coefficient is -.024.[16] We simulate only the semilog form of the demand relation; thus, Y_i is the natural logarithm of quantity. Given the parameters of the generating distributions, equation (6) implies that the average R^2 of this demand relation is approximately 11% and the correlation coefficient between price and the logarithm of quantity is about 0.33. The $\beta_2 X_{2i}$ term induces misspecification bias, in general, and omitted-variable bias, in particular.

As is widely known, omitting a relevant variable from a regression model causes the estimate of a regression coefficient to be biased and inconsistent. Because this bias remains in large samples, it can be mistaken for a genuine effect, potentially causing problems for all summary estimators. Although only omitted-variable biases are simulated here, they serve to represent any type of large-sample, misspecification bias (or inconsistency). Random misspecification bias is induced by making β_2 in equation (6) a random normal variable, $N(0, \sigma_{bias})$, for each study.

This random misspecification bias acts as 'heterogeneity,' which is widely recognized as a key parameter by meta-analysts (Hedges and Vevea 1996; Moreno et al. 2009; Stanley 2008). The most influential magnitude for the performance of the meta-regression methods is the size of the typical misspecification bias, σ_{bias}, relative to the sampling error, σ_{e_i} (Stanley 2005b, 2008). An accepted way to measure the magnitude of this unexplained heterogeneity is I^2, which measures the percent of the total observed variation due to unexplained heterogeneity (*i.e.*, not attributable to random sampling error) (Higgins and Thompson 2002).[17] For freshwater fishing, I^2= 92%. As discussed more fully below, this means that unexplained heterogeneity (or omitted variable bias) routinely swamps sampling error and any underlying price signal, β. In other applications, Stanley (2005b, 2008) and Moreno et al. (2009) find that heterogeneity in excess of 50% renders all known methods to deal with publication selection ineffective and unreliable. Unfortunately, recreation valuation is likely to have such excessive heterogeneity.

Rather than using any arbitrary values of these key parameters, we adjusted the standard deviations of P_i, of β_2, and e_i to approximate the key dimensions found among 752 reported estimates of the value of freshwater fishing and their standard errors. In particular, we force our simulations to have approximately the same relative heterogeneity and coefficient of variation as those observed among all reported freshwater fishing values. The coefficient of variation is used to reflect the extreme skewness found in these distributions. Most importantly, we fine tune the standard deviations of β_2, and e_i until the observed unexplained heterogeneity is approximately the same as what is observed among freshwater fishing values (I^2=92%). Also, the observed ratio of standard deviation among reported consumer surplus values to the average reported standard error of these estimated consumer surplus values is forced to be the same in the simulations as that found in the literature, approximately 4 to 1. Using these key dimensions of freshwater fishing values as bench marks for our data generating process implies that: β = -.02, σ_{e_i}=.50 and σ_{bias}=.03.

Sample sizes chosen for the estimation of demand are {20, 100, 500, 800, 1000}. The meta-regression models are assumed to be estimated using either 80 or 240 studies. Eighty is chosen because there are four specific types of outdoor recreation activities that report a larger number of estimates (freshwater fishing, big-game hunting, waterfowl hunting, and wildlife viewing), while three of these areas have an excess of 240 estimates. MRA sample sizes have little effect on the bias of these alternative estimators of value. To save space, only the simulation results of MRA n=80 are reported in the text. The simulations for n=240 are reported in the Appendix.

Two distinct, yet overlapping, types of publication selection mechanisms are used by these simulations. First, only negative price coefficients are reported in the outdoor recreation literature. Given the small magnitude of the average reported price coefficient (-.024) and the large observed heterogeneity (92%), this could not happen without explicit selection. Even if there is indeed a genuine, but small, price effect, sampling error and the large potential variation due to specification choices would almost guarantee that some of the estimates produced would be positive and a few more would be negative but statistically insignificant. Thus, our simulations select all estimates to be either negative or significant or both. However, different ratios of these two types of selection are explicitly simulated. We begin by assuming that 75% of the reported estimates are selected to be statistically significant (recall that 93% of the outdoor recreation values are). This leaves the remaining 25% to be selected as negative, thereby ensuring that consumer surplus is positive. Three other mixes of these selection mechanisms are simulated (87.5%, 94% and 100%), including the possibility that all price coefficients are selected to be significantly negative (100%). In all cases, price-quantity data are generated according to the above data generating processes, and the demand relation (6) is estimated by conventional regression analysis (OLS). If the resulting price coefficient does not have the desired property, then entirely new random data are generated and the selection processes continue until a suitable estimate is randomly produced and then selected.[18]

5.2 Alternative Summary Estimators of Environmental Value

Alternative conventional meta-analytic methods are simulated and compared to our Root-n MRA, equation (5). First is the simple mean of the reported consumer surplus (CS) values, 'Mean.' Next, two widely employed weighted averages, fixed- and random-effects, are also simulated and compared to Root-n MRA, because these have been recommended for routine use in environmental applications (Nelson and Kennedy 2009). The fixed-effects estimate (FE) is a weighted average of CS; where the weights are the inverse of each estimate's variance, $1/S_i^2$. Similarly, the random-effects (RE) estimate is a weighted average of CS, but here the weights are the inverse of a more complex estimate of each variance.[19] For RE, this variance has two components: an estimate of the heterogeneity variance across studies and sampling error within a study.[20]

We also offer a modified fixed-effect estimate that uses the sample size, $n_{i,}$ as a proxy for the inverse of each estimate's variance and therefore as the weight. The motivation for employing this 'weighted by n' average, or Wn, is the exact same as that used to derive Root-n. As discussed above, the transformation of price coefficients to values distorts both the estimated values and their standard errors. Because sample size is unaffected by this transformation to consumer surplus, it serves as an alternative proxy for the inverse of each estimate's variance.

Lastly, we report the WLS (weighted-least squares) estimates of $\hat{\beta}_0$ from equation (1). This coefficient has been recognized to correct for publication selection (Moreno et al. 2009; Stanley 2008; Sutton et al. 2000). Although this precision-effect estimator, $\hat{\beta}_0$, has been shown to be useful in other contexts (Doucouliagos and Stanley 2009; Krassoi Peach and Stanley 2009), it is extremely biased in recreation valuation applications (see below). Our only purpose in reporting several of these estimators ($\hat{\beta}_0$, FE and RE) is to document that they should not be used in environmental valuation applications.

5.3 Simulation Results

[INSERT TABLE 1 HERE]

Table 1 presents the average values of six alternative estimates of consumer surplus from 10,000 replications when each meta-analysis has access to 80 estimates of nonmarket value, its standard error and the sample size used to make this estimate. For the sake of robustness, key parameters have been both doubled and halved from their benchmark values dictated by freshwater fishing valuation. The most obvious lesson that can be drawn from these simulations is that FE, RE, and the MRA coefficient, $\hat{\beta}_0$, are likely to be very biased, grossly underestimating recreation values. In spite of Nelson and Kennedy's (2009) recent recommendation, these conventional summary statistics are likely to underestimate environmental values, often by a lot. The precise magnitude of this bias depends on a complex interplay of the relative strength of the two selection mechanisms and the underpinning magnitudes of β, σ_{e_i}, and σ_{bias}.[21]

Turning to $\hat{\beta}_0$ first, its mean is but a small fraction, 15 to 40%, of the actual value ($50)—Table 1. Likewise, both FE and RE have large downward biases underestimating value by well over 50%. In all cases, these three estimators are dominated by other approaches, including the simple average, (*i.e.,* 'Mean'). The reason for these large biases is clear. All use some function of the inverse of S_i as precision and weight the individual estimates accordingly. As we discussed in Section 4, such weights are likely to be inversely related to the actual precision of estimation, distorted by the transformation of price coefficients to consumer surplus values. The wrong CS estimates are being given the largest weights. Because these biases tend to be quite large (see Tables 1 and 2), environmental economists would do well to avoid using these conventional meta-analytic methods.

Next, consider the simple average of consumer surplus estimates, 'Mean.' In the large majority of cases it too underestimates value. Across all simulations, this bias represents an underestimate of over 30%. The only case where the average overestimates value is when both σ_{e_i} and σ_{bias} are small. In this case, there is little need for selection for statistical significance, because the estimates are likely to be significant without selection. Also, when σ_{bias} is small, there will be less distortion from misspecification bias, which, due to selection, tends to inflate the price coefficient and thereby lower the CS estimate. Even when the mean overestimates values, it does so by a small amount, under 6%. In this single case where the mean overestimates value, skewness, as measured by the coefficient of variation, is less than the actual reported estimates of the recreational values of freshwater fishing. Thus, the overwhelming practical lesson of these simulations is that past benefit transfer studies that used a simple or weighted average are likely to have underestimated value due to publication selection. And, this underestimate may have been large enough to be of practical significance for policy.

The potential for large bias becomes yet more clear when one realizes what happens if the true CS was $100, rather than $50 as assumed in Table 1. Table 2 presents the average values of these six alternative estimators under all the same conditions as before. The only difference is that the true price coefficient (-.01) is half of what it is in Table 1. Here, selection for statistical significance needs to be considerably more intense, causing larger publication bias, on average. Large underestimates of CS are seen consistently across all estimators and simulation conditions (Table 2).

[INSERT TABLE 2 HERE]

Finally, we turn to the new n-estimators that are introduced in Section 4. First, note that Root-n MRA greatly reduces publication bias (Table 1). Its average bias is only -5.5%. Similarly, the weighted average that uses the sample size as the weight (Wn) nearly halves the downward bias of the simple average (-17.7%). These two 'n-estimators' can be biased in either direction. Like the mean, they tend to overestimate the actual value when selection for statistical significance plays a minor role and is dominated by selection for a negative price effect. These conditions are associated with small values of both sampling errors and heterogeneity. However, none of the conditions that produce this overestimation seen in these simulations match the observed values of freshwater fishing. The conditions in Table 1 that most closely resemble freshwater fishing values are: σ_{e_i}=.50 & σ_{bias}=.03, σ_{e_i}=.25 & σ_{bias}=.045, and σ_{e_i}=.50 & σ_{bias}=.045. In these three cases, both Root-n and Wn greatly reduce publication selection bias, although they too underestimate consumer surplus.

It may also be worthwhile to note the effect that alternative intensities of selection for statistical significance has on bias. As mentioned above, as selection for significance becomes more dominant (from 75% to 100%), all of these estimates of consumer surplus decrease, increasing publication bias. As we move down Table 1 or 2, more intense selection for significance requires larger negative price coefficients, which when transformed, produce smaller consumer surplus estimates. Nonetheless, the overall pattern of bias remains; it only intensifies with more intense selection for statistical significance.

When we look at Table 2, the bias of all methods greatly increases. As discussed above, this inflation of bias is due to the increased intensity of selection for statistical significance required to make up for the fact that the true price coefficient is only half of its previous value (-.01 vs. -.02). When selection for statistical significance is this dominant, all estimators average under 50% of the true value (CS=$100), with the exception of Root-n which averages $56. In all cases, both Wn and Root-n dominate the other summary estimators. Nonetheless, all methods arguably have unacceptably large biases. Focusing on those cases most closely resembling freshwater fishing is of little help if CS=$100. Here, only two cases produce statistics similar to the freshwater fishing distribution (σ_{e_i}=.25 & σ_{bias}=.045 and σ_{e_i}=.50 & σ_{bias}=.045). Although both n-estimators dominate the alternatives for these cases, both estimate consumer surplus to be much less than its actual value.

[INSERT TABLE 3 HERE]

Lastly, we would be remiss if we were to ignore the mean square error (MSE) of these alternative estimators (see Tables 3 and 4). Due to the large publication selection bias of all estimators, bias swamps conventional estimation error in most cases. As clearly seen in Tables 3 and 4, the n-estimators have lower MSE and are thereby more efficient in the great majority of cases and comparisons. The n-estimators are entirely dominant when CS=$100 (Table 4), but the simple mean sometimes outperforms these estimators when the true CS=$50 and heterogeneity are small. By all statistical criteria, we have reason to prefer our n-estimators, especially when the simulations closely resemble conditions found among freshwater fishing values.

[INSERT TABLE 4 HERE]

Simulation results are similar when the meta-analyst has n=240, rather than 80, value estimates (See Appendix Tables 1-4). If anything, having more observations improves the relative performance of our n-estimators.

6. Meta-Regression Analysis of Freshwater Fishing Valuation.

[INSERT FIGURE 3 HERE]

An extensive literature search and meta-analysis identifies 2,594 estimates of the value of various outdoor recreation activities.[22] See Rosenberger and Stanley (2007) for a more complete description of the methods, scope and descriptive statistics of this meta-analysis. To minimize heterogeneity, we confine our attention here largely to a single outdoor recreation activity—freshwater fishing.[23] Freshwater fishing is selected because it has the largest number of reported estimates of consumer surplus per person per day (n=752; 681 of which also report sample sizes). Figure 3 displays the highly skewed, wide distribution of reported consumer surplus values, from $.48 to $420.57, with a mean of $53.81. The standard deviation is $54.36, which makes the coefficient of variation over 100 percent.

Our major concern is whether publication selection is likely to cause large biases in this literature. If so, we seek some method that may reduce these biases. Thus, we turn to the 'n-estimators' introduced in the previous sections. The 'weighted by n' estimate (Wn) for freshwater fishing is $53.05, little different than the simple mean. Likewise, the root-n estimate is not much smaller than the others, $48.68, and within Root-n's sampling error (se = $3.38) (See Table 5 column 1). It is reassuring that the three lesser biased estimators give roughly the same consumer surplus estimate for freshwater fishing, approximately $50.

Fortuitously, there is no evidence of any downward bias among freshwater fishing values (t=1.74), and publication selection explains little of the variation found among reported CS values (R^2 = 0.4%). Recall that testing the coefficient on $1/\sqrt{n_i}$ serves as a test for the presence of publication selection (Stanley 2005a). This coefficient on $1/\sqrt{n_i}$ in Table 5 implies that a study which uses a mere 100 observations would bias consumer surplus upward by $12.37, while a study that has access to 10,000 surveys would have an upward bias of only $1.24. Because large biases in consumer surplus estimates are found only when they are underestimated, our Root-n MRA provides some reassurance that this corrected estimate does not contain a large bias.

[INSERT TABLE 5 HERE]

However, we are not so lucky when we consider the values across all outdoor recreation activities. Here, there is evidence of a modest negative bias (t=-.261; p<.01)—see Table 5 column 2. By this MRA estimate, a study that uses 100 observations would underestimate CS by \$14.84, while 10,000 observations imply an understatement by \$1.48. Needless to say, both of these corrected estimates of outdoor recreation value {\$48.74 & \$64.42} are statistically significant (t={14.5 & 31.1}; p<.0001).[24]

The relative positions of alternative estimators found in our simulations are also confirmed by this application to freshwater fishing. That is, the fixed-effects weighted average is \$22.02, and the random-effects is \$34.74. These are very similar to the simulations reported in Table 1, where FE and RE are much smaller than the simple average and the corrected averages. Likewise, as predicted by our simulations, the freshwater fishing estimate of $\hat{\beta}_0$ (\$14.55) is much lower than all of the other estimates of value. It is comforting to find that the relative values of the alternative estimators found in the simulations are seen among actual reported values.[25]

Unfortunately, we cannot entirely rule out the possibility that all estimators of recreation values have substantial downward bias. As shown in our simulations when the 'true' consumer surplus is \$100 (Table 2), all methods can produce large downward biases. If this worst case scenario were true, our Root-n MRA would, nonetheless, be the least biased. Secondly, when the selection biases are especially large and we have a large sample (recall that n=681 for freshwater fishing), the funnel-asymmetry test on the coefficient of $1/\sqrt{n_i}$ would likely show evidence of such selection, and it does not. Lastly, in those simulations that produce large downward biases for all estimators, the sample mean is much smaller than our n-estimators, which is not true for freshwater fishing. To insure their broad applicability, these n-estimators should be tested on other collections of environmental values that are derived from different methods and estimators.

7. Conclusion

This study reveals how the transformation of price coefficients to environmental values often distorts the observed distributions and can lead to large underestimates of value. The selection of statistically significant negative price coefficients tends to impart an upward bias (in magnitude) to reported price coefficients. However, consumer surplus depends on the inverse of the estimated price coefficient. Thus, any systematic overestimate of price coefficients translates into a downward bias for estimates of environmental values. Monte Carlo simulations, grounded upon the observed characteristics of freshwater fishing values, confirm this downward publication bias and finds that it can be quite large.

Meta-regression analysis (MRA) can help to fill the gap between theory and practice famously identified by Leamer (1983) and others. However, special care needs to be exercised when attempting to model and correct for publication selection in environmental valuation. Unfortunately, conventional meta-analytic methods only make the downward bias of the average environmental value much worse. The most important implication of our simulations is that conventional meta-analytic methods (fixed- and random-effects weighted averages, funnel graphs, trim-and-fill, and meta-regression models of publication bias that use the standard error) should not be employed for environmental values if values are calculated from estimated demand relations or other 'price' variables. In many cases, the downward bias of these conventional meta-analytic methods is very large.

To reduce the likely publication selection bias in consumer surplus estimates, we offer and validate alternative meta-analytic methods—'n-estimators.' In particular, Root-n meta-regression analysis (Root-n MRA) uses the square root of the sample size (n) as a proxy for the precision of the reported value. Unlike the standard error, the sample size cannot be affected by the transformation from price coefficient to consumer surplus and will not, as a result, be endogenously related to the estimated consumer surplus value. In the same spirit, we recommend an alternative weighted average ('weighted by n' or Wn) that uses the sample size as the weight. Simulations show clearly that in most cases these 'n-estimators' reduce publication bias and dominate the alternative estimators by the mean square error criterion. Nonetheless, these simulations also reveal that all estimators, including our n-estimators, can have large downward bias and grossly underestimate environmental values. The major limitation of the n-estimators is that the square root of the sample size is often a poor proxy for the incidence of publication selection. Nonetheless, simulations show that even this poor proxy greatly reduces publication selection bias.

In the absence of some reliable method to validate the overall estimate of environmental value, what might environmental economists do? First, they should be aware of the possibility that all estimators might contain large *downward* biases due to publication selection. Secondly, the field should demand that all estimates, whether significant or not, be publically reported.

To minimize publication selection bias, leading medical journals require the prior registration of clinical trials as a condition for the later publication of any of their findings (Fairman and Curtiss 2009; Krakovsky 2004). A similar effort to increase the accountability in valuation research may be warranted. Environmental economics journals and sponsors of environmental valuation studies might require full disclosure of data and model estimates (Loomis and Rosenberger 2006). As Rosenberger and Stanley (2006, p. 377) note, "an e-journal whose sole purpose is the accurate and complete recording of studies that estimate value" should be established.

Endnotes

[1] At least this is the case in the outdoor recreation valuation literature. An extensive literature review finds no positive price coefficient, and 93% of the reported price coefficients are statistically significant (Rosenberger and Stanley 2007).

[2] We do not wish to imply that selection for statistical significance and negative price coefficients are the only types of selection biases in environmental research. For example, Hoehn (2006) documents how research priority (*i.e.*, the selection of unique or highly valued sites) can cause the overvaluation of wetlands by as much as four-fold. See Rosenberger and Stanley (2006) and Rosenberger and Johnston (2009) for a broader discussion of alternative sources of bias in environmental valuation.

[3] As several court cases have alleged, Paxil has the unfortunate side effect of increased teen suicide, while Vioxx's side effects include the increased risks of heart disease. Life-threatening side effects were well known from clinical trials, but the reporting of these side effects was suppressed by the sponsors of the clinical trials (Foley 2007; Krakovsky 2004; Nesi 2008).

[4] Typically, the weighted least squares (WLS) version of (1) is estimated because the standard error of the $effect_i$ (SE_i) will vary greatly from estimate to estimate. WLS may be derived by dividing (1) by SE_i. As a result, the estimate's reported t-value becomes the dependent variable, while $1/SE_i$ (or precision) becomes an independent variable—hence the name, 'precision-effect test.'

[5] See Rosenberger and Stanley (2007) for details of the recent meta-analysis of outdoor recreation valuation that is used to provide empirical estimates and benchmarks throughout this paper.

[6] We divide the conventional formulas by $\overline{Q}$, because recreation values are expressed in terms of dollars per person per day.

[7] By extension, dichotomous choice models exhibit similar relationships among 'prices' and welfare measures (Hanemann 1984, 1989).

[8] Areas of research that are based upon the 'Law' of demand exhibit significantly greater publication bias than others areas of economics. Nonetheless, economists recognize exceptions to the law of demand (*e.g.*, Giffin goods or 'snob' goods). In other areas where demand is estimated, meta-analysts typically find a few reported positive price coefficients (for example, Dalhuisen et al. 2003; Stanley 2005a). But this is not the case for environmental valuation. In the 983 reported price coefficients that can be found in the outdoor

recreation literature, not a single one is positive. Thus, this constraint to report meaningful consumer surplus values seems to impose a more stringent sign constraint on the price coefficient for environmental services when used for non-market valuation.

[9] Other methods (for example bootstrapping) may also be used, but they will give approximately the same relation between CS and its standard error, S_i.

[10] From the delta method, $S_i = -k\, S_{\beta} \;/\; \hat{\beta}^2$; where k is defined by equation (2) or (3) as $CS = \; k \,/\, \hat{\beta}$.

Equation (4) then follows.

[11] This is not generally a stringent assumption and is routinely satisfied in most areas of empirical research. Research studies that report t-values or confidence intervals have already made the assumption that estimates are independent of their standard errors.

[12] Estimates for a single activity category (freshwater fishing, n=752) were extracted from the broader recreation use values database (n=2,705) in order to minimize heterogeneity across activity category while maintaining a large number of observations. Freshwater fishing provides nearly 28% (n=752 out of 2,705 estimates) of observations in the database out of 27 activity categories identified.

[13] Most (*i.e.*, 60%) of the large CS estimates (those over $500 per person per day) come from models whose authors claim are 'bad.' Hence, for our meta-analysis, we trimmed CS values >$500 on the belief that they are somehow in error or derived from a misspecified model. To be consistent, we also put an upper limit on reporting CS values in our simulations ($300). Simulations without any upper limit clearly show that a very small value of the price coefficient will occur occasionally when selection is for negative price coefficients and that this small price coefficient produces a very large estimate of CS. Rather than a 'bad' model, random but small price coefficients might explain why about 1.5% of the reported consumer surplus values for outdoor recreation exceed $500.

[14] To be more precise, prices are generated as a random normal variable, and quantities are then generated from a demand relation using these randomly generated prices. See below for the details. Author created GAUSS programs are used for all of the simulations. These are available upon request.

[15] As discussed in more detail below. The key parameters, σ_{e_i} and β, are varied to investigate robustness.

[16] This average value is likely to be biased upwards due to publication bias. Thus, we lower the magnitude a little to allow for some publication bias. We also use β=-.01 for the sake of robustness and because some outdoor activities (notably mountain biking and non-motorized boating) have substantially higher average consumer surplus values (hence smaller price coefficients).

[17] The easiest method to calculate I^2 is to run a simple MRA, forced through the origin, between a study's t-value and its precision, $1/S_i$ (Higgins and Thompson 2002, p. 1547). Then, $I^2 = (Se^2 - 1)/Se^2$; where Se^2 is the estimated error variance from this MRA.

[18] These simulations produce funnel graphs that resemble quite closely the graphs of the actual reported values of outdoor recreation.

[19] The terms 'fixed- and random-effects' are used differently in the meta-analysis literature than in econometrics. Here, we are referring to simple weighted averages of the reported estimates not to multivariate panel estimators, as these terms refer to in econometrics. Although these more sophisticated panel models are often employed in meta-analysis (Bateman and Jones 2003; Rosenberger and Loomis 2000), employing them here is beyond the scope of our current paper.

[20] See Sutton et al. (2000) and Cooper and Hedges (1994) for the details of the necessary calculations.

[21] For a given β, the bias, as a percent of the observed mean, can be predicted rather well by a regression on the observed values of the heterogeneity and sampling standard deviations.

[22] This is after we omit any estimate where either the author explicitly identifies as 'bad' or is over $500.

[23] Recall that unexplained heterogeneity is the principal driving force that increases publication bias.

[24] We do not wish to imply that the simple MRAs reported in Table 5 fully explain or adequately estimate outdoor recreation values. Undoubtedly, there is much systematic heterogeneity among these reported estimates of CS that needs to be captured by more complex multivariate MRAs, not the least of which is between different outdoor activities. We have estimated a multivariate root-n MRA for freshwater fishing that explains 61% of the observed variation among CS estimates, but space limitations do not permit a comprehensive analysis or complete discussion of the systematic heterogeneity that can be found among outdoor recreation activity values.

[25] The astute reader will note that the relative values of the better estimators (the simple mean, Wn and Root-n) are not exactly mirrored in the actual reported values of freshwater fishing. However, these three estimates are all within their respective sampling error from one another. Thus, their relative positions are not well defined for this application to freshwater fishing.

References

Adamowicz WL, Fletcher JJ, Graham-Tomasi T (1989) Functional forms and the statistical properties of welfare measures. Am J Agric Econ 71(2):414-421.

Ashenfelter O, Harmon C, Oosterbeek, H (1999) A review of estimates of the schooling/earnings relationship, with tests for publication bias. Labour Econ 6(4):453-470.

Bateman IJ, Jones AP (2003) Contrasting conventional with multi-level modeling approaches to meta-analysis: expectation consistency in U.K. woodland recreation values. Land Econ 79(2):235-258.

Begg CB, Berlin JA (1988) Publication bias: a problem in interpreting medical data. J Royal Stat Society (Series A) 151(3):419-445.

Bergstrom JC, Taylor LO (2006) Using meta-analysis for benefits transfer: theory and practice. Ecol Econ 60(2):351-360.

Berlin JA, Begg CB Louis TA (1989) An assessment of publication bias using a sample of published clinical trials. J Am Stat Assoc 84(406):381-392.

Card D, Krueger AB (1995) Time-series minimum-wage studies: a meta-analysis. Am Econ Rev 85(2):238-243.

Cooper H, Hedges LV (1994) Handbook of research synthesis. Sage, New York.

Coric B, Pugh G (2008) The effects of exchange rate variability on international trade: a meta-regression analysis. Appl Econ (in press). DOI 10.1080/00036840801964500.

Dalhuisen J, Florax RJGM, deGroot HLF, Nijkamp P (2003) Price and income elasticities of residential water demand: a meta-analysis. Land Econ 79(2):292-308.

Davidson R, MacKinnon JG (2004) Econometric theory and methods. Oxford University Press, Oxford.

De Long JB, Lang K (1992) Are all economic hypotheses false? J Polit Econ 100(6):1257-1272.

Doucouliagos C (2005) Publication bias in the economic freedom and economic growth literature. J Econ Surv 19(3):367-388.

Doucouliagos C, Laroche P (2003) What do unions do to productivity: a meta-analysis. Industrial Relations 42(4):650-691.

Doucouliagos C, Stanley TD (2009) Publication selection bias in minimum-wage research? A meta-regression analysis. British J Industrial Relations (in press).

Egger M, Davey Smith G, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. British Med J 315(7109):629-634.

Fairman KA, Curtiss FR (2009) What should be done about bias and misconduct in clinical trials? J Managed Care Pharmacy 15(2):154-160.

Feige EL (1975) The consequence of journal editorial policies and a suggestion for revision. J Polit Econ 83(6):1291-1295.

Foley S (2007) Merck faces renewed legal threat over Vioxx scandal. The Independent (July 20).

Gemmill MC, Costa-Font J, McGuire A (2007) In search of a corrected prescription drug elasticity estimate: a meta-regression approach. Health Econ 16(6):627-643.

Görg H, Strobl E (2001) Multinational companies and productivity: a meta-analysis. Econ J 111(475):723-739.

Hanemann WM (1984) Welfare evaluations in contingent valuation experiments with discrete responses. Am J Agric Econ 66(3):332-341.

Hanemann WM (1989) Welfare evaluations in contingent valuation experiments with discrete responses: reply. Am J Agric Econ 71(4):1057-1060.

Hedges LV, Vevea JL (1996) Estimating effect size under publication bias: small sample properties and robustness of a random effects selection model. J Educ Behav Stat 21(4):299-332.

Higgins JPT, Thompson SG (2002) Quantifying heterogeneity in meta-analysis. Stat Med 21(11):1539-1558.

Hoehn JP (2006) Methods to address selection effects in the meta regression and transfer of ecosystem values. Ecol Econ 60(2):389-398.

Hopewell S, Loudon K, Clarke MJ et al (2009) Publication bias in clinical trials due to statistical significance or direction of trial result. Cochrane Database of Systematic Reviews (1) DOI 10.1002/14651858.MR000006.pub3.

Iovanna R, Griffiths C (2006) Clean water, ecological benefits, and benefits transfer: a work in progress at the U.S. EPA. Ecol Econ 60(2):473-482.

Johnston RJ, Besedin EY, Ranson MH (2006) Characterizing the effects of valuation methodology in function-based benefits transfer. Ecol Econ 60(2):407-419.

Johnston RJ, Besedin EY, Iovanna R et al (2005) Systematic variation in willingness to pay for aquatic resource improvements and implications for benefit transfer: a meta-analysis. Can J Agric Econ 53(2-3):221-248.

Knell M, Stix H (2005) The income elasticity of money demand: a meta-analysis of empirical results. J Econ Surv 19(3):513-533.

Krakovsky M (2004) Register or perish. Sci Am 291:18-20.

Krassoi Peach E, Stanley TD (2009) Efficiency wages, productivity and simultaneity: a meta-regression analysis. J Labor Res (in press).

Leamer EE (1983) Let's take the con out of econometrics. Am Econ Rev 73(1):31-43.

Leamer EE, Leonard H (1983) Reporting the fragility of regression estimates. Rev Econ Stat 65(2):306-317.

Longhi S, Nijkamp P, Poot J (2005) A meta-analytic assessment of the effect of immigration on wages. J Econ Surv 19(3):451-477.

Lovell MC (1983) Data mining. Rev Econ Stat 65(1):1-12.

Moeltner K, Boyle KJ, Paterson RW (2007) Meta-analysis and benefit transfer for resource valuation—addressing classical challenges with Bayesian modeling. J Environ Econ Manage 53(2):250-269.

Moeltner K, Rosenberger RS (2008) Predicting resource policy outcomes via meta-regression: data space, model space, and the quest for 'optimal scope'. B.E. J Econ Analy Policy 8(1) DOI 10.2202/1935-1682.2028.

Mookerjee R (2006) A meta-analysis of the export growth hypothesis. Econ Letters 91(3):395-401.

Moreno SG, Sutton AJ, Ades A et al (2009) Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. BMC Med Res Method 9(2):DOI 10.1186/1471-2288-9-2.

Nelson JP, and Kennedy PE (2009) The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment. Environ Resour Econ 42(3):345-377.

Nesi T (2008) Poison pills: the untold story of the Vioxx drug scandal. St. Martin's Press, New York.

Poe GL, Boyle KJ, Bergstrom JC (2001) A preliminary meta analysis of contingent values for ground water quality revisited. In: Bergstrom JC, Boyle KJ, Poe GL (eds), The economic value of water quality. Edward Elgar Publishers, Northampton, pp137-62.

Roberts CJ, Stanley TD (eds) (2005) Meta-regression analysis: issues of publication bias in economics. Blackwell, Oxford.

Rose AK, Stanley TD (2005) A meta-analysis of the effect of common currencies on international trade. J Econ Surv 19(3):347-365.

Rosenberger RS, Johnston RJ (2009) Selection effects in meta-analysis and benefit transfer: avoiding unintended consequences. Land Econ (in press).

Rosenberger RS, Loomis JB (2000) Panel stratification in meta-analysis of economic studies: an investigation of its effects in the recreation valuation literature. J Agric Appl Econ 32(2):459-470.

Rosenberger RS, Loomis JB (2003) Benefit transfer. In: Champ, PA, Boyle KJ, Brown TC (eds) A primer on nonmarket valuation. Kluwer Academic Publishers, Dordrecht, pp445-482.

Rosenberger RS, Phipps TT (2007) Correspondence and convergence in benefit transfer accuracy: meta-analytic review of the literature. In: Navrud S, Ready R (eds) Environmental value transfer: issues and methods. Springer, Dordrecht, pp23-43.

Rosenberger RS, Stanley TD (2006) Measurement, generalization, and publication: sources of error in benefit transfers and their management. Ecol Econ 60(2):372-378.

Rosenberger RS, Stanley TD (2007) Publication effects in the recreation use value literature: a preliminary investigation. Paper presented at the American Agricultural Economics Association Annual Meeting, Portland, 29 July-1 August 2007. http://purl.umn.edu/9883.

Smith VK, Pattanayak SK (2002) Is meta-analysis a Noah's Ark for non-market valuation? Environ Resour Econ 22(1-2):271-296.

Stanley TD (2005a) Beyond publication bias. J Econ Surv 19(3):309-345.

Stanley TD (2005b) Integrating the empirical tests of the natural rate hypothesis: a meta-regression analysis. Kyklos 58(4):587-610.

Stanley TD (2008) Meta-regression methods for detecting and estimating empirical effect in the presence of publication selection. Oxford Bull Econ Stat 70(1):103-127.

Sutton AJ, Abrams KR, Jones DR et al (2000) Methods for meta-analysis in medical research. John Wiley and Sons, Chichester.

Tullock G (1959) Publication decisions and tests of significance – a comment. J Am Stat Assoc 54(287):593.

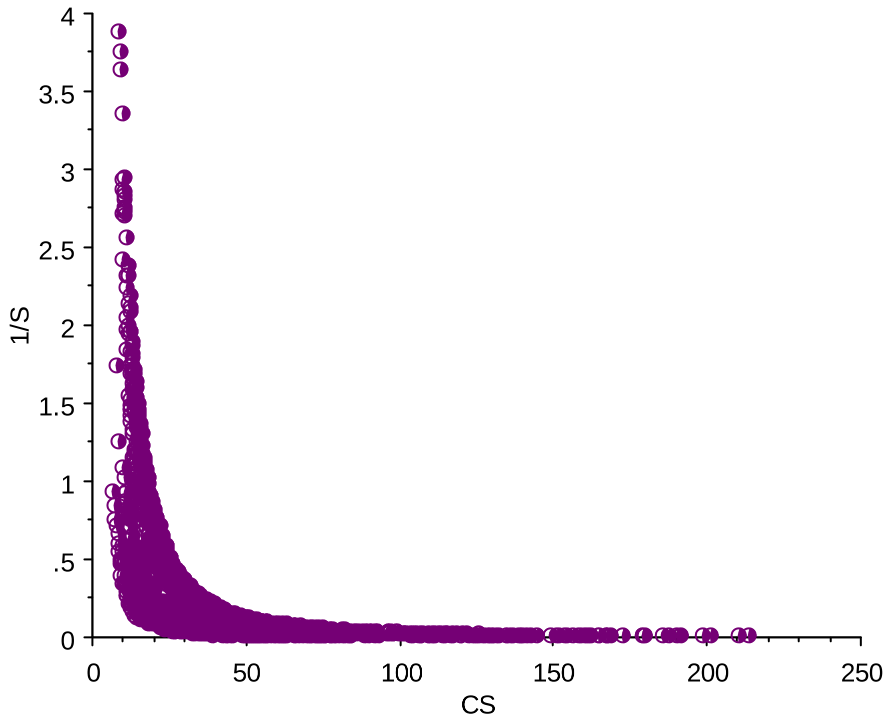
U.S. EPA (2007) Advisory on EPA's issues in valuing mortality risk reduction (EPA-SAB-08-001). U.S. EPA, Science Advisory Board, Office of the Administrator, Washington, DC.

Figure Captions

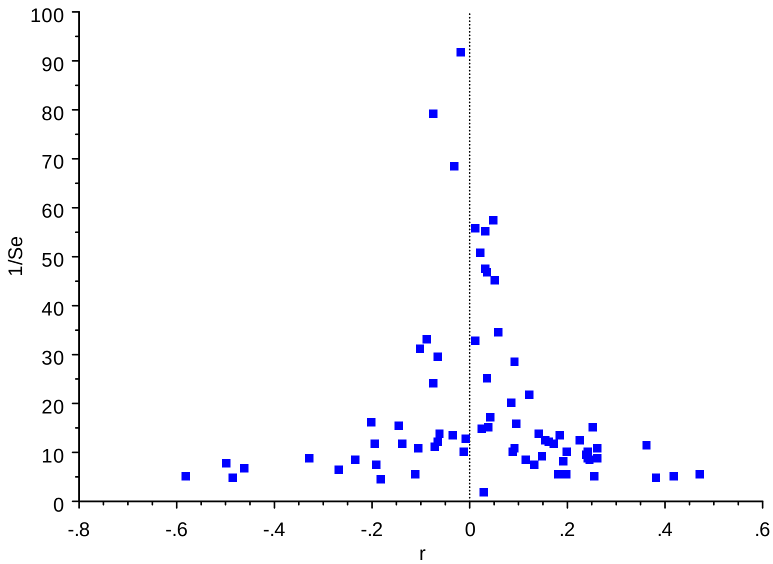
Figure 1: Simulated Values and their Precision, 1/S (No Selection for Significance) (See section 5 for simulation discussion)

Figure 2: Funnel Graph of Union-Productivity Partial Correlations (r) (Source: Doucouliagos and Laroche (2003))

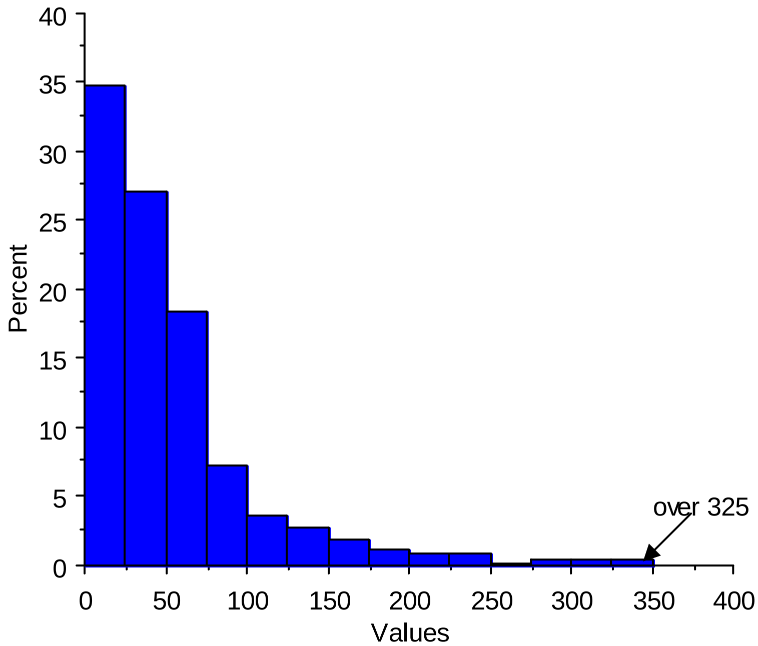
Figure 3: Frequency Distribution of Consumer Surplus Values: Freshwater Fishing



(Figure 1)



(Figure 2)


40
35
30
25
20
15
10
5
0
Percent
over 325
0
50
100
150
200
250
300
350
400
Values

(Figure 3)

Table 1: Means of Alternative Estimators of Value (True CS=$50 & n=80)

Hetero-geneity	σ_e	Selection Incidence	Wn	Mean	FE	RE	$\hat{\beta}_0$	Root-n
σ_{bias}=.015	.25	75%	59.27	52.90	25.83	37.53	20.29	65.09
	.25	87.5%	59.02	51.74	25.81	37.41	20.21	65.68
	.25	94%	59.05	51.21	25.80	37.38	20.17	66.23
	.25	100%	58.91	50.61	25.78	37.28	20.13	66.51
	.50	75%	52.00	44.49	24.20	30.72	18.00	59.02
	.50	87.5%	50.63	42.25	24.09	30.43	17.71	58.51
	.50	94%	49.91	41.08	24.03	30.29	17.55	58.21
	.50	100%	49.23	39.95	23.93	30.14	17.35	57.98
	1.0	75%	42.93	35.06	16.71	22.30	8.64	50.42
	1.0	87.5%	40.28	31.89	16.39	21.92	7.61	48.38
	1.0	94%	38.87	30.26	16.20	21.76	7.01	47.24
	1.0	100%	37.54	28.67	16.07	21.59	6.49	46.20
σ_{bias}=.03	.25	75%	48.39	43.10	15.67	27.10	12.49	53.32
	.25	87.5%	48.20	42.05	15.69	27.02	12.48	53.97
	.25	94%	48.09	41.57	15.67	27.02	12.46	54.19
	.25	100%	48.05	41.10	15.63	26.99	12.41	54.54
	.50	75%	41.81	36.27	15.30	22.80	11.88	47.03
	.50	87.5%	40.48	34.25	15.30	22.64	11.80	46.41
	.50	94%	39.88	33.27	15.26	22.60	11.73	46.17
	.50	100%	39.26	32.27	15.26	22.52	11.69	45.93
	1.0	75%	34.79	29.19	13.63	17.56	9.66	40.09
	1.0	87.5%	32.41	26.38	13.52	17.36	9.32	38.20
	1.0	94%	31.30	25.01	13.46	17.26	9.13	37.38
	1.0	100%	30.09	23.61	13.42	17.17	8.98	36.36
σ_{bias}=.045	.25	75%	40.06	35.83	11.24	21.24	9.02	44.04
	.25	87.5%	39.93	34.99	11.25	21.23	9.01	44.58
	.25	94%	39.84	34.60	11.25	21.21	9.01	44.74
	.25	100%	39.84	34.22	11.21	21.20	8.97	45.12
	.50	75%	34.68	30.40	10.97	18.13	8.67	38.74
	.50	87.5%	33.61	28.72	10.99	18.03	8.64	38.26
	.50	94%	33.07	27.88	10.98	18.00	8.62	38.02
	.50	100%	32.60	27.08	10.98	17.95	8.60	37.86
	1.0	75%	29.15	24.93	10.47	14.55	7.90	33.13
	1.0	87.5%	27.24	22.55	10.42	14.40	7.75	31.75
	1.0	94%	26.34	21.44	10.40	14.34	7.68	31.04
	1.0	100%	25.33	20.23	10.41	14.27	7.63	30.26

Table 2: Means of Alternative Estimators of Value (True CS=$100 & n=80)

Hetero-geneity	σ_e	Selection Incidence	Wn	Mean	FE	RE	$\hat{\beta}_0$	Root-n
σ_{bias}=.015	.25	75%	78.43	67.81	30.55	45.60	23.63	88.36
	.25	87.5%	78.19	66.05	30.60	45.33	23.56	89.58
	.25	94%	77.81	65.02	30.53	45.19	23.43	89.82
	.25	100%	77.67	64.12	30.55	45.05	23.40	90.41
	.50	75%	65.14	54.22	27.27	35.20	19.18	75.54
	.50	87.5%	62.66	50.74	27.07	34.76	18.59	74.10
	.50	94%	61.42	48.97	26.93	34.51	18.25	73.43
	.50	100%	60.14	47.21	26.84	34.33	17.95	72.66
	1.0	75%	50.90	40.70	16.53	23.80	7.23	60.77
	1.0	87.5%	46.48	36.07	16.17	23.34	5.97	56.68
	1.0	94%	44.37	33.85	16.02	23.14	5.35	54.78
	1.0	100%	42.22	31.60	15.85	22.96	4.62	52.77
σ_{bias}=.03	.25	75%	55.96	49.25	17.17	30.23	13.66	62.26
	.25	87.5%	55.90	48.00	17.14	30.17	13.59	63.36
	.25	94%	55.61	47.34	17.11	30.08	13.56	63.36
	.25	100%	55.53	46.70	17.13	30.00	13.56	63.87
	.50	75%	47.39	40.65	16.62	24.87	12.80	53.80
	.50	87.5%	45.79	38.18	16.61	24.66	12.70	53.09
	.50	94%	44.93	36.92	16.62	24.62	12.63	52.63
	.50	100%	44.07	35.65	16.62	24.53	12.57	52.18
	1.0	75%	38.65	32.11	14.36	18.64	9.83	44.88
	1.0	87.5%	35.57	28.66	14.20	18.39	9.38	42.25
	1.0	94%	34.10	26.96	14.11	18.26	9.13	41.03
	1.0	100%	32.66	25.31	14.05	18.17	8.86	39.85
σ_{bias}=.045	.25	75%	44.43	39.45	11.94	22.96	9.58	49.10
	.25	87.5%	44.28	38.45	11.94	22.88	9.56	49.77
	.25	94%	44.16	37.94	11.96	22.88	9.56	50.02
	.25	100%	44.03	37.43	11.96	22.85	9.56	50.26
	.50	75%	37.87	32.98	11.68	19.34	9.21	42.52
	.50	87.5%	36.70	31.08	11.67	19.25	9.14	42.09
	.50	94%	36.10	30.14	11.68	19.20	9.12	41.81
	.50	100%	35.38	29.10	11.65	19.16	9.07	41.41
	1.0	75%	31.49	26.76	10.99	15.24	8.22	36.00
	1.0	87.5%	29.29	24.08	10.99	15.11	8.06	34.31
	1.0	94%	28.17	22.73	10.93	15.04	7.96	33.43
	1.0	100%	26.96	21.34	10.92	14.94	7.87	32.40

Table 3: MSE of Alternative Estimators of Value (True CS=$50 & n=80)

Hetero-geneity	σ_e	Selection Incidence	Wn	Mean	FE	RE	$\hat{\beta}_0$	Root-n
σ_{bias}=.015	.25	75%	129.3	27.6	589.0	159.7	886.5	330.2
	.25	87.5%	123.5	20.1	589.9	162.6	890.9	346.3
	.25	94%	124.4	18.0	590.4	163.4	893.6	364.9
	.25	100%	122.4	16.5	591.5	166.0	895.8	374.5
	.50	75%	27.8	41.7	669.0	373.8	1026.4	138.7
	.50	87.5%	20.3	68.5	674.2	384.8	1045.1	120.0
	.50	94%	17.6	86.5	677.4	390.5	1055.5	109.9
	.50	100%	16.7	106.8	682.5	396.3	1068.5	102.7
	1.0	75%	65.1	231.4	1110.2	768.0	1714.3	36.5
	1.0	87.5%	104.2	332.9	1131.4	789.1	1800.3	25.9
	1.0	94%	130.8	392.8	1144.0	798.1	1850.8	24.6
	1.0	100%	159.6	456.4	1153.4	807.9	1896.3	25.0
σ_{bias}=.03	.25	75%	46.7	67.1	1181.7	527.5	1408.7	115.1
	.25	87.5%	48.3	81.5	1180.4	531.1	1409.7	122.2
	.25	94%	47.4	88.3	1181.5	531.4	1411.5	122.0
	.25	100%	47.4	95.8	1184.4	532.6	1415.2	123.6
	.50	75%	91.8	200.1	1207.0	741.8	1454.5	68.0
	.50	87.5%	111.1	256.7	1206.8	750.1	1460.6	62.4
	.50	94%	120.7	287.2	1209.4	752.8	1465.9	58.4
	.50	100%	131.8	320.4	1209.3	757.1	1468.9	56.2
	1.0	75%	246.5	441.0	1324.2	1052.9	1628.5	134.7
	1.0	87.5%	319.5	562.9	1332.0	1066.3	1655.8	163.7
	1.0	94%	357.3	627.9	1336.5	1072.5	1670.9	177.8
	1.0	100%	401.3	697.9	1339.0	1078.6	1683.9	197.8
σ_{bias}=.045	.25	75%	139.6	218.8	1504.4	829.5	1680.8	131.6
	.25	87.5%	141.3	241.5	1503.7	830.3	1681.6	124.9
	.25	94%	143.8	252.9	1503.4	831.1	1681.5	123.7
	.25	100%	142.6	264.0	1506.4	831.7	1684.7	117.3
	.50	75%	256.3	394.5	1524.8	1016.9	1709.0	178.0
	.50	87.5%	287.7	460.8	1523.6	1023.2	1711.3	183.2
	.50	94%	304.3	496.3	1524.0	1025.6	1713.6	186.1
	.50	100%	318.4	530.8	1523.9	1028.2	1715.2	184.7
	1.0	75%	448.9	635.8	1563.9	1257.6	1773.1	318.3
	1.0	87.5%	527.1	757.9	1567.5	1267.9	1785.3	355.6
	1.0	94%	567.4	819.2	1569.2	1272.5	1791.7	377.1
	1.0	100%	613.5	887.8	1568.2	1277.2	1796.0	402.0

Table 4: MSE of Alternative Estimators of Value (True CS=$100 & n=80)

Hetero-geneity	σ_e	Selection Incidence	Wn	Mean	FE	RE	$\hat{\beta}_0$	Root-n
σ_{bias}=.015	.25	75%	527.8	1062.8	4832.9	2966.1	5838.6	286.1
	.25	87.5%	539.3	1177.9	4826.6	2995.8	5849.0	259.4
	.25	94%	555.1	1247.6	4836.2	3010.7	5868.7	253.5
	.25	100%	561.7	1310.5	4833.6	3026.7	5873.8	241.7
	.50	75%	1247.0	2110.8	5295.2	4202.1	6535.6	675.5
	.50	87.5%	1420.0	2437.2	5324.0	4258.5	6630.5	733.5
	.50	94%	1511.6	2613.5	5342.8	4291.9	6686.4	761.9
	.50	100%	1608.5	2793.5	5356.5	4315.0	6735.2	794.5
	1.0	75%	2431.5	3527.3	6970.3	5806.9	8608.7	1588.5
	1.0	87.5%	2876.4	4092.8	7029.4	5877.8	8843.5	1905.6
	1.0	94%	3102.6	4379.3	7055.0	5907.5	8961.6	2064.7
	1.0	100%	3343.4	4679.5	7084.5	5935.3	9099.7	2241.7
σ_{bias}=.03	.25	75%	1994.1	2599.4	6864.4	4872.6	7457.2	1553.2
	.25	87.5%	2000.2	2725.7	6870.3	4880.1	7469.2	1474.6
	.25	94%	2023.5	2793.3	6874.9	4892.4	7475.0	1467.4
	.25	100%	2031.6	2861.0	6871.9	4904.5	7474.9	1434.0
	.50	75%	2797.1	3536.1	6955.9	5647.3	7605.2	2203.0
	.50	87.5%	2962.6	3832.2	6956.9	5677.8	7624.0	2257.1
	.50	94%	3053.6	3987.8	6954.9	5684.0	7635.1	2294.6
	.50	100%	3146.1	4147.0	6955.8	5698.3	7646.1	2331.7
	1.0	75%	3782.5	4618.2	7335.9	6620.3	8131.2	3081.5
	1.0	87.5%	4162.4	5095.0	7362.8	6661.5	8213.7	3361.6
	1.0	94%	4351.6	5338.4	7377.6	6682.2	8259.2	3497.8
	1.0	100%	4540.2	5580.8	7389.1	6696.3	8308.1	3630.3
σ_{bias}=.045	.25	75%	3133.6	3686.8	7756.8	5937.4	8177.8	2699.4
	.25	87.5%	3150.3	3807.3	7757.5	5949.5	8181.0	2630.2
	.25	94%	3163.0	3868.8	7754.0	5949.9	8180.3	2606.4
	.25	100%	3176.9	3931.6	7753.6	5955.6	8180.7	2580.2
	.50	75%	3884.7	4504.1	7802.7	6508.2	8244.7	3361.6
	.50	87.5%	4026.6	4759.0	7805.1	6522.4	8256.6	3401.4
	.50	94%	4102.7	4887.7	7803.3	6530.6	8259.9	3432.2
	.50	100%	4192.7	5033.2	7808.2	6536.3	8269.4	3474.1
	1.0	75%	4709.8	5372.9	7924.1	7184.5	8424.6	4133.8
	1.0	87.5%	5010.6	5769.6	7924.6	7206.4	8453.1	4341.3
	1.0	94%	5168.1	5973.8	7933.9	7219.5	8473.0	4450.7
	1.0	100%	5340.6	6188.7	7936.9	7235.3	8488.2	4581.7

Table 5: Root-n MRA of Outdoor Recreation Values

Variables: Dependent= CS	Freshwater Fishing WLS equation (5)	All Outdoor Recreation WLS equation (5)
Intercept	48.74 (14.46)*	62.42 (31.05)
$1/\sqrt{n_i}$	123.7 (1.74)	-148.40 (-2.61)
n	681	2222
R^2	0.4%	0.3%

* *t*-values are reported in parenthesis

Appendix: Further Simulations Results

Appendix Table 1: Means of Alternative Estimators of Value (True CS=$50 & n=240)

Hetero-geneity	σ_e	Selection Incidence	Wn	Mean	FE	RE	$\hat{\beta}_0$	Root-n
σ_{bias}=.015	.25	75%	59.30	52.97	25.48	37.46	19.91	65.05
	.25	87.5%	59.13	51.81	25.43	37.30	19.82	65.83
	.25	94%	58.94	51.17	25.43	37.25	19.78	66.04
	.25	100%	58.89	50.61	25.42	37.18	19.75	66.47
	.50	75%	48.47	43.14	15.34	27.00	12.18	53.47
	.50	87.5%	48.14	42.06	15.31	26.91	12.13	53.81
	.50	94%	48.05	41.56	15.32	26.91	12.12	54.10
	.50	100%	48.03	41.08	15.30	26.88	12.09	54.52
	1.0	75%	40.07	35.84	10.98	21.16	8.78	44.03
	1.0	87.5%	39.88	34.98	10.97	21.12	8.75	44.48
	1.0	94%	39.78	34.51	10.98	21.10	8.76	44.74
	1.0	100%	39.82	34.21	10.95	21.09	8.72	45.09
σ_{bias}=.03	.25	75%	52.00	44.54	23.93	30.62	17.75	58.93
	.25	87.5%	50.65	42.26	23.82	30.34	17.46	58.53
	.25	94%	49.90	41.07	23.77	30.20	17.31	58.21
	.25	100%	49.20	39.93	23.72	30.05	17.16	57.93
	.50	75%	41.78	36.28	14.98	22.68	11.59	46.98
	.50	87.5%	40.50	34.25	14.96	22.55	11.49	46.44
	.50	94%	39.92	33.30	14.98	22.49	11.47	46.23
	.50	100%	39.25	32.26	14.97	22.44	11.42	45.90
	1.0	75%	34.72	30.46	10.72	18.04	8.44	38.76
	1.0	87.5%	33.63	28.76	10.72	17.95	8.40	38.25
	1.0	94%	33.04	27.85	10.73	17.92	8.38	38.00
	1.0	100%	32.60	27.08	10.72	17.88	8.36	37.86
σ_{bias}=.045	.25	75%	42.95	35.04	16.58	22.25	8.52	50.52
	.25	87.5%	40.31	31.93	16.24	21.89	7.50	48.41
	.25	94%	38.89	30.26	16.08	21.70	6.99	47.28
	.25	100%	37.53	28.67	15.96	21.55	6.46	46.19
	.50	75%	34.75	29.17	13.46	17.50	9.54	40.04
	.50	87.5%	32.48	26.42	13.36	17.29	9.23	38.30
	.50	94%	31.27	25.03	13.33	17.18	9.09	37.31
	.50	100%	30.07	23.60	13.28	17.10	8.90	36.34
	1.0	75%	29.20	24.97	10.29	14.48	7.74	33.21
	1.0	87.5%	27.25	22.59	10.24	14.35	7.59	31.70
	1.0	94%	26.30	21.43	10.25	14.28	7.54	30.99
	1.0	100%	25.32	20.23	10.23	14.21	7.47	30.24

Appendix Table 2: Means of Alternative Estimators of Value (True CS=$100 & n=240)

Hetero-geneity	σ_e	Selection Incidence	Wn	Mean	FE	RE	$\hat{\beta}_0$	Root-n
σ_{bias}=.015	.25	75%	78.45	67.81	29.96	45.40	23.07	88.37
	.25	87.5%	78.09	65.98	29.96	45.16	22.96	89.44
	.25	94%	77.94	65.09	29.98	45.03	22.92	90.03
	.25	100%	77.68	64.13	29.95	44.89	22.84	90.41
	.50	75%	56.15	49.38	16.71	30.09	13.24	62.49
	.50	87.5%	55.93	48.09	16.68	29.99	13.18	63.30
	.50	94%	55.68	47.33	16.70	29.91	13.18	63.55
	.50	100%	55.54	46.71	16.68	29.85	13.15	63.88
	1.0	75%	44.46	39.49	11.65	22.84	9.31	49.11
	1.0	87.5%	44.32	38.48	11.62	22.78	9.27	49.83
	1.0	94%	44.08	37.88	11.63	22.74	9.26	49.92
	1.0	100%	44.02	37.43	11.64	22.72	9.27	50.22
σ_{bias}=.03	.25	75%	65.15	54.25	26.99	35.07	19.00	75.51
	.25	87.5%	62.65	50.72	26.75	34.63	18.38	74.12
	.25	94%	61.40	48.96	26.63	34.40	18.08	73.43
	.25	100%	60.13	47.21	26.55	34.21	17.80	72.66
	.50	75%	47.52	40.77	16.23	24.78	12.46	53.94
	.50	87.5%	45.79	38.21	16.26	24.58	12.37	53.05
	.50	94%	44.94	36.92	16.27	24.51	12.31	52.64
	.50	100%	44.07	35.65	16.27	24.42	12.26	52.17
	1.0	75%	37.95	33.06	11.38	19.26	8.93	42.59
	1.0	87.5%	36.71	31.08	11.36	19.15	8.86	42.10
	1.0	94%	36.09	30.14	11.36	19.10	8.84	41.81
	1.0	100%	35.40	29.11	11.35	19.07	8.80	41.44
σ_{bias}=.045	.25	75%	50.80	40.59	16.37	23.75	7.10	60.70
	.25	87.5%	46.52	36.12	16.04	23.31	5.95	56.72
	.25	94%	44.41	33.87	15.84	23.09	5.30	54.81
	.25	100%	42.21	31.60	15.70	22.91	4.62	52.75
	.50	75%	38.66	32.13	14.16	18.55	9.71	44.89
	.50	87.5%	35.59	28.67	14.03	18.32	9.29	42.27
	.50	94%	34.16	27.01	13.96	18.21	9.08	41.11
	.50	100%	32.65	25.30	13.90	18.10	8.83	39.84
	1.0	75%	31.57	26.81	10.80	15.21	8.04	36.11
	1.0	87.5%	29.22	24.05	10.73	15.04	7.87	34.19
	1.0	94%	28.10	22.72	10.76	14.97	7.82	33.30
	1.0	100%	26.95	21.34	10.72	14.88	7.71	32.40

Appendix Table 3: MSE of Alternative Estimators of Value (True CS=$50 & n=240)

Hetero-geneity	σ_e	Selection Incidence	Wn	Mean	FE	RE	$\hat{\beta}_0$	Root-n
σ_{bias}=.015	.25	75%	101.0	15.3	603.0	158.7	906.8	260.7
	.25	87.5%	97.7	9.0	605.6	162.5	912.4	284.6
	.25	94%	93.8	6.8	605.6	163.9	914.4	290.2
	.25	100%	93.3	5.8	605.9	165.7	916.4	305.4
	.50	75%	17.1	53.5	1202.7	530.3	1431.2	47.1
	.50	87.5%	18.2	69.1	1204.7	534.2	1434.9	49.7
	.50	94%	18.4	76.9	1203.8	534.4	1435.5	51.7
	.50	100%	18.5	85.2	1205.4	535.5	1437.9	54.7
	1.0	75%	112.2	206.3	1522.8	832.5	1699.7	67.9
	1.0	87.5%	115.8	231.0	1524.1	834.9	1702.0	62.3
	1.0	94%	118.1	245.3	1522.8	836.0	1701.5	59.9
	1.0	100%	116.8	254.5	1525.9	836.4	1704.4	55.6
σ_{bias}=.03	.25	75%	12.0	33.8	680.8	376.3	1041.0	98.8
	.25	87.5%	7.0	62.7	686.2	387.3	1059.5	88.7
	.25	94%	5.8	82.1	689.0	392.8	1069.4	81.3
	.25	100%	6.2	103.4	691.9	398.6	1079.5	76.3
	.50	75%	75.6	192.1	1227.0	747.2	1475.6	28.2
	.50	87.5%	96.9	250.7	1228.7	753.9	1483.3	28.7
	.50	94%	107.9	281.4	1227.6	757.1	1485.4	29.3
	.50	100%	120.9	316.6	1227.7	760.2	1488.8	29.4
	1.0	75%	241.2	385.7	1543.3	1021.6	1727.6	144.4
	1.0	87.5%	274.1	453.7	1543.4	1027.3	1731.3	153.0
	1.0	94%	293.5	493.0	1542.7	1029.6	1732.6	157.8
	1.0	100%	308.2	527.1	1543.3	1032.3	1734.6	160.2
σ_{bias}=.045	.25	75%	54.5	226.5	1117.6	770.6	1721.9	11.9
	.25	87.5%	97.2	328.3	1140.7	790.4	1807.2	10.3
	.25	94%	125.7	390.7	1151.1	800.8	1851.0	13.1
	.25	100%	156.9	455.5	1159.7	809.5	1896.8	18.0
	.50	75%	238.0	436.8	1335.3	1056.3	1637.3	111.8
	.50	87.5%	310.4	557.6	1342.8	1070.1	1662.6	144.9
	.50	94%	353.2	624.8	1345.0	1077.2	1674.0	167.1
	.50	100%	398.8	697.3	1349.0	1082.5	1689.5	190.7
	1.0	75%	437.4	628.8	1577.5	1261.8	1786.1	293.2
	1.0	87.5%	520.7	752.9	1581.3	1271.3	1798.5	342.3
	1.0	94%	564.0	817.5	1580.5	1276.0	1802.7	367.1
	1.0	100%	610.7	887.1	1581.8	1280.9	1808.6	394.4

Appendix Table 4: MSE of Alternative Estimators of Value (True CS=$100 & n=240)

Hetero-geneity	σ_e	Selection Incidence	Wn	Mean	FE	RE	$\hat{\beta}_0$	Root-n
σ_{bias}=.015	.25	75%	486.2	1045.1	4909.8	2983.9	5920.5	186.7
	.25	87.5%	501.4	1165.5	4908.8	3009.6	5938.2	162.6
	.25	94%	507.4	1226.5	4905.9	3024.4	5944.1	148.3
	.25	100%	519.0	1294.4	4910.1	3039.6	5955.9	140.4
	.50	75%	1940.9	2570.0	6939.1	4888.8	7529.1	1448.5
	.50	87.5%	1960.0	2702.1	6943.3	4902.3	7539.2	1389.7
	.50	94%	1983.4	2781.3	6940.8	4913.9	7539.5	1372.5
	.50	100%	1994.5	2846.8	6943.5	4922.6	7544.8	1347.4
	1.0	75%	3100.1	3668.4	7806.4	5954.9	8225.5	2625.6
	1.0	87.5%	3115.8	3791.2	7811.6	5963.1	8233.2	2552.5
	1.0	94%	3115.8	3791.2	7811.6	5963.1	8233.2	2552.5
	1.0	100%	3149.3	3920.9	7807.8	5972.7	8233.0	2514.0
σ_{bias}=.03	.25	75%	1224.8	2097.6	5331.8	4216.8	6562.4	624.6
	.25	87.5%	1403.8	2432.5	5367.2	4274.0	6663.2	691.2
	.25	94%	1497.3	2608.4	5384.3	4304.8	6711.9	724.5
	.25	100%	1595.7	2789.5	5395.8	4328.9	6758.5	762.9
	.50	75%	2764.1	3512.9	7018.2	5658.9	7664.9	2144.3
	.50	87.5%	2946.8	3821.9	7014.4	5688.7	7680.0	2222.7
	.50	94%	3039.2	3981.4	7012.0	5699.0	7689.9	2260.5
	.50	100%	3134.5	4143.3	7012.0	5713.0	7699.7	2302.6
	1.0	75%	3859.3	4485.3	7854.7	6519.0	8294.6	3315.8
	1.0	87.5%	4013.1	4753.5	7858.2	6536.8	8306.3	3369.0
	1.0	94%	4090.1	4882.3	7857.5	6545.7	8310.3	3401.0
	1.0	100%	4179.1	5028.1	7860.3	6550.6	8318.6	3443.0
σ_{bias}=.045	.25	75%	2427.1	3532.5	6995.0	5813.9	8630.7	1561.2
	.25	87.5%	2864.2	4082.9	7050.9	5881.0	8846.3	1882.4
	.25	94%	3093.2	4374.3	7083.2	5915.8	8969.6	2048.9
	.25	100%	3341.2	4678.8	7106.9	5942.7	9098.6	2235.8
	.50	75%	3768.6	4610.0	7369.1	6634.4	8151.9	3051.5
	.50	87.5%	4152.1	5089.6	7392.1	6672.6	8228.1	3341.8
	.50	94%	4338.2	5328.4	7402.8	6690.4	8267.7	3474.5
	.50	100%	4537.9	5580.5	7413.8	6707.8	8312.4	3623.1
	1.0	75%	4688.0	5359.5	7957.9	7190.0	8456.5	4095.2
	1.0	87.5%	5013.2	5770.1	7969.2	7219.1	8489.1	4339.3
	1.0	94%	5172.3	5974.0	7964.1	7230.0	8497.8	4455.6
	1.0	100%	5337.8	6187.8	7971.6	7245.0	8517.0	4574.3