The Benefit-Transfer Challenges[1]

Kevin J. Boyle
Nicolai V. Kuminoff
Chris Parmeter
and
Jaren C. Pope[2]

Abstract

Presidential Executive Order 12866 requires federal agencies to design "cost-effective" regulations, and assess "costs and benefits" of these regulations based "on the best reasonably obtainable scientific, technical, economic, and other information." Benefit transfers are one economic approach used to estimate these benefits and costs and the use of existing economic information to predict the effects of new policies is well established. However, advancing the practice of benefit transfers is crucial if economists are to play a role in developing federal policies. In this paper we review contributions to the benefit-transfer literature and present a unified conceptual framework to guide the design and evaluation of benefit-transfer guidelines.

Key Words: policy evaluation, benefit transfer, validity

1. Introduction

Economic analyses play a key role in the evaluation of policies promulgated by federal agencies in the U.S. Presidential Executive Order 12866 (1993) requires federal agencies to design "cost-effective" regulations, and assess "costs and benefits" of these regulations that are based "on the best reasonably obtainable scientific, technical, economic, and other information." Benefit transfers are one economic approach used to estimate these benefits and costs. A benefit transfer occurs when an estimated value, based on original studies (study sites), is transferred to a new application (policy site).

Benefit transfers can be transfers through time or space; the key feature is that study-site value(s) are used to estimate a value for a policy that is different from the original policy

[1] Funding for this research was provided by the U.S. EPA. The authors appreciate Kerry Smith's helpful comments on an earlier draft of this paper.
[2] Boyle, Parmeter and Pope are Professor, Assistant Professor and Assistant Professor, respectively, in the Department of Agricultural and Applied Economics at Virginia Tech. Kuminoff is an Assistant Professor in the department of Economics at Arizona State University.

objective. Benefit transfers are typically employed when policy analysts are faced with time or monetary constraints that preclude the conduct of an original study. Benefit transfer are also used to develop liability payments for court cases.

There are two types of benefit transfers, value transfers and function transfers. A value transfer uses a single value from a study site or a mean from multiple study sites to provide a policy-site value estimate. A function transfer uses an estimated valuation function to compute a transfer estimate that is calibrated to policy-site conditions using the variables in the equation. A function transfer could be an estimated preference function from a single study site or a meta-analysis of results from multiple study sites.

The practice of conducting benefit transfers for changes in environmental quality came under considerable academic scrutiny in the early 1990s and a special issue of *Water Resources Research* (1992, Vol. 28, No. 3) focused on this topic. Since then, benefit transfers have become standard practice to evaluate benefits and costs in federal regulatory impact analyses of environmental programs. For example, the Economic Report of the President (2009) includes net benefits of federal policies to improve air quality (Table 3-1, p. 115). These estimates were obtained from regulatory impact analyses conducted by the U.S. Environmental Protection Agency (EPA) that relied on benefit-transfer procedures to develop some of the clean-air benefit estimates (e.g., U.S. EPA, 2005, p. 4-48, Table 4-11).

Despite becoming common practice, even a cursory review of the benefit-transfer literature displays a wide variety of implementation procedures with no consensus on which procedure actually results in the lowest transfer error beyond two general principles:

- study sites and policy sites should be similar, and
- equation transfers are more accurate than value transfers.

These principles are related because equations can be used to predict welfare estimates that are calibrated to policy-site conditions, which is one approach to imposing similarity. No consensus exists on what actually constitutes similarity and what type of equation transfer might work best. This lack of consensus undermines the credibility of benefit transfers as an accepted economic tool for inferring the nonmarket benefits of public actions. In this paper we review selected contributions to the benefit-transfer literature and present a unified conceptual framework that formally addresses similarity, provides guidelines for equation transfers, and provides a basis for evaluating and comparing future contributions to the benefit-transfer literature.

2. Benefit Transfers in Context

Advancing the practice of benefit transfers is crucial if economists are to play a role in the development of major federal policies such as climate change and health care. The Economic Report of the President (2009) states that "energy policy will continue to be one of the major challenges facing the United States for many years to come" (p. 126). The Office of Management and Budget states "the Administration believes that comprehensive health reform should … invest in prevention and wellness (and) improve patient safety and quality of care …"[3] Economic analyses that arise from these national policies can only be accomplished if some information is inferred from existing economic studies.

The practice of using existing economic information to predict the effects of new policies is well established. Transfers are often made to predict the benefits and costs of future policies, although the process is rarely called "benefit transfer" outside of environmental economics. For example, Hines (1999) notes that "… to quantify the economic costs of (taxes, regulations, externalities, monopolistic practices, etc.) … it is standard practice – and has been since the

[3] http://www.whitehouse.gov/issues/health-care/, accessed January 27, 2010.

1960s – to use a small number of assumptions and selected elasticities to estimate areas of the relevant 'Harberger triangles'" (p. 167).[4]

In fact, the transfer of existing information to new applications is not unique to economics:

- engineers use "steam tables" to predict pressure and flow in power facilities,
- architects use "weight-load" tables to predict weight holding capacity of floors and roofs, and
- the U.S. government publishes an actuarial table of the probability of death.

Each type of transfer avoids the need to conduct a potentially time-consuming and expensive study by using past results to predict future outcomes.[5]

Property appraisals are a specific example of transfers of monetary values. A typical appraisal uses sale prices of three nearby and similar properties to develop a calibrated appraisal value for the subject property based on an appraiser's professional judgment of differences between sold properties and the appraised property.[6] Thus, a real estate appraisal is similar to a benefit transfer in that the three selected properties are "study sites" and the appraised property is the "policy site".

Some degree of error is unavoidable in any information transfer. Weather conditions can affect weight-holding capacities of roofs and pre-existing health conditions can affect the average probability of death. In real estate appraisals, an appraiser may not pick the best set of

[4] Harberger triangles (1964 and 1971) are based on elasticities from Marshallian demand functions that are estimated using choices people make in markets.

[5] For links to tables for steam, weight-load, and probability of death, see: (http://www.engineeringtoolbox.com/saturated-steam-properties-d_101.html, accessed October 21, 2009), (http://www.awc.org/technical/spantables/tutorial.htm, accessed October 21, 2009), and (http://www.ssa.gov/OACT/STATS/table4c6.html, accessed September 14, 2009). Additional examples of data transfers include dose-response functions developed by physical scientists that have been used to identify changes in quality for original applications at study sites (Spash and Vatn 2006).

[6] See: http://www.appraisalinstitute.org/profession/appraiser.aspx, accessed October 21, 2009.

comparables, may not know constraints faced by sellers, or may not account for buyers with unique preferences. For example, Fisher, Miles and Webb (1999) found an average error in commercial real estate appraisals in the U.S. between 1980 and 1998 of 11% with 89% of the prediction errors falling between 0 and 25%.

Errors can also arise when using existing information, such as the Harberger triangles, to compute the welfare implications of policies. Hausman (1981) showed that for a single price change, which is similar to many benefit-transfer applications, it is possible to derive estimates of Hicksian consumer surplus and the error associated with using Marshallian elasticities can be small for compensating variation (3.2%) and much larger for deadweight loss (32%). More recently, Chetty, Looney and Kroft (2009) have shown that consumers react differently to different presentations of a tax, which can affect estimated price and income elasticities. Their finding suggests that elasticities estimated in one policy context may involve error when transferred to another policy context with a different framing of the policy.

All of these examples involve transfers of existing information to predict outcomes in new situations. Environmental benefit transfer is perhaps the most difficult type of information transfer. What makes it difficult is that the provision of environmental services almost always occurs outside formal markets. As a result, study-site analysts must define the units of measurement for quality (or quantity) and estimate a Marshallian virtual price or Hicksian willingness to pay. The way in which this is done varies from study to study. The benefit-transfer practitioner is left to develop a strategy for synthesizing existing information on Marshallian and Hicksian willingness to pay for potentially different proxy measures of environmental quality in order to predict welfare outcomes at a policy site.[7]

[7] For example, consider a benefit transfer for a policy that targets air quality. Some studies have used revealed preference methods, such as hedonic models, to estimate Marshallian virtual prices for marginal changes in ozone

To see the contrast between market and non-market benefit transfers consider the Harberger triangle example. Harberger triangles rely on price and income elasticities to infer the welfare consequences of a new policy, while environmental benefit transfers typically use existing welfare estimates to infer the welfare consequences of a new policy. Harberger triangles are typically based on market data for private goods, whereas environmental benefit transfers rely on values estimated using nonmarket valuation methods for publicly provided goods. While Harberger triangles have been used for over four decades, they still face challenges when evaluating public policies. Thus, it is not surprising that environmental benefit transfer, which has only been in use for about two decades, also faces challenges to the credibility of this approach to welfare evaluations of public policies.

3. A Historical Perspective

The scrutiny the special issue of *Water Resources Research* (1992, Vol. 28, No. 3) brought to benefit transfers is not surprising; the library of nonmarket values was growing rapidly and, in response, the use of benefit transfers was expanding. President Reagan's 1981 Executive Order 12291 stated that "*regulatory action shall not be undertaken unless the potential benefits to society for the regulation outweigh the potential costs to society*" (Section 2 (b)). This motivated analysts and decision makers to seek data on benefits and costs to support regulatory impact analyses.

This was a time when researchers were beginning to seriously ask how accurate study-site value estimates were. Cummings, Brookshire and Schulze (1986) concluded that contingent valuation was accurate to ±50% (p. 244). Smith and Kaoru (1990) asked whether travel-cost

concentrations. Others have estimated Hicksian willingness to pay for nonmarginal changes in particulate matter using stated–preference methods.

models were presenting value "signals or noise". Contingent valuation and travel cost were (and continue to be) the methods most frequently used to estimate study-site values. If there was error in the original study-site values, the rhetorical question was "how accurate could benefit transfers be that are based on data with measurement errors?" Although this question is still relevant today, the conduct of nonmarket valuation studies has improved substantially over the last 25 years (Champ, Boyle and Brown, 2003; Mäler and Vincent, 2003). In this paper we will take study site values as given (but not necessarily unbiased estimates) and focus on discussing benefit-transfer protocols and the accuracy of these protocols .

The *Water Resources Research* (*WRR*) articles began to address benefit-transfer protocols and accuracy. Several of the articles presented examples of the state-of-the-art in benefit transfers (Boyle and Bergstrom, 1992; Desvousges, Naughton and Parsons, 1992; Loomis, 1992; Luken, Johnson and Kibler, 1992). Smith (1992) considered convergent validity of the benefit transfers presented in the Desvousges and Luken articles. McConnell (1992) began the conceptual framework for benefit-function transfers through the use of travel-cost models. Walsh, Johnson and McKean (1992) presented a meta-analysis of recreation demand studies as another function-transfer approach that uses the collective information from multiple studies. Atkinson, Crocker and Shogren (1992) investigated Bayesian exchangeability in a hedonic model as a third approach.

The Boyle and Desvousges articles set the foundation for benefit-transfer guidelines that were codified in the U.S. Environmental Protection Agency's *Guidelines for Preparing Economics Analyses* (2000)[8] That include:

- describing the policy case;

[8] The EPA guidelines are currently being revised (http://yosemite.epa.gov/sab/sabproduct.nsf/c91996cd39a82f648525742400690127/2dd3f407cb483bd685257352004b72b7!OpenDocument, accessed October 21, 2009).

- identifying existing, relevant studies;
- reviewing available studies for quality and applicability;
  - basic commodities must be essentially equivalent,
  - baseline and extent of change should be similar, and
  - affected populations should be similar;
- transfer the benefit estimates; and
- address uncertainty (p. 86-87).

The Office of Management and Budget (2003) (OMB) subsequently developed similar guidelines for benefit transfers. A notable difference between the EPA and OMB guidelines is that the OMB guidelines advocated that "*you should transfer the entire demand function (referred to as benefit function transfer) rather than adopting a single point estimate (referred to as benefit point transfer)*" (p. 25).

While the primary use of benefit transfers in the U.S. has been in the assessments of benefits and costs as components of federal Regulatory Impact Analyses, another use has been to develop liability payments for court cases. This latter use of benefit transfers requires adherence to the Daubert standard[9] for scientific evidence that require the benefit transfer procedure:

- be tested for validity,
- have a known error rate,
- be peer reviewed and published, and
- have general acceptance.

[9] http://www.daubertexpert.com/basics_daubert-v-merrell-dow.html, accessed November 21, 2009.

These guidelines and criteria are very general and do not focus on key issues such as what constitutes similarity between study and policy sites, and what type of function transfer works best.

4. Formalizing Benefit Transfers

Benefit transfers, like any economic policy analysis, seek to measure the economic consequences. The analysis starts with a conceptual definition of the value to be estimated. Suppose that the policy-relevant dimension is a change in environmental quality (q_j) and willingness to pay might be defined as:

$$V_i(P, x, q_j^{''}, I_i - WTP_{i,j}; d_i) = V_i(P, x, q_j^{'}, I_i; d_i) \qquad (1)$$

where *V* is the indirect utility function for individual i, *P* is a vector of market prices, q_j is the dimension of environmental quality that is changing (q_j"<q_j'), *x* is a vector of other quality attributes, I is income, $WTP_{i,j}$ is willingness to pay for the increase in q_j, and d_i is demographic characteristics of individual i. Let us assume that q_j is a change in coastal water quality due to the presence of a harmful algal bloom (HAB).[10] HABs can negatively affect the desirability of coastal waters for a variety of activities and nonmarket valuation methods can be used to estimate the benefits of avoiding exposure to a HAB. Common estimation methods include revealed-preference approaches such as travel-cost models (Herriges and Kling, 1999; Parsons, 2003), hedonic models (Taylor, 2003; Baranzini et al., 2008), and stated-preference approaches such as contingent valuation (Bateman et al., 2002; Boyle, 2003) and choice modeling (Bennett and Blamey, 2001; Holmes and Adamowicz, 2003).

[10] http://www.cdc.gov/hab/, accessed February 15, 2010.

The different nonmarket-valuation approaches present a number of challenges to the application of benefit transfers. Revealed-preference methods estimate Marshallian and Hicksian values, while the stated-preference methods estimate Hicksian values. Travel-cost models are capable of estimating nonmarginal changes in consumer surplus, while most hedonic models estimate marginal prices. Revealed-preference methods are most frequently used to estimate use values, while stated-preference methods are commonly used to measure both use and nonuse values. Thus, the benefit-transfer analyst must sort through the relevance and appropriateness of the welfare measures represented in the available studies.

In the case of a HAB, which can occur anywhere along the Atlantic Coast, Gulf of Mexico or Pacific Coast in the U.S., a key challenge in benefit transfer is making sense of the available economic information in a coherent manner.[11] There may have been travel-cost, hedonic and choice-modeling studies done at three different locations, all of which are different from the new policy site, and the analyst must decide how to best utilize this information in the benefit transfer. The travel-cost study might provide Marshallian use values, the hedonic study might provide a Marshallian implicit price, and the choice-modeling study might provide a Hicksian measure of total value that includes use and nonuse values. Thus, just as in an original study where a definition of the value to be estimated, such as equation (1), is needed to guide data collection and analysis, a similar framework is needed for benefit transfers to select studies (data) and compute benefit transfers via an equation.

Estimated willingness to pay (*WTP*) in any benefit transfer may depend on the demographic characteristics of consumers (d_i), the levels of other site characteristics (X_j), a vector of econometric parameter estimates ($\hat{\beta}_j$), and quality:

[11] http://tidesandcurrents.noaa.gov/hab/, accessed February 15, 2010.

$$WTP_{i,j} = f\left(d_{i,j}, q_j, X_j, \hat{\beta}_j\right) \tag{2}$$

where the observed relationships have been estimated from study site(s) j based on equation (1). Equation (2) provides the first insight into similarity between study sites and policy sites. Similarity includes characteristics of affected individuals (d_i), site and other location characteristics (q_j and X_j), and preferences of affected individuals (β_j). Benefit transfers use the relationship in (2) to predict the benefits from a change in *q* that is calibrated to policy site *k's* conditions. This is typically done through value transfers or equation transfers.

Value transfers. Value transfers apply a single statistic (usually an average from one or more study sites) to the policy site. For example, in the case of a single study site, aggregate benefits for a comparable change in *q* at the policy site is predicted by $M \times \left(\sum_{i=1}^{N} WTP_{i,j} \Big/ N \right)$, where *M* and *N* are the number of people at the policy and study sites. An alternative, and more common, value transfer is to average the means from *K* study sites and the benefit-transfer estimate is $M \times \left(\sum_{s=1}^{K} MWTP_s \Big/ K \right)$, where $MWTP_s$ is mean willingness to pay from study-site s and K is the number of studies. The U.S. Environmental Protection Agency used this approach to estimate the annualized benefit of reduced mortality due to limits on particulate matter imposed by the Clean Air Act (U.S. EPA, 1999). This approach to benefit estimation is appealing for its simplicity, but value transfers do not account for any differences between the study sites and the policy site. If *MWTP* is increasing in income, for example, and if income distributions differ between the two sites, one would need to adjust for this difference to develop an accurate benefit measure. Making adjustment to calibrate transfer predictions to policy-site conditions is the focus of benefit-function transfers.

Function transfers. Function transfers use an econometric model such as equation (2) to predict a calibrated value for a new policy as a function of variables describing characteristics of the policy site and people at the policy site. In this case, aggregate benefits for a quality change at site *k* could be approximated by $\sum_{i=1}^{M} f\left(d_{i,k}, q_k, X_k, \hat{\beta}_j\right)$. This was the approach the U.S. Department of Agriculture used to predict the water quality and wildlife habitat benefits associated with "environmentally friendly" farming practices subsidized by the Conservation Reserve Program (USDA, 2005). There are a variety of function transfers and the more common types are described in the following section.

5. A Diversity of Function Transfer Methods

A recent special issue of Ecological Economics (EE) focused on benefit-transfer methods (Wilson and Hoehn, 2006). This was accompanied by edited books by Rolfe and Bennett (2006) and Navrud and Ready (2007). Together, these assemblages of studies represent the current state of the art in benefit-function transfer. The current literature clearly focuses on function transfers to develop calibrated value predictions for policy sites. Most function transfers fit within one of two broad categories: "reduced-form" meta-analysis or "structural" transfers of preference functions.

5.1. Reduced Form Meta-Analysis

This approach provides an informative way to summarize results across a large number of distinct study sites that investigate a common policy value. A group of studies can collectively cover variation in site and population characteristics that is not possible in a single study.

The meta-analysis approach begins by regressing study-site value estimates on a set of variables describing the data (site and population characteristics) and methods used to estimate values. For example, a linear version of the equation could be

$$WTP_j = \alpha + \beta_1 d_j + \beta_2 q_j + \beta_3 X_j + \beta_4 A_j \qquad (3)$$

where d_j is the demographic characteristics of consumers at site *j*, q_j is the size of the quality change that was valued, X_j contains other relevant site characteristics, and A_j is a set of variables describing study-design features that may help explain variation in value estimates (e.g. data collection procedures and valuation methods). Plugging policy-site values for d_k, q_k and X_k into (3) provides a calibrated prediction about willingness to pay at policy site *k*. A meta-analysis by Mrozek and Taylor (2002) was the basis of the EPA benefit transfer of values for air quality improvements reported in the Economic Report of the President (2009).

Meta-analyses have been used by researchers and government agencies to conduct benefit transfers, but the validity of this approach is relatively unexplored (Johnston et al., 2005; Rosenberger and Loomis, 2000; Shrestha and Loomis, 2001; Smith and Pattanayak, 2002). Lindhjem and Navrud (2008) "*question whether the use of meta-analysis for practical benefit transfer achieves reliability gains*" (p. 425). However, their conclusion is based on the results from a benefit transfer across different *countries*. Other investigators have reported that sub-national transfers tend to be more accurate (Loomis, 1992; Vandenberg, Poe and Powell, 2001).

The error in benefit-transfer predictions from meta-analysis equations depend on the bias in the estimates of the β vector and the spurious error associated with omitted variables that may be unique to the policy site. Nelson and Kennedy (2009) highlight five potentially important issues: (i) sample selection; (ii) data summary; (iii) data heterogeneity; (iv) heteroskedasticity;

and (v) dependency. Some, but not all, of these issues have been addressed in the context of benefit transfer.

A unique feature of the meta-analysis approach is the ability to address some potential sources of bias and error explicitly. The results from past valuation studies may reflect systematic decisions about what questions to research and what methodologies to use. Controlling for these investigator decisions through the A_j term in (3), one can measure their influence on past value estimates and choose the set of study-design conditions to be imposed on the benefit-transfer estimate. Rosenberger and Johnston (2009) conclude that "*meta-analysis offers … an ability to correct for selection effects that is not available in other mechanisms for benefits transfer*" (p.424). Stapler and Johnston (2009) conclude *"…methodological covariates within benefit transfer is neither the source of the substantial errors sometimes intimated by the literature, nor can it be ignored as trivial in all cases."* (p. 244).

Finally, Smith and Pattanayak (2002) note that *"Meta-analyses summarizing non-market valuation studies have often not met the goal of measuring `identical' concepts"* (p.274). Even if numerous studies exist to conduct a meta-analysis, the results can be misleading if the objects being valued are dissimilar. For example, hedonic and travel-cost models produce Marshallian measures of value whereas contingent valuation and random utility models produce Hicksian measures of value. One approach to this problem is to be judicious in the studies selected for the meta-analysis. Another solution is to adopt a "structural" approach to benefit transfer that uses a specification for the preference function to make consistent adjustments between different concepts of value.

5.2. Structural Preference Functions and Preference Calibrations

Two types of function transfers are capable of ensuring the benefit-transfer process is theoretically consistent: preference transfers and preference calibration. Both approaches seek to estimate structural preference parameters for consumers at one or more study sites and then transfer information from the corresponding utility function(s) to the policy site. The difference between the two methods is that preference transfer uses the results from a previous study to transfer a utility (or demand) function to the policy site (Parsons and Kealy, 1994; Zanderson, Termansen, and Jensen, 2007), whereas preference calibration uses previous estimates of nonmarket values to calibrate the parameters of a utility function posited by the analyst (Smith, Van Houtven and Pattanayak, 2000). In either case, the resulting utility parameters can be combined with information on site quality, income and other consumer demographics to compute the transfer value calibrated to policy-site conditions.

Preference-Function Transfers. This approach begins by specifying a parametric form of the utility function at the study site. It is common to use an indirect utility function that is linear in parameters. In this case, individual *i*'s utility from choice *j* at the study site can be expressed as

$$V_{ij} = \alpha_{1,i}\left(y_i - p_{ij}\right) + \alpha_{2,i} q_j + \alpha_{3,i} X_j + \varepsilon_{ij} \qquad (4)$$

where $\varepsilon_{ij} \sim \text{type I extreme value}$, *y* is income, and *p* is the cost of consuming choice *j*, and the preference parameters may be a function of consumer demographics and/or a random error, $\alpha_i = f\left(d_i\right)$. Estimates of the preference parameters for this random utility model can be combined with data on the characteristics of the policy site and demographics to predict a calibrated willingness to pay for a change in *q*. An advantage of this approach is that many objects of valuation are multi-attribute goods/services and (4) allows for marginal values to be estimated for these component attributes. The valuation methods best able to support preference-function transfers are choice modeling, site-choice travel-cost models (RUMs), and structural hedonic models where the estimated preference functions include study-site and demographic characteristics.

Preference-function transfers provide theoretically consistent welfare measures for the policy site. Yet, theoretical consistency comes at a cost. The preference function that is transferred is based on the results of a single study site rather than all of the relevant research that has been done. Useful information from other studies may be excluded.

Preference-Calibration Transfers. Smith and co-authors (2000, 2002, 2006) addressed the multiple-study limitation to structural benefit transfers by proposing an approach to calibrating the preference function that would use all congruent and available information to estimate a common economic value. Their proposal for "preference calibration" is the structural equivalent to meta-analysis.

The process of preference calibration begins by choosing a parametric specification for the preference function, $U(q, X, d; \alpha)$. Then the analyst derives the relationship between the parameters of that function from benefit measures reported by previous studies. For example, suppose that one study estimated the Marshallian consumer surplus (MCS_j) for a change in *q* at study site *j* and another study estimated the compensating variation (CV_k) for a different change in *q* at study site *k*. Both welfare measures can be expressed in terms of the underlying arguments of utility,

$$MCS_j = f(q_j, X_j, d_j; \alpha_1) \quad \text{and} \quad CV_k = g(q_k, X_k, d_k; \alpha_2), \tag{5}$$

where $\alpha_1 \cup \alpha_2 = \alpha$. Parametric versions of the two equations in (5) serve as moment conditions for GMM estimation of the parameter vector α, using data on MCS_j, q_j, X_j, d_j and

CV_k, q_k, X_k, d_k. Intuitively, α is chosen to calibrate the preference function to approximately reproduce the estimates reported by the two previous studies. Finally, the calibrated preference function can be used to predict the welfare implications of a quality change at the policy site:

$$CV_{ps} = h(q\ , X\ , d\ ; \alpha) \tag{6}$$

where CV_{ps} is estimated compensating variation for the policy site.

Unlike the "reduced-form" approach to meta-analysis, preference calibration guarantees that the transfer process is consistent with economic theory. It ensures that the model will not make predictions for benefit measures at the policy site that exceed income, for example. It also forces the analyst to be explicit in the assumptions that they make about how to specify the utility function, which original studies to select, and how to translate the information from different studies into common units of measurement. Preference calibration is subject to the same data limitations that exist for reduced form meta-analyses.

5.3. A Unified Conceptual Framework for Benefit-Function Transfers

The choice among function transfer methods has yet to be carefully studied. There is no consensus on which type of function transfer provides the most accurate predictions. However, a common set of assumptions is required of all function transfers to ensure the transfer process will yield a consistent measure of value at the policy site. Boyle et al. (2009) illustrate this using a unified conceptual framework for the initial estimation of benefits and the subsequent transfer of values. They demonstrate that four "***S***" conditions are necessary to guarantee that benefit estimates at the policy site will be econometrically consistent: (i) utility must be ***S***eparable in unobserved characteristics at the study and policy sites; (ii) the study site and policy-site models must be ***S***pecified correctly; (iii) people must not be ***S***orted between the study and policy sites according to unobserved features of their preferences; and (iv) data on the characteristics of consumers and their choices must be free of ***S***election problems. The four "***S***" assumptions are necessary for the consistency of all benefit transfers and sufficient to establish the consistency of preference function transfers.

The ***S***eparability assumption implies that it is possible to explain the values that consumers assign to the amenity of interest in terms of the observable features of the data. This assumption is routinely invoked in random-utility models for computational convenience. For example, in equation (4) the time-constant unobserved variables, represented by ε_{ij}, cancel out of individual measures of compensating variation. While there is no simple test for separability in variables that are unobserved, intuition may provide some guidance. For example, it is likely problematic to assume separability in situations where a consumer's valuation of the relevant amenity depends on their fitness, or the fitness of their family members, or other metrics for health that are difficult to observe outside of specialized surveys.

Nelson and Kennedy (2009) conclude that numerous studies in environmental and natural resource economics have abused the meta-analysis methodology by failing to adequately deal with ***S***pecification issues (p.372).[12] When we move from meta-analysis to preference functions that have been estimated or calibrated, there is less precedent for developing specification tests. One possibility is to exploit unexpected shocks. If one can observe consumer behavior before and after a shock to the market, there may be an opportunity to judge the out-of-sample performance of competing specifications for the preference function (Provencher and Bishop, 2004). Even without such data, Banzhaf and Smith (2007) and Kuminoff (2009) demonstrate it may be possible to assess the sensitivity of value estimates to subjective modeling decisions.

[12] See also Stanley (2001).

The "no ***S***orting" assumption is made to rule out any systematic variation in unobserved preferences between the study-site populations and the policy-site population. Put differently, any systematic variation in preferences between the two populations is assumed to be fully explained by differences in their observable demographic characteristics. This is most likely to be true for transfers based on unexpected events, such as the discovery of a cancer cluster, where people did not have a prior opportunity to sort themselves according to their preferences over risk (Davis, 2004). The assumption is more problematic for transfers between the site of an unexpected event and a site with long-term contamination, such as a Superfund site, where people have knowledge of the quality change and had considerable time to move.

When considering the scope for sorting, it is important to keep in mind that sorting is simply a metaphor for the way that market forces partition consumers across the landscape; the process need not be premeditated or explicit. If preferences for spatially delineated amenities are a function of past experience with those amenities, for example, then "sorting" may arise even if mobility is limited or nonexistent. Unfortunately, there is no simple test for the presence of sorting. The fact that participants in the annual American Housing Survey consistently choose "looks/design of neighborhood" as one of the top factors influencing where they choose to live suggests that sorting does occur. Yet, there is very little evidence on the spatial and temporal scales over which sorting occurs in response to specific amenities (Rhode and Strumpf, 2003; Banzhaf and Walsh, 2008).

The final assumption, ***S***election, is that data are available on the relevant demographic characteristics of the policy-site population. This is important because we often expect willingness to pay for environmental quality to depend on income, and perhaps other variables. Without direct data on the policy-site population, one approach would be to use the Census of

Population and Housing to predict their demographic characteristics. Yet, this approach raises the possibility of "site selection bias". The people who choose to visit beaches and parks, for example, may have demographic characteristics that differ from the average household living in the area. We are unaware of any existing strategies to detect and correct for this problem.

6. Convergent Validity of Benefit Transfers

Research on the validity of benefit transfers has largely focused on convergent validity, which is a special type of what Carmines and Zeller (1979) call "criterion-related validity". Convergent validity asks whether parallel studies, one conducted at the study site and one conducted at the policy site, provide statistically similar estimates of value or if study-site data can be used to predict a calibrated policy-site value; the policy-site estimate being the validity criterion. A benefit transfer is deemed valid in this context if one fails to reject the null hypothesis of no difference between the policy-site value estimate and the predicted value obtained by calibrating the study-site estimate(s) to policy-site conditions.

Unfortunately, the literature has not adopted a consistent approach to evaluating the validity of benefit transfers. Studies of convergent validity have:

- tested for differences between mean values;
- tested for differences between individual parameters in value functions;
- tested for differences in entire vectors of parameters in value functions;
- measured transfer prediction error, but not in a consistent manner; and
- made other statistical and qualitative comparisons.

In addition, the validity studies have not consistently addressed similarity between study and policy sites in terms of site and population characteristics.

Four papers demonstrate different approaches that have been implemented to investigate the validity of benefit transfer (Table 1).[13] All of these studies are benefit transfers where a stated-preference method was used to estimate study-site values.

Downing and Ozuna (1996) investigated the transfer of values for eight saltwater fishing sites in Texas with observations from three different years at each site. They rejected the null hypothesis of no difference in the parameter estimates from benefit functions between 41% and 63% of the time over four sets of comparisons. From these results they concluded that "*benefit functions transfer … is unreliable*" (p. 322). The authors also observed that confidence intervals on welfare estimates did not overlap in over 90% of the 16 comparisons conducted. The results of this study are somewhat surprising since this is actually a test of temporal reliability (values for the same site for three consecutive years). Other authors have shown that contingent-valuation estimates, the valuation method used by Downing and Ozuna, are reliable (Carson et al., 1997; Loomis, 1989; Teisl et al., 1995; Whitehead and Hoban, 1999).

Kirchhoff, Colby and LaFrance (1997) investigated transfers between two natural area sites in Arizona and to two white-water boating sites on the Rio Grande River in New Mexico. Their transfers were over different sites and different populations of people. They reject the null hypothesis of equality of benefit-function parameters for the Arizona and New Mexico comparisons. They were not able to reject the null hypothesis of no difference in welfare estimates. The authors conclude that "*empirical results indicate that benefit function transfer is more robust than transfer of average site benefits. Our results suggest that the circumstances under which benefit function transfer provides valid, policy-relevant information may be limited*

[13] We selected the benefit-transfer validity studies from peer-reviewed journal articles that had 100 or more *Google scholar* citations.

and that errors from applying benefit transfer can be quite large, even across seemingly similar amenities" (p. 75).

Brouwer and Spaninks (1999) investigated transfers between two different areas of wildlife management on agricultural peat meadows in the Netherlands. They conducted transfers where only the populations varied and where both the populations and sites varied. They rejected the equality of benefit-function parameters in three of the four tests and rejected the null hypothesis in one of the two comparisons of mean values. The authors concluded "*that in the case of statistically valid benefits transfer, the function approach results in a more robust benefit transfer than the unit value approach*" (p. 95).

Finally, Morrison et al. (2002) investigated the transfer of values from two study sites to one policy site for wetland values in Australia. Their transfers were for different populations for the same site and the same population for different sites. They could not reject the null hypothesis of no difference in the implicit prices for wetland attributes in five of the eight comparisons. They rejected the null hypothesis of equality of benefit-function parameters for two comparisons. The authors concluded that "*transfers across sites showed greater evidence of convergent validity than across population transfers*" (p.170). A key element of the Morrison study is that choice modeling was used to estimate values at the original study sites, and the estimated implicit prices for site characteristics facilitated the calibration of transfer estimates to policy-site characteristics.

Three insights can be drawn from these studies:

1. Different tests of convergent validity result in different conclusions regarding the validity of benefit transfers. There is a need for discussion of which tests are most appropriate.

2. Function transfers, because of their ability to customize predictions to policy-site conditions, tend to have lower transfer errors.
3. More research is needed before we can draw broad conclusions about the precise conditions under which benefit-function transfers can be expected to work.

A number of other studies have investigated the validity of benefit transfers with results that are as mixed as the findings from the four studies discussed above.

Transfer errors from these four papers are reported in Table 1 and range from 2 to 87%. If the Downing transfer errors are removed because of their inconsistency with the findings from studies of temporal reliability, average transfer errors range from 32 to 39% (~35%).

All of the studies reported in Table 1 used stated-preference methods. Parsons and Kealy (1994) conducted the first convergent-validity study of benefit transfers using a revealed preference method, a random utility travel cost model. They observed an average error of 21% (n=11) with a range from 1 to 75%. More recently, Zanderson, Termansen, and Jensen (2007), in another travel-cost transfer, observed an average transfer error of approximately 53% (n=53) with a range from 1 to 229% using the authors' preferred transfer function. The mid-point of the Parsons and Zanderson means is 37%, about the same as the 35% mid-point of the latter three studies in Table 1. The Parsons range of errors is similar to that of the latter three studies in Table 1 (2-87% vs. 1-75%), but this is not the case for the Zanderson study (2-87% vs. 1-299%).

Two recent convergent-validity studies of benefit transfers report transfer errors based on choice-modeling data that allow for value predictions calibrated to policy-site conditions. Colombo, Calatrava-Requena, and Hanley (2007) report an average transfer error of 66% (n=27) with a range from 11 to 366% (authors preferred preference-function). Johnston (2007) estimated an average error of 37% (n=24) with a range from 7 to 101%. Thus, Johnston's study confirms a transfer error of about 35% and the error range is similar to the latter three studies in Table 1. However, the Colombo study indicates that the average error could be twice as large (66/35=1.9) and the maximum error could be four times as large (366/87=4.2).

The rhetorical question is how accurate are benefit-transfer errors compared to other estimation errors. Six of the eight studies just discussed used stated-preference methods to estimate study-site values. Murphy et al. (2005) report a median error in original study-site values of 35% when stated-preference methods are used and the error could be as large as 2500% (n=87). The error in commercial real estate appraisals cited earlier is only 11%, but this level of accuracy may be unrealistic for benefit transfers where there are not a lot of repeat studies of similar populations and resources.

If we assume that study-site values are estimated consistently, the weight of evidence from the cited studies suggests that benefit transfers might have an average error of about 35% with exceptions indicating that the error could be even higher. Two clarifications are important for interpreting the reported errors. First, the errors reported are absolute errors and benefit transfers may result in over estimates or under estimates. Second, our discussion takes transfer errors at face value and does not explore the quality of the original study-site values or the quality of the transfer. Controlling for study quality, which is not easy, may reduce the transfer errors.

With an average transfer error rate of 35% or greater there is clearly room for improvement in the conduct of benefit transfers. At least three issues must be addressed in future convergent-validity studies of benefit transfers. First, studies of convergent validity must explain how they address EPA and OMB guidelines for implementing benefit transfers. If these guidelines are not followed in validity research, a schism arises between research and practice that limits the applicability of research results. In particular, a key topic in the benefit-transfer guidelines is establishing "*similarity*" between study sites and policy sites in terms of:

- the affected populations whose values are elicited,
- the physical conditions of the object being valued,
- the institutional settings,
- the availability of substitutes, and
- other site-specific characteristics.

Similarity was implicitly addressed in a number of the studies discussed above, but not in a systematic or rigorous manner. Downing and Ozuna compared the same sites at different points in time, but site and user characteristics could have changed through time. Kirchhoff, Colby and LaFrance compared two different sites on the same river. Morrison et al. compared the same wetlands with different affected populations and the same affected populations with different wetlands. Parsons and Kealy investigated different subsets of fishing sites. None of these studies presented a framework for quantifying, documenting, and testing for similarities and differences between study and policy sites.

The second issue is determining which tests of validity are most appropriate. For the most part, the literature has been agnostic about the choice of validity test, but this is an important issue; differences in study conclusions can be driven by the use of different test procedures. We suggest that testing the equality of benefit-function parameters is fundamental because equations allow study-site estimates to be combined with policy-site data to predict a calibrated transfer value; equality of parameters is indicative that different populations value a common resource similarly, which is a key dimension of the EPA and OMB guidelines. A potential limitation of testing for equality of benefit-function parameters is that, in a worst-case scenario of type-I error, the null hypothesis could be rejected if a single pair of parameters is significantly different. The significant difference could be driven by a parameter that describes a variable that has little or no relevance to the transfer value. A related concern is that well-designed studies may tend to have smaller error components, which would imply that better designs might be more likely to reject the null hypothesis of no difference in estimated benefit functions, but may still possess small transfer errors. One can hedge against these limitations by carefully crafting a test of equality to reflect the relevant subset of parameters and balancing results from statistical tests with computed transfer errors.

The third issue is that tests of convergent validity are not strongly convincing about whether benefit transfers are valid or not. Convergent validity only establishes whether study and policy-site estimates are equivalent. Both estimates can (and are likely to) include biases. If these biases are induced by study design choices, then tests of convergent validity can not identify such endogenous effects. In addition, convergent-validity tests may be compromised by features of models and data that have nothing to do with the validity test itself, but nevertheless affect value estimates. For example, an attribute may be measured with error at the study site or the model of utility may be misspecified at the policy site. Thus, it is possible for convergent-validity tests to spuriously confirm or reject the validity of a benefit transfer.

Finally, it worth noting that some authors have considered alternatives to standard testing for equivalence in convergent-validity studies of benefit transfers. Muthke and Holm-Mueller (2004) introduced “accuracy-t-test” and Kristofersson and Navrud (2005) applied the concept of “equivalence testing”. Neither of these approaches overcomes the fundamental issues with convergent-validity tests. In contrast, Bateman and Jones (2003) introduce a hierarchal modeling framework that allows clustering of "similar" study-site data that provides a rigorous approach to reducing transfer errors when meta-analyses are employed.

7. Moving Forward

Two key features of benefit transfers are important to consider: 1) better data and 2) transfer functions that account for the inherent uncertainty in predicting policy-site values. Data is crucial because once preference-function parameters are estimated it is the availability and quality of characteristic data that will affect the accuracy of calibrated policy-site value predictions.

7.1. Enhancing Function Transfers Through Enhanced Data

Troy and Wilson (2006) investigated the use of a Geographic Information System (GIS) to make resource conditions explicit between study sites and policy sites (see also Eade and Moran, 1996). This data-rich environment provides an opportunity to include quality and demographic characteristics for calibrated transfer predictions that are spatially explicit. This approach will only be successful if original studies are designed to include the level of detail available in GIS databases. This approach is only relevant for natural-resource applications where the valuation attributes are spatially explicit (e.g., the valuation of wetlands and other types of natural areas) and is not relevant for applications where this is not the case (e.g., health risks from drinking contaminated municipal tap water).

Brouwer (2000) notes that while some value estimates may be precisely defined in terms of the increment of change (e.g., hedonic-price functions) other studies may not present clearly defined values (e.g., stated-preference studies of nonuse values). Both of these types of studies present challenges to benefit-transfer practitioners. The former, while having a precise definition of the change valued, may not have the change defined correctly or in a manner that is too narrow to transfer. However, if the definition is correct and matches policy-site conditions, this greatly facilitates accomplishing a credible benefit transfer. In the case of poorly defined changes in values there is little hope for a credible transfer. These limitations apply to all types of function transfers and highlight the need for careful deign and documentation of study-site designs.

Benefit transfers can also be enhanced by improving the use of existing data. Moeltner, Boyle and Paterson (2007) use Bayesian methods to consider challenges to classical meta-analyses: i) the trade-off between increasing sample size and a reduced set of regressors common to all studies, and ii) the treatment of original study valuation method in the analysis. This is done in the context of allowing for random parameters to allow for heterogeneity and correlation between study-site value estimates. Over four welfare scenario comparisons the difference between the classical meta-analysis and Bayesian estimates is about 8%. Thus, the Bayesian framework provides an approach that explicitly addresses important data issues that and reduce transfer errors.[14]

7.2. A New Approach to Function Transfers--Bounding

Both structural and reduced form approaches to benefit function transfer are typically used to predict mean or median values for quality changes at the policy site. Confidence intervals on these predictions rely on maintained assumptions about the study-site model. For example, preference function transfers based on random parameter logit models (e.g. Zanderson Termansen and Jensen, 2007; Colombo et al., 2007) are conditioned by assumptions about the

[14] Leon-Gonzalez and Scarpa (2008) investigated Bayesian averaging to identify the set of study sites that best describe policy-site conditions. They found that failure to address this site similarity can have a substantial effect on benefit-transfer estimates.

statistical distributions used to describe variation in unobserved sources of preference heterogeneity (e.g. normal, type I extreme value, etc.). The extent to which these maintained assumptions ultimately drive transfer values is unclear. This is important because the final step in EPA's benefit transfer guidelines requires practitioners to address uncertainty about transferred values. At present, there is virtually no discussion in the literature of how to quantify uncertainty about the structure of the study-site model.

Recent advances in discrete-choice estimation have provided new ways to measure structural uncertainty. Berry and Pakes (2007) and Bajari and Benkard (2005) develop semiparametric estimators that require a specification for the preference function, but avoid the need for distributional assumptions. Relaxing distributional assumptions means that individual preferences are no longer point-identified. Instead, observed choices identify a set of preference parameters for each consumer. Kuminoff (2009) illustrates how these preference sets can be translated into bounds on welfare measures for changes in environmental quality at a study site. The bounds reflect the analyst's uncertainty about the true structure of the study site model.

The bounding logic from the literature on discrete choice estimation can be adapted to preference-function transfers. The first step of the process is to solve for bounds on the set of values for the preference parameters that are capable of explaining the observed behavior of each individual at the study site. These preference sets are then used to define bounds on the preference function. For example, a bounded version of the indirect utility function in (3) can be expressed as

$$V_{ij} = \alpha_{1,i}\left(y_i - p_{ij}\right) + \alpha_{2,i}q_j + \alpha_{3,i}X_j + \xi_j, \quad \text{where} \tag{5}$$

$$\left(\alpha_{1,i}, \alpha_{2,i}, \alpha_{3,i}\right) \in \left\{ \tilde{\alpha}_{ij} : j^* = \max_j V_{ij}\left(\alpha_i; y_i, p_{ij}, q_j, X_j, \xi_j\right) \right\}.$$

In words, $\tilde{\alpha}_{ij}$ is the set of preference parameter vectors for which the model predicts that consumer *i*'s utility-maximizing choice, $j = j^*$, is the same choice they were observed making at the study site.[15] Finally, the model in (5) is transferred to the policy site and used to predict upper and lower bounds on the value that each individual would assign to a change in quality, $CV_{i,\min} = \min_{\alpha_i \in \tilde{\alpha}_{ij}} CV_{ps}(\alpha_i; y_i, p, q, X, \xi)$ and $CV_{i,\max} = \max_{\alpha_i \in \tilde{\alpha}_{ij}} CV_{ps}(\alpha_i; y_i, p, q, X, \xi)$. Since there is a range of predicted values for each individual, the process is not technically a "function transfer". It is a "correspondence transfer". By providing bounds on point estimates for aggregate transfer values, Boyle et al.'s conceptual model for correspondence transfer begins to address EPA's guidelines for addressing uncertainty.

Data transfers based on maxima and minima have been quite useful in other contexts, such as the "weight-load" tables used by architects to judge the maximum weight holding capacity of floors and roofs. In the context of environmental policy, a correspondence transfer could allow policymakers to judge the maximum cost of a potential deterioration in quality at the policy site or the minimum benefit from an improvement. While Boyle et al. (2009) focus on preference functions, the logic of correspondence transfer is not limited to structural estimation. A similar approach could be developed for reduced form meta-analysis following Manski (1990).

7.3. Concluding Comments

A common theme in the benefit-transfer literature is that function transfers outperform value transfers because transfer values can be calibrated to policy-site conditions. This is a key

[15] Notice that the idiosyncratic error term (ε_{ij}) from (3) has been replaced by a composite unobserved characteristic (ξ_j) in (5). To highlight this distinction, Berry and Pakes (2007) label (5) the "pure characteristics" model of consumer behavior to distinguish it from the random utility model in (3).

element in addressing similarity between study and policy-site conditions. These function transfers can be supported by any of the existing nonmarket valuation methods (Champ, Boyle and Brown, 2003). Just like original valuation studies, benefit transfers are hard to do well and require rigorous analysis. The lack of time and funds for the conduct of an original valuation study is no excuse for not conducting a rigorous analysis of the available valuation data. All future function transfers, whether they are practical transfers to support public decision making or studies of the validity of this approach, should explicitly address the 4S assumptions that underlie all function transfers. These analyses will enhance the policy credibility of transfers and make investigator assumptions, which are required in every empirical analysis, explicit and transparent.

As noted in the introduction to this paper, Presidential Executive Order 12866 (1993) requires federal agencies to evaluate changes in public policies based "on the **best reasonably obtainable** scientific, technical, economic, and other information" (emphasis added). The discussions presented here indicate benefit transfers are capable of providing estimates for specific changes that are based on economic theory, have a rigorous analysis to support a scientific estimate, and have a known error range. However, the tendency towards an average error of approximately 35% must be interpreted with caution. This is a relative average error, not an absolute error because original study-site values are estimated with error.

The final benefit-transfer challenge comes from how transfer estimates are used in decision making. Let us use an example where study-site values were estimated using a stated-preference study, which is common of many benefit transfers. Assume the study-site values are overestimates by 35 percent (median error from stated-preference validity studies). If the benefit transfer underestimates a policy-site value by 35%, then the real transfer error is 0% (35%-35%).

On the other hand, if the benefit transfer overestimates by 35%, the real transfer error is 70% (35%+35%). From the information presented above, one could assume that the original study error and the transfer error are as likely to be less than 35% than they are to be greater than 35%.

One approach to addressing this uncertainty in the real error associated with benefit transfers may be to consider using a decision framework such as the "safe minimum standard" proposed by Bishop (1978) with adjustments on whether the policy goal is to remove a "bad" (e.g., contamination at a super fund site) or to attain a "good" (e.g., prevention of future global warming effects) and the reversibility of the decision. Another approach is to use a bounding approach as described above, which estimates lower and upper bounds, rather than a specific point estimate of central tendency. If the transfer bounds are used in estimating benefits, then a lower-bound aggregate benefit that exceeds costs would be strong evidence the policy was efficient. Or, if the transfer bounds were used to estimate costs, an upper-bound aggregate cost that exceeds benefits would be strong evidence that the policy was inefficient. This "real transfer error" has not and cannot be addressed by convergent-validity studies. Future research on the validity of benefit transfers must address "real" transfer error rates and conditions that lead to over or underestimates, not just the magnitude of error.

References

Atkinson, Scott E., Thomas D. Crocker, and Jason F. Shogren. 1992. "Bayesian Exchangeability, Benefit Transfer, and Research Efficiency." *Water Resources Research*, Vol. 28 (3): 715-722.

Bajari, Patrick and C. Lanier Benkard. 2005. "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach." *Journal of Political Economy*, 113(6): 1239-76.

Banzhaf, H. Spencer and Randall P. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism." *American Economic Review*, 98(3): 843-63.

Banzhaf, H. Spencer and V. Kerry Smith. 2007. "Meta-Analysis in Model Implementation: Choice Sets and the Valuation of Air Quality Improvements." *Journal of Applied Econometrics*, 22: 1013-1031.

Baranzini, Andrea, Jose Ramirez, Caroline Schaerer and Phillippe Thalman (eds). 2008. *Hedonic Methods in Housing Markets: Pricing Environmental Amenities and Segregation*. Springer: New York, NY.

Bateman, Ian J., Richard T. Carson, Brett Day, W. Michael Hanemann, Nic Hanley, Tanis Hett, Michael Lee Jones, Graham Loomes, Susan Mourato, Ece Ozdemiriglu and David W. Pearce. 2002. *Economic Valuation with Stated preference Techniques: A Manual*. Edward Elgar: Cheltenham, UK.

Bateman, Ian J., and Andrew P. Jones. 2003. "Contrasting Conventional with Multi-level Modeling Approaches to Meta-analysis: Expectation Consistency in UK Woodland Recreation Values." *Land Economics*, Vol. 79 (2): 235-258.

Bennett, Jeff, and Russell Blamey. 2001. *The Choice Modeling Approach to Environmental Valuation*. Edward Elgar: Cheltenham, UK.

Berry, Steven and Ariel Pakes. 2007. "The Pure Characteristics Demand Model." *International Economic Review*, 48(4): 1193-225.

Bishop, Richard C. 1978. "Endangered Species and Uncertainty: The Economics of a Safe Minimum Standard." *American Journal of Agricultural Economics*, Vol. 60 (1): 10-18.

Boyle, Kevin J. 2003. "Contingent Valuation in Practice." In *A Primer on Nonmarket Valuation*, P.A. Champ, K.J. Boyle and T.C. Brown (eds). Kluwer Academic Publishers: Dordrecht.

Boyle, Kevin J., and John C. Bergstrom. 1992. "Benefit Transfer Studies: Myths, Pragmatism, and Idealism." *Water Resources Research*, Vol. 28 (3): 657-663.

Boyle, Kevin J., Nicolai V. Kuminoff, Christopher F. Parmeter, and Jaren C. Pope. 2009. "Necessary Conditions for Valid Benefit Transfers." *American Journal of Agricultural Economics*, Vol. 91 (5): 1328-1334.

Brouwer, Roy. 2000. "Environmental Value Transfer: State of the Art and Future Prospects." *Ecological Economics*, Vol. 32 (1): 137-152.

Brouwer, Roy, and Frank A Spaninks. 1999. "The Validity of Environmental Benefits Transfer: Further Empirical Testing." *Environmental and Resource Economics*, Vol. 14 (1): 95-117.

Carmines, Edward G. and Richard A. Zellner. 1979. *Reliability and Validity Assessment*. Sage Publications: Beverly Hills, CA.

Carson, Richard T., W. Michael Hanemann, Raymond Kopp, John A. Krosnick, Richard C. Mitchell, Stanley Presser, Paul A. Rudd, V. Kerry Smith, Michael Conaway and Kerry Martin. 1997. "Temporal Reliability of Estimates from Contingent Valuation." *Land Economics*, Vol. 73 (2): 151-163.

Champ, Patricia A., Kevin J. Boyle and Thomas C. Brown. 2003. *A Primer on Nonmarket Valuation*. Kluwer Academic Publishers: Dordrecht.

Chetty, Raj, Adam Looney and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." American Economic Review, Vol. 99 (4); 1145-1177.

Colombo, Sergio, Javier Calatrava-Requena and Nick Hanley. 2007. "Testing Choice Experiment for Benefit Transfer with Preference Heterogeneity." *American Journal of Agricultural Economics*, Vol. 89 (1): 135–151.

Cummings, Ronald G., David S. Brookshire, and William D. Schulze. 1986. *Valuing Environmental Goods: An assessment of the contingent valuation method*. Rowman and Littlefield Publishers, Inc.

Davis, Lucas. 2004. "The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster." *American Economic Review*, 94(5): 1693-704.

Desvousges, William H., Michael C. Naughton, and George R. Parsons. 1992. "Benefits Transfer: Conceptual Problems in Estimating Water Quality Benefits Using Existing Studies." *Water Resources Research*, Vol. 28 (3): 675-683.

Downing, Mark, and Teofilo Ozuna. 1996 "Testing the Reliability of the Benefit Function Transfer Approach." *Journal of Environmental Economics and Management*, Vol. 30 (3): 316-322.

Eade, Jeremy D. O., and Dominic Moran. 1996. "Spatial Economic Valuation: Benefits Transfer Using Geographical Information Systems." *Journal of Environmental Management*, Vol. 48 (2): 97-110.

Fisher, Jefferey D., Mike E. Miles and R. Brian Webb. 1999. "How Reliable Are Commercial Appraisals? Another Look." *Real Estate Finance*, Vol. 16 (3): 9-15.

Harberger, Arnold C. 1964. "The Measurement of Waste." *American Economic Review*, Vol. 54 (3): 58-76.

Harberger, Arnold C. 1971. "Three Basic Postulates for Applied Welfare Analysis. *Journal of Economic Literature*, Vol. 9 (3): 785-797.

Hausman, Jerry A. 1981. "Exact Consumer Surplus and Deadweight Loss." *American Economic Review*, Vol. 71 (4): 662-676.

Herriges, Joseph A., and Catherine L. Kling (eds.). 1999. *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*. Edward Elgar: Cheltenham, UK.

Hines, James R. 19099. "Three Sides of Harberger Triangles." *The Journal of Economic Perspectives*, Vol. 13 (2): 167-188.

Holmes, Thomas P., and Wicktor L. Adamowicz. 2003. "Attribute Based Methods." In *A Primer on Nonmarket Valuation*, P.A. Champ, K.J. Boyle and T.C. Brown (eds). Kluwer Academic Publishers: Dordrecht.

Johnston, Robert J. 2007. "Choice Experiments, Site Similarity and Benefits Transfer." *Environmental and Resource Economics*, Vol. 38 (3): 331-351.

Johnston, Robert J., Elena Y. Besedin, Richard Iovanna, Christopher J. Miller, Ryan F. Wardwell and Matthew H. Ranson. 2005. "Systematic Variation in Willingness to Pay for Aquatic Resource Improvements and Implications for Benefit Transfer: A Meta-Analysis." *Canadian Journal of Agricultural Economics,* Vol. 53 (2-3): 221-248.

Kirchhoff, Stefanie, Bonnie G. Colby and Jeffrey T. LaFrance. 1997. "Evaluating the Performance of Benefit Transfer: An empirical inquiry." *Journal of Environmental Economics and Management*, Vol. 33 (1): 75-93.

Kristofersson, Dadi, and Ståle Navrud. 2005. "Validity Tests of Benefit Transfer – Are We Performing the Wrong Tests?" *Environmental and Resource Economics*, Vol. 30 (3): 279-286.

Kuminoff, Nicolai V. 2009. "Decomposing the Structural Identification of Nonmarket Values." *Journal of Environmental Economics and Management*, Vol.57 (2): 123-39.

Leon-Gonzalez, Roberto, and Riccardo Scarpa. 2008. "Improving Multi-site Benefit Functions Via Bayesian Model Averaging: A New Approach to Benefit Transfer." *Journal of Environmental Economics and Management*, Vol. 56 (1): 50-68.

Lindhjem, Henrik, and Stale Navrud. 2008. "How Reliable Are Meta-Analyses for International Benefit Transfers?" *Ecological Economics*, Vol. 66 (2-3): 425-435.

Loomis, John B. 1989. "Test-Retest Reliability of the Contingent Valuation Method: A Comparison of General Population and Visitor Responses." *American Journal of Agricultural Economics*, 71 (1): 76-84.

Loomis, John B. 1992. "The Evolution of a More Rigorous Approach to Benefit Transfer: Benefit Function Transfer." *Water Resources Research*, Vol. 28 (3): 701-705.

Luken, Ralph A., F. Reed Johnson, and Virginia Kibler. 1992. "Benefits and Costs of Pulp and Paper Effluent Controls Under the Clean Water Act." *Water Resources Research*, Vol. 28 (3): 665-674.

Mäler, Karl-Göran, and Jeffrey R. Vincent. 2003. Handbook of Environmental Economics: Valuing Environmental Changes (Vol.2). North-Holland: Amsterdam, The Netherlands.

Manski, C. F. 1990. "Nonparametric Bounds on Treatment Effects." *The American Economic Review*, Vol. 80, No. 2 :319-323.

McConnell, Kenneth E. 1992. "Model Building and Judgment: Implications for Benefit Transfers with Travel Cost Models." *Water* Resources Research, Vol. 28 (3): 695-700.

Moeltner, Klaus, Kevin J. Boyle and Robert W. Paterson. 2007. "Meta-analysis and benefit transfer for resource valuation-addressing classical challenges with Bayesian modeling." *Journal of Environmental Economics and Management*, Vol. 53 (2): 250-269.

Morrison, Mark, Jeff Bennett, Russell Blamey and Jordan Louviere. 2002. "Choice Modeling and Tests of Benefit Transfer." *American Journal of Agricultural Economics*, Vol. 84 (1): 161-170.

Mrozek, Janusz R., and Laura O. Taylor. 2002. "What Determines the Value of Life? A meta-analysis." *Journal of Policy Analysis and Management*, Vol. 21 (2): 253-270.

Murphy, James J., P. Geoffrey Allen, Thomas H. Stevens and Darryl Weatherhead. 2005. "A Meta-Analysis of Hypothetical Bias in Stated Preference Valuation." *Environmental and Resource Economics*, Vol. 30 (3):313-325.

Muthke, Thilo, and Karin Holm-Mueller. 2004. "National and International Benefit Transfer Testing with a Rigorous Test Procedure." *Environmental and Resource Economics*, Vol. 29 (3) : 323-336.

Navrud, Ståle, and Richard Ready (eds.). 2007. *Environmental Value Transfer: Issues and Methods*. Springer: Dordrecht, The Netherlands.

Nelson, Jon P., and Peter E. Kennedy. 2009. "The Use (and Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assessment." *Environmental and Resource Economics*, Vol. 42 (3): 345-377.

Parsons, George R. 2003. "The Travel Cost Model." In *A Primer on Nonmarket Valuation*, P.A. Champ, K.J. Boyle and T.C. Brown (eds). Kluwer Academic Publishers: Dordrecht.

Parsons, George R. and Mary Jo Kealy. 1994. "Benefit Transfer in a Random Utility Model of Recreation." *Water Resources Research*, Vol. 30 (8): 2477-84.

Rhode, Paul W. and Koleman S. Strumpf. 2003. "Assessing the Importance of Tiebout Sorting: Local Heterogeneity from 1850 to 1990." *American Economic Review*, 93(5): 1648-77.

Rolfe, John, and Jeff Bennett. 2006. *Choice Modeling and the Transfer of Environmental Values*. Edward Elgar: Northampton, MA.

Rosenberger, Randall S., and Robert J. Johnston. 2009. "Selection Effects in Meta-Analysis and Benefit Transfer: Avoiding Unintended Consequences." *Land Economics*, Vol. 85 (3): 410-428.

Rosenberger, Randall S., and John B. Loomis. 2000. "Using Meta-Analysis for Benefit Transfer: In-Sample Convergent Validity Tests of an Outdoor Recreation Database." *Water Resources Research*, Vol. 36 (4): 1097-1107.

Shrestha, Ram K., and John B. Loomis. 2001. "Testing a Meta-analysis Model for Benefit Transfer in International Outdoor Recreation." *Ecological Economics*, Vol. 39 (1): 67-83.

Smith, V. Kerry. 1992. "On Separating Defensible Benefits From 'Smoke and Mirrors'." *Water Resources Research*, Vol. 28 (3): 685-694.

Smith, V. Kerry, and Yoshiaki Kaoru. 1990. "Signals or Noise? Explaining the Variation in Recreation Benefit Estimates." *American Journal of Agricultural Economics*, Vol. 72 (2): 419-433.

Smith, V. Kerry, and Subhrendu K. Pattanayak. 2002. "Is Meta-Analysis a Noah's Ark for Non-Market Valuation?" *Environmental and Resource Economics*, Vol. 22 (1-2): 271-296.

Smith, V. Kerry, Subhrendu K. Pattanayak and G. Van Houtven. 2006. "Structural Benefits Transfer: An Example Using VSL Estimation." *Ecological Economics*, Vol. 60 (2): 361-371.

Smith, V. Kerry, George Van Houtven, and Subhrendu K. Pattanayak. 2000. "Benefit Transfer via Preference Calibration: 'Prudential Algebra' for Policy." *Land Economics*, Vol. 78 (1): 132-152.

Spash, Clive L., and Arild Vatn. 2006. "Transferring Environmental Value Estimates: Issues and Alternatives." *Ecological Economics*, Vol. 60 (2): 379-388.

Stanley T.D. 2001. "Wheat from Chaff: Meta-Analysis as Quantitative Literature Review." *Journal of Economic Perspectives*, Vol. 15(3): 131-150.

Stapler, Ryan W., and Robert J. Johnston. 2009. "Meta-Analysis, Benefit Transfer, and Methodological Covariates: Implications for Transfer Error." *Environmental and Resource Economics*, Vol. 42 (2): 227-246.

Taylor, Laura O. 2003. "The Hedonic Method." In *A Primer on Nonmarket Valuation*, P.A. Champ, K.J. Boyle and T.C. Brown (eds). Kluwer Academic Publishers: Dordrecht.

Teisl, Mario F., Kevin J. Boyle, Daniel W. McCollum and Stephen D. Reiling. 1995. "Test-Retest Reliability of Contingent Valuation with Independent Sample Pretest and Posttest Control Groups." *American Journal of Agricultural Economics*, Vol. 77 (3): 613-619.

Troy, Austin, and Matthew A. Wilson. 2006. "Mapping Ecosystem Services: Practical Challenges and Opportunities in Linking GIS and Value Transfer." *Ecological Economics*, Vol. 60 (2): 435-449.

U.S. Department of Agriculture (USDA). 2005. *Conservation Security Program: Amendment to the Interim Final Rule Benefit Cost Assessment*. Washington DC, March.

U.S. Environmental Protection Agency. 1999. *The Benefits and Costs of the Clean Air Act 1990 to 2010*. EPA 410-R-99-001.

U.S. Environmental Protection Agency. 2000. *Guidelines for Preparing Economic Analyses*. EPA 240-R-00-003.

U.S. Environmental Protection Agency. 2005. *Regulatory Impact Analysis for the Final Clean Air Visibility Rule of the Guidelines for Best Available Retrofit technology (BART) Determinants Under the Regional Haze Regulations.* EPA-452/R-050-004.

U.S., Executive Office of the President. 2009. *Economic Report of the President.* U.S. Government Printing Office, ISVB-13: 978-0-16-077918-3.

U.S., Executive Office of the President. 1981. "Executive Order 12291: Federal Regulation." *Federal Register*, 46 FR 13193, 3 CFR (http://www.archives.gov/federal-register/codification/executive-order/12291.html, accessed September 14, 2009).

U.S., Executive Office of the President. 1993. "Executive Order 12866: Regulatory Planning and Review." *Federal Register*, Vol. 58 (160) (http://www.archives.gov/federal-register/executive-orders/pdf/12866.pdf, accessed January 26, 2010).

U.S. Office of Management and Budget. 2003. *Circular A-4: Regulatory Analysis.* (http://www.whitehouse.gov/OMB/circulars/a004/a-4.pdf, accessed October 21, 2009).

Vandenberg, Timothy P., Gregory L. Poe, and John R. Powell. 2001. "Assessing the Accuracy of Benefits Transfers: Evidence from a Multi-site Contingent Valuation Study of Ground Water Quality." *The Economic Value of Water Quality*. J.C. Bergstrom, K.J. Boyle and G.L. Poe (eds.). Edward Elgar: Cheltenham, UK.

Walsh, Richard G., Donn M. Johnson, and John R. McKean. 1992. "Benefit Transfer of Outdoor Recreation Demand Studies, 1968-1988." Water Resources Research, Vol. 28 (3): 707-713.

Whitehead, John C., and Timothy J. Hoban. 1999. "Testing for Temporal Reliability in Contingent Valuation with Time for Changes in Factors Affecting Demand." *Land Economics*, Vol. 75 (3): 453-465.

Wilson, Matthew A., and John P. Hoehn. 2006. "Valuing Environmental Goods and Services Using Benefit Transfer: The State-of-the Art and Science." *Ecological Economics*, Vol. 60 (2): 335-342.

Zanderson, Marianne, Mette Termansen, and Frank S. Jensen. 2007. "Testing Benefits Transfer of Forest Recreation Values Over a Twenty-Year Time Horizon." *Land Economics*, 83 (3): 412-40.

Table 1. Benefit Transfer Convergent Validity Studies (listed in chronological order)

Authors	Valuation Method	Location	Application	Prediction Errors[a,b] (absolute values)
Downing and Ozuna (1996)	Contingent Valuation	Texas	Saltwater fishing	15-164% Mean=50% (n=16)
Kirchhoff, Colby and LaFrance (1997)	Contingent valuation	Arizona & New Mexico	Day hikes and white-water boating	2-87% Mean=33% (n=12)
Brouwer and Spaninks (1999)	Contingent valuation	Netherlands	Peat meadow land	22-60% Mean=39% (n=12)
Morrison et al. (2002)	Choice modeling (Attribute-based valuation)	Australia	Wetlands	4-65% Mean=32% (n=9)

[a] Percent errors not reported by Downing and Ozuna and Morrison et al. Percent errors are computed as (((study site value – policy site value)/Policy site value)x100).

[b] Kirchoff, Colby and LaFrance computed prediction errors as (((study site #1 value – study site #2 value)/ study site #2 value)x100)and then flipped the order and computed (((study site #2 value – study site #1 value)/ study site #1 value)x100). Only data from the first set of error calculations are reported here.