**Context Similarity and the Validity of Benefits Transfer:
Is the Common Wisdom Correct?**

Robert J. Johnston
Department of Agricultural and Resource Economics
University of Connecticut

April 6, 2006

Contact Information:

Dr. Robert J. Johnston
Connecticut Sea Grant Office
University of Connecticut at Avery Point
1080 Shennecossett Rd.
Groton, CT 06340-6048
(860) 405-9278 (office) / (860) 405-9109 (fax) / robert.johnston@uconn.edu

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Long Beach, California, July 23-26, 2006

 This research was funded by the USDA Fund for Rural America #97-36200-5219. Opinions belong solely to the author and do not imply endorsement by the funding agency. This research has not been previously presented at any other scholarly meeting.

Abstract

Choice experiments are designed to account for variations in environmental resources and site characteristics, as well as potential implications of these variations for willingness to pay. This may render choice experiment results highly suitable for benefits transfer. It is unclear, however, whether the flexibility of choice experiments renders the similarity of study and transfer sites less critical for transfer validity. Drawing from identical choice experiments conducted in different Rhode Island communities, this model assesses the extent to which error in function-based benefits transfer is related to the similarity of communities across a variety of observable dimensions. Results suggest that site similarity, at least across some dimensions, influences the validity of choice experiment benefits transfers. However, the use of some measures of similarity as indicators of transfer error may provide misleading results.

Introduction

Generalization error may be defined as the error that occurs when benefit estimates from a study site (or combination of sites) are used or adapted to forecast benefits at a policy or transfer site; it is the difference between the transferred and actual, generally unknown, value (Rosenberger and Stanley 2005). The likelihood and magnitude of such errors are critical to both the validity and accuracy of benefits transfer (Jiang et al. 2005). As a general rule, generalization errors are typically assumed to be smaller if the sites or contexts over which transfers occur are more similar (Boyle and Bergstrom 1992; Desvousges et al., 1992; Kask and Shogren 1994; Rosenberger and Loomis 2001). Indeed, similarity policy contexts is often considered a fundamental prerequisite of benefits transfer (Jiang et al. 2005). Transfers conducted over dissimilar sites or contexts—even if addressing willingness to pay (WTP) for otherwise similar resources—are often treated with skepticism (e.g., Bergstrom and De Civita 1999; Rosenberger and Loomis 2003). This skepticism persists despite the ability of some transfer methods (e.g., meta-analysis; function-based transfer) to adjust for attributes of the valuation context, and for some valuation models (e.g., choice experiments) to adjust for multi-attribute distinctions between resources and/or valuation contexts (Rosenberger and Loomis 2003; Jiang et al. 2005; Johnston et al. 2005).

A significant literature demonstrates the importance of resource and context similarity for transfer or generalization error (e.g., Loomis 1992; VandenBerg et al. 2001; Piper and Martin 2001; Rosenberger and Loomis 2001). Transfers reported in this work, however, rely almost exclusively on methodologies unable to adjust for differences between transfer and policy sites (e.g., contingent valuation, travel cost methods). In contrast, only a small number of studies (e.g., Morrison et al. 2002, 2004; Jiang and Swallow 2005) have assessed the performance of benefits transfer using choice experiments—a methodology designed to account for variations in environmental resources and site characteristics, as well as potential implications of these variations for willingness to pay.

The ability of choice experiments to account for inter-site variations may render the resulting benefit functions highly suitable for transfer. It is unclear, however, whether this flexibility of choice experiments renders the similarity of study and transfer sites less critical for transfer validity. Existing assessments of choice experiment benefit transfer provide limited insight into the importance of site similarity. For example, while Morrison et al. (2002) show that between site transfers perform better than between population transfers, these results provide little evidence regarding the potential role of site or context similarity on the validity of these transfers. Similarly, Jiang et al.'s (2005) analysis considers only two potential sites (Massachusetts and Rhode Island coastal areas), and hence cannot assess potential implications of varying site similarity on transfer error. In sum, the limited literature addressing choice experiment benefits transfer offers little systematic evidence indicating the extent to which site similarity can influence the extent of transfer error.

This paper seeks to fill this gap in the literature. Drawing from identical choice experiments conducted in distinct Rhode Island communities, the model assesses the extent to which generalization error in function-based benefits transfer is related to the similarity of communities across a variety of observable dimensions. Of particular emphasis is similarity with regard to attributes that might reflect the relative availability of substitutes or complements in different communities—here denoted " policy context similarity"—a feature of particular focus in the literature (Bergstrom and De Civita 1999). The model also distinguishes between context similarity and other forms of likeness, including spatial or geographical proximity. Results suggest that context similarity, at least across some indicators, does influence the validity of choice experiment benefits transfer. However, reliance on other measures of likeness as an indicator of the potential performance of benefits transfer (e.g., geographical proximity) may provide misleading results.

Choice Experiments and Benefits Transfer

The suitability of choice experiments for benefits transfer is discussed by both Morrison et al. (2002, 2004) and Jiang and Swallow (2005); these discussions are only summarized here. Choice experiments ask respondents to evaluate alternative goods or programs (often including a do-nothing option, or status quo) that may differ across a variety of attributes, and choose the option that offers the greatest satisfaction or utility (Adamowicz et al. 1998). The framework forces respondents to acknowledge and react to tradeoffs among attributes, including money cost. As a result, unlike contingent valuation—which typically estimates values for a single or very small number of policy or good configurations—choice experiments generate an empirical (econometric) estimate of a valuation or utility function. This function typically allows analysts to estimate utility theoretic values for a wide range of policy or environmental good outcomes, and assess how these values change when policy configurations are altered. This property of choice experiments renders them highly suitable for benefits transfer, at least in theory. Simply put, the ability of choice experiments to explicitly adjust for differences in the attributes of environmental goods or policies provides an increased capacity to adjust for differences between study and policy sites—thereby improving the potential accuracy of benefits transfer (Morrison et al. 2002; Jiang et al. 2005).

As a practical matter, however, choice experiments cannot account for all possible attributes that might distinguish study and policy contexts. Assessments of choice experiment benefits transfer generally suggest improved performance compared to fixed-value transfers

(Morrison et al. 2002; Jiang et al. 2005). However, statistically significant transfer errors nonetheless remain common. Moreover, unlike transfers conducted using simpler methods (contingent valuation, travel cost), the existing literature provides little information regarding the conditions under which choice experiment benefit transfers are likely to generate small or large transfer errors. This begs the question as to whether the performance of choice experiment-based benefit transfers are improved in cases in which the study and transfer sites are highly similar.

Methods and Conceptual Approach

The data are drawn from the *Rhode Island Rural Land Use Survey*, a choice experiment survey designed to assess rural residents' tradeoffs among attributes of residential development and conservation (Johnston et al. 2003b). Respondents from four Rhode Island rural communities (Burrillville, Exeter, West Greenwich, and Coventry) were asked to consider alternative, multiattribute development options for hypothetical tracts of forested land located in their local town. Attributes of choice options characterized land use features and amenities identified as important by focus groups and interviews with growth management practitioners.

The four sampled communities were chosen to provide exemplars of contexts with varying degrees of likeness. Table one illustrates demographic and development characteristics of the four communities. Two of the communities (Exeter and West Greenwich) are neighboring communities that are highly similar over a wide range of attributes that might be considered relevant to choices concerning land development and conservation. Indeed, many in the area consider these communities to be virtually identical, distinguished only by lines on a map. The population density of both communities is nearly identical (~100 persons per square mile), with similar numbers of housing units and mean family incomes.

A third community (Coventry) borders West Greenwich to the north, and hence is geographically proximate to the first two communities. However, the population and development characteristics of this community differ markedly from both Exeter and West Greenwich (table 1). The population is density in Coventry, for example, is nearly six times greater than that in West Greenwich—a difference that may be highly relevant for residents' land use preferences. The fourth community (Burrillville) is located in the far northwestern corner of Rhode Island, and is hence geographically separated from the first three communities. However, with regard to population and development attributes, it provides a middle-ground between the rural communities of Exeter and West Greenwich and the more developed town of Coventry.

The selection of communities allows for an assessment of two distinct types of similarity that often remain undistinguished in the literature. The first is *geographical proximity* of the study and policy site across geographical space. The second is *context similarity*, as characterized above. In the present case, the choice experiment survey addresses development and conservation issues. Hence, as an indicator of context similarity one might consider the likeness of communities with respect to potential substitutes and complements offered by land use policies (e.g., indicators of development and land use). The four communities provide examples of both: communities that are geographically similar yet divergent in land use indicators, and communities that are geographically distant yet more similar in terms of land use.

Methods follow the general approaches of Morrison et al. (2002, 2004) and Jiang et al. (2005), with the exception that neither prior work addresses the role of site similarity on the performance of choice experiment benefits transfer. The data are systematically split such that individual choice models may be estimated for each of the four communities, together with set of pooled models that imposes identical preference structures across pairs of different communities. Given four distinct communities, this provides six different pairwise contrasts, allowing tests of transfer errors across each pair. A contrast of these six models provides information necessary to assess the role of site similarity on the validity of function-based benefits transfer between communities, as well as to identify those aspects of similarity that appear most relevant.

For example, if geographical proximity is an important indicator of potential transfer error, one would expect to see substantial and statistically significant transfer errors involving the community of Burrillville, a community spatially distant from the other three contiguous communities. If, in contrast, context similarity is a more important indicator of transfer validity, one might expect to see more substantial transfer errors between highly developed Coventry and its much less developed neighbors—Exeter and West Greenwich—with greater similarity between Coventry and Burrillville. Clear differences between values estimated in Coventry and Burrillville, and between these two communities and Exeter/West Greenwich, would suggest that both geographical and context similarity are relevant. Finally, significant differences between results for Exeter and West Greenwich would suggest that transfers are likely to involve substantial error even between sites that are nearly identical.

Choice experiment models are estimated based on the familiar random utility framework. Hypothesis tests address differences in estimated preference functions (i.e., estimated coefficients), scale parameters (Swaite and Louviere 1993), and marginal willingness to pay for policy attributes. We also quantify correlations among selected indicators of site similarity and the magnitude of error in marginal WTP transfer.

The Random Utility Model

Survey responses are analyzed using a standard random utility model, which allows welfare measures to be derived from choice experiment data. To model a respondent's choice, we define a utility function that includes attributes of a rural development or conservation plan and the net cost of the plan to the respondent (Hanemann 1984; McConnell 1990):

(1) $$U(\cdot) = U(\mathbf{X}_c, Y\text{-}F_c) = v(\mathbf{X}_c, Y\text{-}F_c) + \varepsilon_c$$

where

$\mathbf{X}_c$ = a vector of variables describing attributes of development or conservation plan c;

Y = disposable income of the respondent.

F_c = the change in mandatory taxes paid by the respondent under plan c;

$v(\cdot)$ = a function representing the empirically measurable component of utility;

ε_c = econometric error.

If one compares Plan A ($c{=}A$) to Plan B ($c{=}B$), the change in utility (dU) may be modeled as

(2) $$dU = U(\mathbf{X}_A, Y\text{-}F_A) - U(\mathbf{X}_B, Y\text{-}F_B) = [v(\mathbf{X}_A, Y\text{-}F_A) - v(\mathbf{X}_B, Y\text{-}F_B)] - [\varepsilon_B\text{-}\varepsilon_A]$$
$$= dv - \theta$$

The model assumes a respondent assesses the difference between utility under the two plans and indicates the sign of dU by either choosing Plan A ($dU>0$) or Plan B ($dU<0$). If θ is assumed to have a logistic distribution then the familiar logit model applies, in which the probability of selecting a given option is a logistic function of the utility difference dv (Maddala 1983). As prior attempts at estimating random parameters (mixed logit) specifications for these data failed to converge (Johnston et al. 2003), we illustrate results for standard logit estimation.

Although the literature offers no firm guidance regarding the choice of specific functional forms for dv, in practice linear forms are often used. Hence,

(3) $$dv = v(\mathbf{X}_A, F_A)\text{-}v(\mathbf{X}_B, F_B) = \beta_x(\mathbf{X}_A\text{-}\mathbf{X}_B) + \beta_f(F_B\text{-}F_A),$$

where β_x is a conforming vector of coefficients associated with the vector of attribute differences $(\mathbf{X}_A\text{-}\mathbf{X}_B)$ and β_f as a scalar coefficient associated with the tax difference $(F_B\text{-}F_A)$. The parameter vector β_x may be interpreted as the marginal utility of development or conservation attributes, while β_f represents the marginal utility of income.[2]

Six models are estimated—one for each possible pair of communities. Each model pools data from relevant community pair, but allows systematic variations in slope and intercept coefficients. Formally, this approach redefines dv in (3) to provide a separate utility estimate for respondents in each community. We define a binary variable D_j to equal one for respondents from community j, and to equal zero for the second community $i \neq j$. We then estimate a simple extension of (3) allowing for systematically varying slopes,

$$dv = \beta_x(\mathbf{X}_A - \mathbf{X}_B) + \beta_f(F_B - F_A) + \beta_{xj}D_j(\mathbf{X}_A - \mathbf{X}_B) + \beta_{fj}D_j(F_B - F_A) \quad (4)$$

where β_x and β_f represent marginal utility parameters for respondents from community i, and the conforming sums $(\beta_x + \beta_{xj})$ and $(\beta_f + \beta_{fj})$ represent marginal utilities for respondents from community j. To test for the equivalence of model parameters using a pooled model such as (4), one must also account for the potential confounding effect of the scale parameter (or heteroskedasticity in the residual variance across communities) on coefficient estimates. This may be accomplished using established methods (e.g., Allison 1999; Swait and Louviere 1993).

The Survey

As noted above, the *Rhode Island Rural Land Use Survey* was designed to assess rural residents' tradeoffs among attributes of residential development and conservation (Johnston et al. 2003b). Survey development required approximately eighteen months, and involved background research, interviews with policy makers and stakeholders, and a large number of focus groups. Individual and group pretests ensured that survey language and format could be easily understood by respondents, and that respondents shared consistent interpretations of survey scenarios (cf. Johnston et al. 1995).

Each choice experiment scenario presented respondents with two development options, a current development plan (CDP) and an alternate development plan (ADP), where each plan could differ across a set of land use (development and conservation) attributes. Attributes distinguishing management plans were chosen based on focus groups and interviews, and characterized such features as protected open space, residential development, unprotected undeveloped land, scenic views, wildlife habitat, and household cost (taxes). Table 2 characterizes attributes distinguishing hypothetical management plans.

Prior to presenting respondents with development choices, the survey provided background information on community land use and tradeoffs implicit in development choices. Contingent choice instructions and questions were then presented. Each respondent considered three potential pairs of current and alternate plans for the same 400 acre undeveloped site. Respondents were instructed to consider each pair independent of previous choices, and to assume that all choices applied to the same parcel. Respondents were told that "if you do not vote for either plan, development will automatically occur as shown by the current development plan," thereby specifying the status quo that would occur if no choice were made (Adamowicz et al. 1998). This framework was chosen to mimic actual community considerations of development proposals, wherein a landowner possesses the property rights necessary to permit development. However, officials may seek to influence the configuration of the development, delaying permits unless changes are made. As a result, officials may exert control over the ultimate form of development (Johnston et al. 2003b).

A fractional factorial design was used to construct survey questions with an orthogonal array of attribute levels.[1] Attributes were free to vary over their full range for both the current and alternate plans, with no imposed ordering of attribute levels between the two plans. This resulted in 128 unique contingent choice questions divided among 43 different survey booklets (three questions per booklet). Surveys were mailed to 4000 randomly selected residents of the four Rhode Island towns during March-May 2000 (1000 surveys per town), following the total survey design method (Dillman 2000). Of 3702 deliverable surveys, 2157 were returned, providing 6062 (94% of the potential 6471) complete and usable responses to dichotomous choice questions. The number of completed surveys per town ranged from 505 in Coventry to 580 in West Greenwich, with response rates varying from 53% to 61% across the four communities.

Assessments of Convergent Validity and Transfer Error

Discussion here emphasizes implications of model results for transfer error between community pairs. The focus on benefits transfer emphasizes a comparison of results *across* communities (convergent validity) rather than detailed individual results for *each* community. As a basis for initial comparison, however, table 3 presents individual results for each of the four communities, based on the random utility (logit) model outlined above. All models are statistically significant at $p<0.01$. In all cases, the substantial majority of variables are statistically significant, with the sign of significant variables matching prior intuition. Here, we emphasize only the general similarity of results across communities, subject to more rigorous subsequent testing. Primary model results, however, are not based on these independent models, but rather on models that pool community data pair-wise to test hypotheses relevant to benefits transfer. The six pair-wise pooled models used for convergent validity testing follow (4) above,

[1] The statistical design was conducted by Don Anderson of STATdesign, Inc.

and are suppressed for the sake of brevity.

Based on the pooled models for each community pair, a variety of tests relevant to the convergent validity of benefits transfer may be conducted, including tests of utility model parameters and implicit prices (marginal WTP) (Morrison et al. 2002). Following Jiang et al. (2005) and Morrison et al. (2002), we first test for differences in the overall utility structure (or estimated model parameters) across communities, where these parameters are tested in community pairs. This is followed by a test of implicit prices of selected attributes, or marginal WTP. This implies six sets of hypothesis tests associated with unique community pairs, for each of the two categories noted above (e.g., tests of model parameters; tests of marginal WTP).

As shown by Allison (1999) and Swait and Louviere (1993), the confounding effect of the scale parameter (or residual variance) on coefficient estimates requires that tests of the equivalence of utility function parameters across community pairs be decomposed into two parts. Specifying the vector of coefficient estimates (associated with the vector of model variables) as $\boldsymbol{\beta}$ and the scale parameter as μ, the first test is of hypothesis H_{1A}, that $\boldsymbol{\beta_i}=\boldsymbol{\beta_j}=\boldsymbol{\beta}$ for communities $i \neq j$, while allowing μ to vary across communities.[2] If we reject H_{1A}, then we conclude that the utility structure (e.g., population parameters of the utility function) differs across communities *i* and *j*. If we fail to reject H_{1A}, we continue to test hypothesis H_{1B}, that $\mu_i = \mu_j = \mu$. Rejection of H_{1B} also implies rejection of the equivalence of population parameters across the two groups.

Models required for the testing of H_{1A} and H_{1B} are estimated following Allison (1999). Table 4 summarizes results of the test of hypothesis H_{1A}. Results show only two instances in which we reject the null hypothesis (the Coventry-Exeter and Burrillville-Coventry pairs); in four out of six instances we fail to reject the null hypothesis that $\boldsymbol{\beta_i}=\boldsymbol{\beta_j}=\boldsymbol{\beta}$, if μ is free to vary across communities. Table 5 summarizes results of the test of hypothesis H_{1B}. Here Wald χ^2

---

[2] That is, $\boldsymbol{\beta} = [\boldsymbol{\beta_x} \ \ \beta_f]$ from (4).

tests (Allison 1999) universally fail to reject the null hypothesis of equal residual variances (scale parameters) across community pairs. Combining results for H_{1A} and H_{1B}, we fail to reject parameter equality for four out of the six possible community pairs, with the two rejections both involving the community of Coventry. This is an encouraging result for benefits transfer, and stands in contrast to results such as those of Jiang et al. (2005), which show wide scale rejection of parameter equality across groups (Swait and Louviere 1993).

The second set of tests involves the equivalence of implicit prices, or WTP for marginal changes in individual attributes. As noted by both Jiang et al. (2005) and Morrison et al. (2002), the implicit price for the k^{th} attribute, assuming a linear approximation for utility, is given by $-\beta_k / \beta_{\cos t}$, where β_k is the parameter on the k^{th} attribute, and β_{cost} is the parameter on the household cost of the program (i.e., the marginal utility of income). As above, results are drawn from models pooled pair-wise by communities that allow parameter estimates to vary systematically. A variety of statistical approaches are available for testing the convergent validity of marginal implicit prices; here we conduct hypothesis tests using standard Wald tests (Greene 2003, p. 487). For illustration, we test the equivalence of implicit prices for four attributes that are highly significant in all four community models (table 3) and are also a primary focus of community land use policy: open space isolated from developments (*iso_open*), open space adjacent to developments (*adj_open*), development size (*size_dif*), and development housing density (*dense_dif*) (table 2; cf. Johnston et al. 2003b).[3] Hypothesis tests are conducted for each community pair, resulting in 6 sets of pair-wise hypothesis tests and 24 tests. Results are shown in table 6, along with the point estimate magnitudes of marginal WTP

[3] In all cases, marginal WTP estimates for *iso_open* and *adj_open* are positive, and marginal WTP estimates for *size_dif* and *dense_dif* are negative. This is the expected result associated with positive preferences for preserved open space and negative preferences for developed acres.

differences. For comparison, associated percentage differences in marginal WTP point estimates are shown in table 7.

Hypothesis test results (table 6) mirror prior findings of Morrison et al. (2002), but stand in contrast to those of Jiang et al. (2005), finding a high degree of statistical correspondence in implicit prices across community pairs. Of 24 individual WTP differences, only six are shown to be statistically significant at $p<0.10$ or better. Hence, compared to some past assessments of benefits transfer (Bergstrom and DeCivita 1999; Rosenberger and Loomis 2001), results here are generally more supportive of the convergent validity of WTP across sites and the potential appropriateness of benefits transfer.

Percentage differences in marginal WTP point estimates for identical attributes, between different communities, vary from -101.33% to 67.97%. In absolute values, percentages vary from 7.04% to 101.33%, with an average absolute value of 39.08%. Across communities, the average (absolute value) transfer error in implicit prices ranges from 32.6% in pairs with Exeter, to 47.67% in pairs with Coventry (table 7). Error ranges such as this are typical for function based transfer (Rosenberger and Loomis 2003). This suggests that while marginal WTP values cannot, in the majority of cases, be shown to differ from a statistical perspective (table 6), the percentage differences in estimated marginal WTP can in some cases be substantial (table 7). Hence, results here validate prior suggestions that the appropriateness of benefits transfer, even between similar sites, may be largely a function of the size of error that policymakers judge to be acceptable (Shrestha and Loomis 2003).

As a final assessment of relationships between benefit transfer error and the similarity of communities, one might consider the correlation between differences in indicators of community land use and differences in marginal WTP for land use attributes. Here, housing density is used as an illustrative indicator of community similarity with regard to land use—although it is certainly not the only indicator that might be used for such purposes. Given data for community-wide housing density (table 1) and differences in marginal WTP for land use attributes (table 6), it is possible to calculate the correlation across community pairs between these two measures. Greater correlation coefficients (positive or negative) would indicate greater correlation between similarity in housing density and similarity in implicit prices for land use attributes—a result that would correspond to common intuition regarding transfer error.

Results of the correlation analysis are shown in table 8. Overall, results suggest a relatively high degree of correlation between measures of community similarity in land use (here, housing density) and differences in implicit prices, or transfer error, associated with land use attributes. Pearson correlation coefficients range in absolute value from 0.37 to 0.82, with an average absolute value of 0.68. For example, the correlation between the difference in housing density and the difference in marginal WTP for housing acreage (*size_dif*) is 0.81 across the six community pairs;[4] greater positive differences in housing density are associated with greater positive differences in marginal WTP. The correlation coefficient for open space isolated from developments (*iso_open*) is -0.82; greater positive differences in housing density are associated with greater negative differences in marginal WTP. This result is perhaps counter-intuitive in one sense, in that it suggests that communities that are more densely developed are less willing to pay for open space preservation. Nonetheless, it supports the intuition that community similarity in housing density is related to similarity in the implicit price of preserved open space.

Implications and Discussion—Site Similarity and Generalization Error

The most immediate implication from model results is the relatively high degree of convergent validity of choice experiment results across communities, at least from a statistical

[4] Recall, marginal WTP estimates for *size_dif* and *dense_dif* are negative; respondents are willing to pay to prevent increases in housing acres and density. Results must be interpreted accordingly.

perspective. For example, results suggest that transfers of marginal WTP for open space preservation (*iso_open*, *adj_open*) would be most often statistically appropriate, given the general lack of statistically significant differences in implicit prices for these attributes across communities. The convergent validity of utility parameters and implicit prices across communities is perhaps less surprising in light of prior findings, however, if one considers that many of the policy contexts (i.e., communities) considered are quite similar, at least by the standards of prior assessments of benefits transfer. Nonetheless, hypothesis test results are generally supportive of the use of choice experiment results from study communities to forecast marginal utilities and implicit prices in other Rhode Island rural communities (table 6), notwithstanding sometimes substantial point-estimate differences in marginal WTP (table 7).

The primary focus of this assessment, however, is on the importance of site similarity for transfer validity and accuracy. Here, the analysis support the hypothesis that transfers are more appropriate between communities that are more similar. The analysis also suggests that similarity in terms of the general policy environment (e.g., availability of substitutes and complements)—here the similarity of land use attributes—is more critical than geographical proximity.

For example, with respect to the equivalence of utility parameters, we fail to reject the null hypothesis of equal parameter estimates in four of the six community pairs. The two instances in which we reject the null hypothesis (i.e., find evidence of differences in the utility function) both involve the community of Coventry—a community distinguished by a much greater population and housing density than any of the other communities sampled (table 1). A third hypothesis test involving this community (Coventry-West Greenwich) narrowly misses rejection of the null hypothesis of parameter equality (p=0.12; table 4). In contrast, parameter estimates for Exeter and West Greenwich—neighboring communities that are very similar from a land use perspective—show a high degree of correspondence. Geographical distance alone, however, seems to play a lesser role, with utility parameters for the least spatially proximate community (Burrillville) differing only from those of Coventry. Of the three neighboring communities (Exeter, West Greenwich, and Coventry), parameter estimates for Coventry are more likely to differ from those of other communities—regardless of geographical proximity. In sum, similarity of the policy context (e.g., availability of substitutes and/or complements) seems to influence equivalence of utility parameters, while geographical proximity cannot be shown to have a clear impact.

Results are similar for implicit prices or marginal WTP. Four of the six statistically significant differences (out of 24 tested) involve the marginal WTP to avoid additional acres or density of housing developments in the more heavily-developed community of Coventry (i.e., the implicit prices of *size_dif* and *dense_dif*). Results suggest that Coventry residents are willing to pay less to prevent increases in housing acres and density. This result is intuitive if viewed from the perspective of the substantial amount of already existing housing stock. Compared to the other communities in the sample, marginal increases in housing acreage in Coventry are more trivial relative to existing housing acreage—and hence residents are willing to pay less to avoid such changes. In terms of percentage differences in implicit prices, Coventry again shows the evidence of larger transfer errors, with an average absolute value of a 47.67% error, compared to errors ranging from 32.60% to 38.31% in the other three communities (table 7). Geographical proximity alone, however, plays a less clear role, with average implicit price transfer errors involving the more distant community of Burrillville (37.85%) similar to those involving both Exeter (32.60%) and West Greenwich (38.21%).

Pearson correlation coefficients (table 8) further support the intuition that context similarity is related to the potential for transfer error. Correlation coefficients reveal relatively strong relationships between differences in housing density across community pairs and divergences in estimated benefit estimates associated with land use outcomes (marginal WTP for land use attributes). Such patterns again suggest that communities that are more similar in terms of housing density are also expected to have marginal WTP estimates (for land use attributes) that comport more closely. Relatively strong positive or negative correlations hold for implicit prices associated with housing acres (0.81), housing density (0.72) and open space acres isolated from open space (-0.82), with a much weaker correlation found for open space acres adjacent to developments (-0.37).

Statistical equivalence of implicit prices, however, cannot always be shown to be related to context similarity. For example, marginal WTP for open space preservation (*iso_open*, *adj_open*) for the most part cannot be shown to differ across communities, at least from a statistical perspective. This result is not universal, however, with results suggesting a statistically significant difference in marginal WTP for acres of open space adjacent to housing developments (*adj_open*) between Burrillville and West Greenwich residents (table 6). Despite sometimes substantial point estimate differences in marginal WTP, however, we cannot reject the equivalence of these implicit prices (for open space acres) across most community pairs. Such findings suggest that reliance on single indicators of site context similarity (here, using housing density) may not always provide appropriate guidance regarding the likely statistical equivalence of implicit prices across sites.

Conclusions

Model results suggest that common intuition regarding site similarity and transfer error is largely justified. Reliance on choice experiments for benefits transfer does not invalidate standard guidance regarding the importance of site similarity for transfer error. Here, generalization errors between communities are smaller and less likely to be statistically significant in cases where policy contexts—here proxied by housing density—are more similar. Residents of communities that are more similar across land use attributes are more likely to have similar WTP for land use policies, and hence WTP measures that may be transferred with greater confidence.

Results also suggest, however, that relationships between the context similarity and the validity of benefits transfer are more complex than is often assumed. Here, the importance of context similarity for transfer error appears to vary across attributes. In addition, while context similarity does influence the validity of function-based transfers, it is also possible to conduct statistically valid transfers, for some implicit prices, across sites that might be considered relatively dissimilar. Moreover, geographical proximity is not sufficient to justify benefits transfer when other attributes of the valuation context are not comparable; for communities in close proximity but highly dissimilar in terms of population and land use attributes, the model is more likely to reject the convergent validity of implicit prices.

Results, of course, must be viewed within the context of the relatively small sample of studies from which they are drawn. It should also be emphasized that these communities—all similar size communities in a single northeastern US state—are somewhat more similar overall then typically study and policy sites between which benefits transfer is conducted. Also, there are a number of vectors across which sites may differ, only a small number of which (e.g., housing density, geographical proximity) are addressed here. These caveats notwithstanding, model results suggest that common intuition regarding the role of site similarity in benefit transfer is for the most part appropriate, and holds for function-based, choice experiment benefit transfers as it does for simpler, fixed-value transfers.

References

Adamowicz, W., P. Boxall, M. Williams, and J. Louviere. 1998. Stated Preference Approaches for Measuring Passive Use Values: Choice Experiments and Contingent Valuation. *American Journal of Agricultural Economics* 80(1): 64-75.

Allison, P.D. 1999. Comparing Logit and Probit Coefficients Across Groups. *Sociological Methods and Research* 28(2): 186-208.

Bergstrom, J.C. and P. De Civita. 1999. Status of Benefits Transfer in the United States and Canada: A Review. *Canadian Journal of Agricultural Economics* 47: 79-87.

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

Desvousges, W.H., M.C. Naughton, and G.R. Parsons. 1992. Benefit transfer: Conceptual Problems in Estimating Water Quality Benefits Using Existing Studies. *Water Resources Research* 28(3): 675-683.

Dillman. D.A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. New York, NY: John Wiley and Sons.

Hanemann, W.M. 1984. Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses. *American Journal of Agricultural Economics* 66(3): 332-341.

Jiang, Y. S.K. Swallow and M. McGonagle. 2005. Context-Sensitive Benefit Transfer Using Stated Choice Models: Specification and Convergent Validity for Policy Analysis. *Environmental and Resource Economics* 31: 477–499.

Johnston, R.J., E.Y. Besedin, R. Iovanna, C. Miller, R. Wardwell, and M. 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* 53: 221-248.

Johnston, R.J., S.K. Swallow, D. Marie Bauer, and C.M. Anderson. 2003. Preferences for Residential Development Attributes and Support for the Policy Process: Implications for Management and Conservation of Rural Landscapes. *Agricultural and Resource Economics Review* 32(1): 65-82.

Johnston, R.J., S.K. Swallow, T.J. Tyrrell and D.M. Bauer. 2003b. Rural Amenity Values and Length of Residency. *American Journal of Agricultural Economics* 85: 1009-1024.

Johnston, R.J., T.F. Weaver, L.A. Smith, and S.K. Swallow. 1995. Contingent Valuation Focus Groups: Insights from Ethnographic Interview Techniques. *Agricultural and Resource Economics Review* 24 (1): 56-69.

Kask, S.B. and J.F. Shogren. 1994. Benefit Transfer Protocol for Long-Term Health Risk Valuation: A Case of Surface Water Contamination. *Water Resources Research* 30(10): 2813-2823.

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

McConnell, K.E. 1990. Models for Referendum Data: The Structure of Discrete Choice Models for Contingent Valuation. *Journal of Environmental Economics and Management* 19(1): 19-34.

Morrison, M. and J. Bennett. 2004. Valuing New South Wales Rivers for Use in Benefit Transfer. *The Australian Journal of Agricultural and Resource Economics* 48(4): 591-611.

Morrison, M., J. Bennett, R. Blamey, and J. Louviere. 2002. Choice Modeling and Tests of Benefit Transfer. *American Journal of Agricultural Economics* 84(1): 161-170.

Piper, S. and W. E. Martin. 2001. Evaluating the Accuracy of the Benefit Transfer Method: A Rural Water Supply Application in the USA. *Journal of Environmental Management* 63(3): 223–235.

Rosenberger, R.S. and Loomis, J.B., 2001. Benefit Transfer of Outdoor Recreation Use Values: A Technical Document Supporting the Forest Service Strategic Plan (2000 Revisions). General Technical Report RMRS-GTR-72. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 59p. (www.fs.fed.us/rm/pubs/rmrs_gtr72.html).

Rosenberger, R.S. and J.B. Loomis. 2003. Benefit Transfer. In *A Primer on Non-Market Valuation* edited by P.A. Champ, K.J. Boyle and T.C. Brown. Dordrecht, The Netherlands: Kluwer Academic Publishers (pp. 445-482).

Rosenberger, R.S. and T.D. Stanley. 2005. Measurement, Generalization and Publication: Sources of Error in Benefit Transfers and Their Management. Paper presented at *Benefit Transfer and Valuation Databases: Are We Heading in the Right Direction? An International Workshop Sponsored by US EPA and Environment Canada.* March 21-22. Washington DC.

Shrestha, R.K. and J.B. Loomis. 2003. Meta-Analytic Benefit Transfer of Outdoor Recreation Economic Values: Testing Out-of-Sample Convergent Validity. *Environmental and Resource Economics* 25: 79–100.

Swait, J. and J. J. Louviere. 1993. The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. *Journal of Marketing Research* 30(3): 305–314.

VandenBerg, T.P., G.L. Poe, and J.R. Powell, 2001. Assessing the Accuracy of Benefits Transfers: Evidence from a Multi-site Contingent Valuation Study of Groundwater Quality. In J.C. Bergstrom, K.J. Boyle, and G.L. Poe, eds., The Economic Value of Water Quality. Northampton, MA: Edward Elgar.

Table 1. Demographic and Land Use Indicators for Sampled Communities

	Burrillville	Coventry	Exeter	West Greenwich
Population	15,796	33,668	6,045	5,085
Population Density (persons/sq. mile)	284	566	105	100
Housing Units	5,821	13,059	2,196	1,809
Housing Density (units / sq. mile)	104.77	219.33	38.05	35.73
Mean Family Income	58,979	60,315	74,157	71,332

Source: US Census Data. All indicators are for the year 2000, the same year that the survey was conducted.

Table 2. Model Variables: Definitions and Summary Statistics

Variable Name	Description	Units and Measurement[a]	**Mean** (Std. Dev.)
adj_open	The difference between acres of open space adjacent to developments and roads in the CDP and ADP.	Acres in CDP minus acres in ADP. (Range: -200 to 200)	-3.41967 (95.091)
iso_open	The difference between acres of open space not adjacent to developments and roads in the CDP and ADP.	Acres in CDP minus acres in ADP. (Range: -200 to 200)	2.62028 (53.724)
size_dif	The difference between acres of residential development in the CDP and ADP.	Acres in CDP minus acres in ADP. (Range: -200 to 200)	-1.77646 (90.806)
dense_dif	The difference in housing density in the CDP and ADP.	Houses/acre in CDP minus houses/acre in ADP. (Range: -2 to 2)	-0.00666 (0.9759)
lg_mammal	Difference between habitat quality for large mammals in CDP and that in ADP.	Difference in wildlife habitat quality scale (1=worst; 5=best).	0.00370 (1.2193)
sm_mammal	Difference between habitat quality for small mammals in CDP and that in ADP.	Difference in wildlife habitat quality scale (1=worst; 5=best).	-0.01628 (1.2194)
com_bird	Difference between habitat quality for common birds in CDP and that in ADP.	Difference in wildlife habitat quality scale (1=worst; 5=best).	0.05107 (1.7511)
uncom_bird	Difference between habitat quality for uncommon birds in CDP and that in ADP.	Difference in wildlife habitat quality scale (1=worst; 5=best).	0.00370 (1.7038)
wet_sp	Difference between habitat quality for wetland species in CDP and that in ADP.	Difference in wildlife habitat quality scale (1=worst; 5=best).	-0.04663 (1.7359)
tax_dif	Difference in additional annual taxes and fees between CDP and ADP (resulting from management plan).	Dollars in CDP minus dollars in ADP. (Range: -$325 to $325)	-1.22132 (154.33)
lowvis	Difference between dummy variables indicating the presence of development either highly screened or not visible from the main road; in the CDP and ADP. Survey versions included eight different photographs characterizing different development visibility levels; four of these photographs are characterized as low visibility development.	Difference between dummy variables for CDP and ADP.	-0.00740 (0.6928)

[a] CDP=Current Development Plan; ADP=Alternate Development Plan.

Table 3. Choice Model (Logit) Results: Independent Community Models

Variable	Parameter Estimates (std. error)			
	Burrillville	Coventry	Exeter	West Greenwich
intercept	-0.0272 (0.0627)	0.0192 (0.0662)	-0.1600 (0.0641)**	-0.1266 (0.0599)**
dense_dif	-0.7253 (0.0705)***	-0.6806 (0.0747)***	-0.7404 (0.0722)***	-0.6260 (0.0658)***
size_dif	-0.0094 (0.0007)***	-0.0059 (0.0008)***	-0.0099 (0.0008)***	-0.0089 (0.0007)***
iso_open	0.0052 (0.0015)***	0.0031 (0.0014)**	0.0034 (0.0015)**	0.0046 (0.0013)***
adj_open	0.0028 (0.0008)***	0.0044 (0.0009)***	0.0051 (0.0009)***	0.0055 (0.0008)***
lowvis	0.0830 (0.0886)	0.2047 (0.920)**	0.1749 (0.0877)**	0.2075 (0.0834)**
lg_mammal	0.1337 (0.0524)**	0.0821 (0.0546)	0.1223 (0.0517)**	0.0827 (0.0492)*
sm_mammal	-0.0573 (0.0511)	0.0329 (0.0532)	-0.0664 (0.0511)	0.0263 (0.0483)
com_bird	0.0893 (0.0363)**	0.0681 (0.0379)*	0.1348 (0.0369)***	0.1082 (0.0350)***
uncom_bird	-0.0023 (0.0360)	0.0474 (0.0377)	0.0134 (0.0359)	0.0552 (0.0344)
wet_sp	0.0449 (0.0372)	0.0081 (0.0382)	0.0849 (0.0379)**	0.0476 (0.0353)
tax_dif	-0.0044 (0.0004)***	-0.0061 (0.0005)***	-0.0051 (0.0004)***	-0.0052 (0.0004)***
Likelihood Ratio χ^2	438.10 (p<0.0001)	393.20 (p<0.0001)	502.92 (p<0.0001)	505.37 (p<0.0001)
Obs (N)	1431	1297	1453	1593

* p<0.10
** p<0.05
*** p<0.01

Table 4. Hypothesis Test Results: Equivalence of Estimated Coefficients Allowing Residual Variance to Vary Across Community Samples (β_i=β_j=β)

	Burrillville	Coventry	Exeter	West Greenwich
Coventry	$\chi^2 = 27.11$ $p < 0.01$	--	--	--
Exeter	$\chi^2 = 6.42$ $p = 0.70$	$\chi^2 = 22.99$ $p < 0.01$	--	--
West Greenwich	$\chi^2 = 12.77$ $p = 0.17$	$\chi^2 = 14.05$ $p = 0.12$	$\chi^2 = 5.91$ $p = 0.75$	--

Table 5. Hypothesis Test Results: Equivalence of Residual Variance Across Community Samples ($\mu_i = \mu_j = \mu$)

	Burrillville	Coventry	Exeter	West Greenwich
Coventry	$\chi^2 = 0.12$ $p = 0.73$	--	--	--
Exeter	$\chi^2 = 1.24$ $p = 0.27$	$\chi^2 = 1.89$ $p = 0.17$	--	--
West Greenwich	$\chi^2 = 0.16$ $p = 0.69$	$\chi^2 = 0.37$ $p = 0.54$	$\chi^2 = 0.72$ $p = 0.39$	--

Table 6. Hypothesis Test Results for Implicit Price (Marginal WTP) Differences: Open Space and Development Acres[a,b]

	Burrillville	Coventry	Exeter	West Greenwich
Coventry				
iso_open	0.677 (χ^2=2.47)	--	--	--
adj_open	-0.073 (χ^2=0.09)	--	--	--
size_dif	-1.150 (χ^2=17.50)***	--	--	--
dense_dif	-53.05 (χ^2=4.91)**			
Exeter				
iso_open	0.514 (χ^2=1.23)	-0.163 (χ^2=0.18)	--	--
adj_open	-0.353 (χ^2=1.81)	-0.281 (χ^2=1.50)	--	--
size_dif	-0.177 (χ^2=0.32)	0.974 (χ^2=16.43)***	--	--
dense_dif	-19.48 (χ^2=0.56)	33.58 (χ^2=2.50)		
West Greenwich				
iso_open	0.299 (χ^2=0.45)	-0.379 (χ^2=1.10)	-0.216 (χ^2=0.29)	--
adj_open	-0.423 (χ^2=2.71)*	-0.350 (χ^2=2.48)	-0.070 (χ^2=0.08)	--
size_dif	-0.400 (χ^2=1.83)	0.750 (χ^2=11.63)***	-0.223 (χ^2=0.72)	--
dense_dif	-43.33 (χ^2=3.12)*	9.72 (χ^2=0.25)	-23.86 (χ^2=1.19)	

[a] Implicit price differences are denominated in dollars and are calculated as the implicit price for the community named in the row subtracted from the implicit price for the community named in the column, for each attribute. For example, results suggest that the implicit price of *iso_open* in Burrillville is $0.677 greater than that in Coventry. In all cases, implicit prices for *iso_open* and *adj_open* are positive, and for *size_dif* are negative.

[b] Numbers in parentheses are Wald χ^2 values with one degree of freedom (Greene 2003, p. p. 487) for the null hypothesis that the WTP difference is equal to zero.

* p<0.10
** p<0.05
*** p<0.01

Table 7. Percentage Differences (Transfer Errors) in Marginal WTP: Open Space and Development Acres[a]

	Burrillville	Coventry	Exeter	West Greenwich
Coventry				
iso_open	57.57%	--	--	--
adj_open	-11.47%	--	--	--
size_dif	54.50%	--	--	--
dense_dif	32.40%			
Exeter				
iso_open	43.71%	-32.73%	--	--
adj_open	-55.58%	-39.55%	--	--
size_dif	8.36%	-101.33%	--	--
dense_dif	11.89%	-30.33%		
West Greenwich				
iso_open	25.34%	-75.95%	-32.55%	--
adj_open	-66.47%	-49.36%	-7.04%	--
size_dif	18.94%	-78.04%	11.54%	--
dense_dif	67.97%	-8.78%	16.54%	
Absolute Value of Percentage Transfer Errors: Averages by Community[b]				
Burrillville	37.85%			
Coventry	47.67%			
Exeter	32.60%			
West Greenwich	38.21%			

[a] Percentage differences are calculated as the difference in implicit price for a given attribute between community pairs, divided by the baseline implicit price for the community in the column. For example, results (table 5) suggest that the implicit price of *iso_open* in Burrillville is $0.677 greater than that in Coventry. This represents a 57.57% increase compared to the baseline WTP of $1.18 per acre in Burrillville.

[b] Calculated as the average of the absolute value of all percentage differences in implicit prices (*iso_open*, *adj_open*, *size_dif*, *dense_dif*), between the noted community and all other communities.

Table 8. Correlations Across Community Pairs: Differences in Housing Density and Differences in Implicit Prices for Land Use Attributes

Attribute	Pearson Correlation Coefficient
iso_open (acres open space isolated from developments)	-0.8234
adj_open (acres open space adjacent to developments)	-0.3703
size_dif (acres in housing developments)	0.8144
dense_dif (housing density in developments)	0.7201