Benefit Transfer from Multiple Contingent Experiments: A Flexible Two-Step Model Combining Individual Choice Data with Community Characteristics

Klaus Moeltner, Robert J. Johnston, Randall S. Rosenberger, Joshua M. Duke

The potential of past Choice Experiments (CEs) to provide useful information for a yet unstudied policy site or context has received increasing recognition in recent years (e.g. Morrison and Bergland 2006; Johnston 2007). By design CEs can address a flexible mix of site or context attributes which are likely to include the relevant set of attributes for the policy context for which a transfer of information or "benefits" is sought. In addition it is conceivably more practical for a researcher to calibrate CE designs to match past examples on similar topics and thus contribute to a "homogenization" of research

---

Klaus Moeltner is associate professor in the Department of Resource Economics at the University of Nevada, Reno. Robert J. Johnston is director of the George Perkins Marsh Institute (GPMI) and professor in the Department of Economics at Clark University, Worcester, MA. Randall S. Rosenberger is assistant professor in the Department of Forestry Resources at Oregon State University. Joshua M. Duke is professor in the Department of Food and Resource Economics at the University of Delaware.

This article was presented in an invited paper session at the 2009 AAEA annual meeting in Milwaukee, WI. The articles in these sessions are not subjected to the journal's standard refereeing process.

This work was funded by EPA STAR grant # RD-83346001, NRI / CSREES grant # 2003-35400-13875, an Agricultural Viability grant from the State of Connecticut Department of Agriculture, and the UNR and OSU Experiment Stations.

instruments than it is to align survey questionnaires and data collection in a real-world, revealed preference setting.

The focus of this study is on benefit transfer (BT) based on combined information from *multiple* CEs. In principle there are two general approaches to build a candidate transfer function from several CE sources: (i) *The aggregate approach*, which uses the reported parameter estimates from original CE studies and combines them with attribute settings pertinent to the policy context; or (ii) *The choice-level approach*, which combines the raw choice data from source studies to generate a new set of estimates of transfer parameters. The first approach is illustrated by Johnston, Duke, and Kukielka (2008), who generate and average point estimates of policy-relevant welfare measures, and by Kukielka, Johnston, and Duke (2008), who feed welfare estimates corresponding to different attribute settings and sources into a second-stage meta-regression model. The second approach is implemented in Morrison and Bennett (2004), Johnston (2007), and Johnston and Duke (2009) with varying numbers of source studies and pooling constraints.

The aggregate approach is attractive to the time-constrained policy maker in that it does not require "chasing after" original data. However, it also has serious shortcomings. Specifically, the averaging over point estimates approach cannot utilize publicly available secondary information, such as geographical characteristics or community statistics, which could lead to a richer and thus more accurate transfer function. Furthermore, averaging over point estimates suppresses much of the underlying study-specific heterogeneity in preferences and may result in a misleadingly tight estimated distribution of transferred benefits. The meta-analytical variant, while able to incorporate secondary, community-level information, suffers from the usual pitfalls of mismatched sets of regressors across sources, and the dilemma of how to handle study-methodological attributes in the transfer function (see Moeltner, Boyle, and Paterson 2007). In addition neither aggregate variant preserves the link with a utility-theoretic framework.

The second approach, building on raw choice data from all original studies, naturally provides more flexibility in this latter respect: the BT analyst has the option to adopt the utility-theoretic framework chosen by the original authors of each source study or alternatively re-estimate the raw data under a different utility-theoretic umbrella. Other challenges, however, remain. Most notably, it is not clear how to pool data from choice experiments that differ in their design matrix (i.e. in attributes or attribute settings). It is thus not surprising that all existing contributions that have taken the choice-level approach build on CE data from *identical* experiments administered at different locations.

This study aims to capitalize on the strengths of both strategies. We propose a flexible two-step approach that combines raw choice data from potentially heterogeneous CE experiments with community-level information to generate a predictive distribution of policy-relevant benefits. Unlike existing contributions, our approach does not impose any cross-study pooling constraints on underlying preference structures or parameters. Specifically, we first re-estimate each original CE model separately in a Random Utility Model (RUM) framework, allowing for a maximum degree of unobserved individual heterogeneity in preferences for CE attributes. Each source model yields a predictive distribution of policy-relevant benefits. In the second stage we then generate a mixture distribution of benefits by combining these individual densities with discrete model weights composed of spatial and community-level characteristics. Since these weights are functionally independent of underlying preferences, the analyst has considerable flexibility in their construction.

For the dual reasons of computational convenience and intuitive interpretation of predictive constructs, we use a Bayesian estimation framework for the first analytical step. Our key finding is that predicted benefit distributions flowing from our proposed mixture model have substantially better overlap with directly estimated benefits based on actual data than the worst-case transferred benefits building on a single source study. At the same time we find that our empirical weights based on spatial and community statistics have only limited ability to improve over perfectly uniform weights. However, we believe that further gains in BT accuracy are possible with richer community-level data. This will be a potentially fruitful subject for future research.

Modeling Framework

Random Utility Model

Our empirical application is based on eight existing farmland preservation studies that use a CE format to elicit implicit prices and welfare measures for different bundles of farmland attributes. A set of four studies each use identical CE formats. All eight CE designs have choice menus with three mutually exclusive options: preservation of parcel one, preservation of parcel two, or non-preservation of either. An interesting feature in all eight studies is that one of the stipulated attributes of a hypothetical parcel is the probable time horizon of development if the land is left unpreserved. This attribute was treated as a direct argument in the indirect utility function in the original studies. We propose an arguably more intuitive strategy to introduce these development probabilities into a RUM framework.

Let the non-stochastic component of annual indirect utility to a respondent from the presence of Q_j acres of a specific type of farmland in her community be given as:

(1) $U^*_{(j)} = Q_j\left(\mathbf{L}'_\mathbf{j}\boldsymbol{\gamma}_\mathbf{L} + \mathbf{A}'_\mathbf{j}\boldsymbol{\gamma}_\mathbf{A}\right) + \delta\left(M - P_j\right),$

where $\mathbf{L}_\mathbf{j}$ is a vector of indicators summing to one for land use (food production, idle, orchard, etc.), $\mathbf{A}_\mathbf{j}$ is a vector of indicators summing to one for the level of public accessibility (none, walking, hunting, etc.), *M* is annual income, P_j is the stipulated annual preservation cost to the respondent, and $\boldsymbol{\gamma}_\mathbf{L}$, $\boldsymbol{\gamma}_\mathbf{A}$, and δ are corresponding slope coefficients. We use the notation (j) to distinguish the utility associated with an individual *parcel* from the utility flowing from a selected *choice option* (see below).

Each of the two proposed parcels is associated with a development probability $\pi_j, j = 1,2$. When contemplating the three options, the individual will have to weigh the expected benefits of preservation against the certain costs. Specifically, if she chooses option *j*, she will preserve parcel *j* for the coming year at cost P_j, but there is also a $\left(1-\pi_{k \neq j}\right)$ probability that parcel $k \neq j$ remains undeveloped in the coming year as well. Thus, the choice of either option results in the following expected indirect utility:

$$U_j^* = Q_j\left(\mathbf{L}_\mathbf{j}'\boldsymbol{\gamma}_\mathbf{L} + \mathbf{A}_\mathbf{j}'\boldsymbol{\gamma}_\mathbf{A}\right) + \left(1-\pi_k\right)Q_k\left(\mathbf{L}_\mathbf{k}'\boldsymbol{\gamma}_\mathbf{L} + \mathbf{A}_\mathbf{k}'\boldsymbol{\gamma}_\mathbf{A}\right) + \delta\left(M - P_j\right) =$$

$$(2)\quad \left(Q_j\mathbf{L}_\mathbf{j} + \left(1-\pi_k\right)Q_k\mathbf{L}_\mathbf{k}\right)'\boldsymbol{\gamma}_\mathbf{L} + \left(Q_j\mathbf{A}_\mathbf{j} + \left(1-\pi_k\right)Q_k\mathbf{A}_\mathbf{k}\right)'\boldsymbol{\gamma}_\mathbf{A} + \delta\left(M - P_j\right), \quad j,k \in \{1,2\}, k \neq j.$$

Similarly, a decision to protect neither parcel results in:

$$U_3^* = \left(1-\pi_j\right)Q_j\left(\mathbf{L}_\mathbf{j}'\boldsymbol{\gamma}_\mathbf{L} + \mathbf{A}_\mathbf{j}'\boldsymbol{\gamma}_\mathbf{A}\right) + \left(1-\pi_k\right)Q_k\left(\mathbf{L}_\mathbf{k}'\boldsymbol{\gamma}_\mathbf{L} + \mathbf{A}_\mathbf{k}'\boldsymbol{\gamma}_\mathbf{A}\right) + \delta M =$$

$$(3)\quad \left(\left(1-\pi_j\right)Q_j\mathbf{L}_\mathbf{j} + \left(1-\pi_k\right)Q_k\mathbf{L}_\mathbf{k}\right)'\boldsymbol{\gamma}_\mathbf{L} + \left(\left(1-\pi_j\right)Q_j\mathbf{A}_\mathbf{j} + \left(1-\pi_k\right)Q_k\mathbf{A}_\mathbf{k}\right)'\boldsymbol{\gamma}_\mathbf{A} + \delta M,$$

$$j,k \in \{1,2\}, k \neq j.$$

Thus, rather than following the customary procedure of setting the non-stochastic component of the "opt-out" utility to zero, we propose a more realistic version that affords to the respondent positive expected benefits at zero cost unless $\pi_j = \pi_k = 1$, which does not apply to our case.

Adding an i.i.d. error term with zero mean to each equation yields the following decision rule for the choice of option *j*:

$$(4)\quad \begin{aligned} &\pi_j Q_j\left(\mathbf{L}_\mathbf{j}'\boldsymbol{\gamma}_\mathbf{L} + \mathbf{A}_\mathbf{j}'\boldsymbol{\gamma}_\mathbf{A}\right) - \pi_k Q_k\left(\mathbf{L}_\mathbf{k}'\boldsymbol{\gamma}_\mathbf{L} + \mathbf{A}_\mathbf{k}'\boldsymbol{\gamma}_\mathbf{A}\right) + \delta\left(P_k - P_j\right) + \left(\varepsilon_j - \varepsilon_k\right) > 0 \quad \text{and} \\ &\pi_j Q_j\left(\mathbf{L}_\mathbf{j}'\boldsymbol{\gamma}_\mathbf{L} + \mathbf{A}_\mathbf{j}'\boldsymbol{\gamma}_\mathbf{A}\right) - \delta P_j + \left(\varepsilon_j - \varepsilon_3\right) > 0. \end{aligned}$$

This is intuitively sound. The first equation states that under equal attributes and prices and holding errors at zero, the respondent chooses the parcel that is at a higher risk of development. In addition, as expressed by the second equation, the expected loss in utility from developing the parcel has to exceed the preservation price. Naturally, this also implies that if *development* is generally preferred (i.e. $\mathbf{L}_\mathbf{j}'\boldsymbol{\gamma}_\mathbf{L} + \mathbf{A}_\mathbf{j}'\boldsymbol{\gamma}_\mathbf{A} < 0,\ j = 1,2$), the respondent will always opt out of preserving either parcel.

Econometric Model

As in most modern CE applications, each respondent $i = 1 \cdots N_s$ in our $s = 1 \cdots S$ source studies receives $t = 1 \cdots T_s$ sequential, design-independent choice menus, each consisting of J choice options. Thus, there are JxT_s observations per respondents in the data set corresponding to study *s*. Allowing for the possibility that all land use and access level indicators may be associated with unobserved heterogeneity in individual preferences, our econometric model can be expressed at the panel level as:

$$(5) \quad \mathbf{U_i^*} = \mathbf{X_i}\boldsymbol{\beta} + \mathbf{X_{ri}}\boldsymbol{\alpha_i} + \boldsymbol{\varepsilon_i} \qquad \boldsymbol{\varepsilon_i} \sim n\left(\mathbf{0}, \mathbf{I}_{JxT_s}\right), \ \boldsymbol{\alpha_i} \sim n\left(\mathbf{0}, \boldsymbol{\Sigma}\right),$$

where $\mathbf{U_i^*} = \begin{bmatrix} U_{i11}^* & U_{i21}^* & \cdots & U_{iJT_s}^* \end{bmatrix}'$, $\mathbf{X_i}$ includes all regressors in (2) and (3), $\mathbf{X_{ri}}$ is a subset of $\mathbf{X_i}$ that captures all regressors that are paired with random parameters, and $\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\gamma}_\mathbf{L}' & \boldsymbol{\gamma}_\mathbf{A}' & \delta \end{bmatrix}'$. Since we have no priors regarding the sign of attribute coefficients, we model all random parameters to follow a joint normal density, as indicated in the second line of (5). Furthermore, as shown in the first line, we set the i.i.d. error variance for all equations to one and all covariances to zero.[1] In essence this yields the random parameters multinomial probit (MNP) model of Hausman and Wise (1978). A Bayesian version of this model is presented in Layton and Levine (2003 and 2005) .

A given respondent will exhibit an observed choice sequence of

$\mathbf{y_i} = \begin{bmatrix} k_1 & k_2 & \cdots & k_{T_s} \end{bmatrix}$ if

$$\left\{ \max\left\{ \left\{ U_{ij1}^* \right\}_{j=1}^J \right\} = U_{i,k1}, \max\left\{ \left\{ U_{ij2}^* \right\}_{j=1}^J \right\} = U_{i,k2}, \cdots, \max\left\{ \left\{ U_{ijT}^* \right\}_{j=1}^J \right\} = U_{i,kT} \right\}.$$

As illustrated in Layton and Levine (2003), this can be efficiently modeled by subtracting all other utilities from the winning utility in an individual menu via an appropriate $\left(T * (J-1)\right) x \left(T * J\right)$ differencing matrix $\mathbf{D_i}$. Letting $\mathbf{U_i} = \mathbf{D_i}\mathbf{U_i^*}$, an individual's contribution to the likelihood function can then be written as:

(6) $p\left(\mathbf{y_i} \mid \mathbf{X_i}, \boldsymbol{\beta}, \boldsymbol{\Sigma}\right) = \int_{\boldsymbol{\alpha}_{ri}} \Phi\left(\mathbf{0}, \mathbf{D_i V_i D_i'}; \mathbf{R_i}\right) f\left(\boldsymbol{\alpha}_i\right) d\boldsymbol{\alpha}_i, \quad \mathbf{V_i} = \mathbf{X_{ri} \Sigma X_{ri}'} + \mathbf{I_{JxT_s}},$

where (with slight abuse of notation) $\Phi(.)$ denotes the *cdf* of the truncated multivariate normal density with mean $\mathbf{0}$, variance matrix $\mathbf{D_i V_i D_i'}$, and truncation region $\mathbf{R_i}$, and $f\left(\boldsymbol{\alpha}_i\right)$ denotes the multivariate normal density. The truncation region will always be bounded by $\mathbf{-D_i X_i} \boldsymbol{\beta}$ on the left and infinity on the right.

This model would be cumbersome to estimate in a classical framework. We thus opt for a Bayesian approach that stipulates prior densities for all parameters and that, via a Gibbs Sampler (GS), consecutively and repeatedly draws from the following conditional densities:

(7) $p\left(\boldsymbol{\beta} \mid \boldsymbol{\Sigma}, \mathbf{U}, \mathbf{X}\right), p\left(\boldsymbol{\alpha}_i \mid \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{U_i}, \mathbf{X_i}\right), i = 1 \cdots N_s, p\left(\boldsymbol{\Sigma} \mid \boldsymbol{\alpha}_i\right)$, and $p\left(\mathbf{U_i} \mid \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{y_i}, \mathbf{X_i}\right), i = 1 \cdots N_s$

where $\mathbf{X} = \begin{bmatrix} \mathbf{X_1} & \mathbf{X_2} & \cdots & \mathbf{X_{N_s}} \end{bmatrix}$, and $\mathbf{U} = \begin{bmatrix} \mathbf{U_1} & \mathbf{U_2} & \cdots & \mathbf{U_{N_s}} \end{bmatrix}$. After an appropriate number of discarded draws ("burn-ins"), this posterior sampler will converge to the joint posterior density of the main model parameters $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$, i.e. $p\left(\boldsymbol{\beta}, \boldsymbol{\Sigma} \mid \mathbf{y}, \mathbf{X}\right)$.[2]

Posterior Predictive Densities

In the first step of our analysis, we estimate a separate hierarchical MNP model for each of our *S* underlying source studies and corresponding data sets. We are primarily interested in the posterior predictive density (PPD) of the annual compensating surplus for a prototypical resident from sub-population *s* for a farmland with attributes $\mathbf{x_p} = Q_p * \begin{bmatrix} \mathbf{L_p'} & \mathbf{A_p'} \end{bmatrix}'$, relative to a "full development" scenario with $Q_0 = 0$. The settings for Q_p, $\mathbf{L_p}$, and $\mathbf{A_p}$ are chosen to reflect the farmland attributes at the policy site (i.e. the BT "target"). Under an "identical error" assumption (i.e. $\varepsilon_0 = \varepsilon_p$) and price invariance (i.e. $P_0 = P_p$, see Morey and Rossmann 2008) this welfare metric takes the following form, conditional on model parameters and a given draw of random deviations $\boldsymbol{\alpha}$:

(8) $C_s\left(\mathbf{x_p}\right) | \boldsymbol{\alpha},\boldsymbol{\beta_s} = C_{sp} | \boldsymbol{\alpha},\boldsymbol{\beta_s} = -\delta_s^{-1}\left(\mathbf{x'_p}\boldsymbol{\beta}_{-P,\mathbf{s}} + \mathbf{x'_{pr}}\boldsymbol{\alpha}\right)$

where $\mathbf{x_{pr}}$ comprises the random regressors in $\mathbf{x_p}$. It is important to note that the true, unknown error scale drops out of this expression, which enables us to directly compare the welfare measures flowing from the *S* studies without further adjustments for differences in scale. The PPD for C_{sp}, conditioned only on observed choices and the CE design matrix for study *s*, is then given as:

(9) $p\left(C_{sp} | \mathbf{y_s},\mathbf{X_s}\right) = \int\limits_{\boldsymbol{\theta_s}}\left(\int\limits_{\boldsymbol{\alpha}}\left(C_{sp} | \boldsymbol{\alpha},\boldsymbol{\beta_s},\delta_s\right) f\left(\boldsymbol{\alpha} | \boldsymbol{\Sigma_s}\right) d\boldsymbol{\alpha}\right) p\left(\boldsymbol{\theta_s} | \mathbf{y_s},\mathbf{X_s}\right) d\boldsymbol{\theta_s}$

where vector $\boldsymbol{\theta_s}$ comprises all elements of $\boldsymbol{\beta_s}$ and $\boldsymbol{\Sigma_s}$. Given, say, *R* draws of $\boldsymbol{\theta_s}$ from the original Gibbs Sampler, it is straightforward to obtain draws from this PPD. The details of this process are given in Moeltner, Johnston, and Rosenberger (2009).

Weighted Mixture Distribution

In the second step of our analysis, we combine the informational content of all *S* welfare distributions in a finite mixture framework. Specifically, we stipulate that the true, unknown distribution of compensating surplus at the policy site follows a weighted mixture distribution with the *S* PPDs from step one as its continuous components, i.e.:

$$(10)\quad p\left(C_{pp}\right)=\sum_{s=1}^{S}\psi_s\, p\left(C_{sp}\mid \mathbf{y_s},\mathbf{X_s}\right)\quad \text{with}\quad \sum_{s=1}^{S}\psi_s=1.$$

Setting $\psi_s = 1/S$ would allocate equal weight across source studies. However, ideally we would like to assign relatively larger weights to sources that are "more similar" to the target site. The quest for such "intelligent" weights is the focus of the second step of our analysis, as described in the next section.

Empirical Application

Data

Our eight source studies flow from two separate research projects: (i) a CE on farmland preservation administered in the Delaware Communities of Georgetown (GT) and Smyrna (SM) and the Connecticut towns of Mansfield (MF) and Preston (PR) in 2005 and 2006; and (ii) a similar but not identical CE implemented between 2005 and 2007 in the Connecticut communities of Brooklyn (BR), Pomfret (PO), Thompson (TH), and Woodstock (WO). We will henceforth refer to these two clusters of communities as "set 1" and "set 2." For details on these projects see and Johnston, Duke, and Kukielka (2008). Respondents received three menus for set 1 and four menus for set 2. After eliminating observations with missing demographic information, we retain 1,066 individuals (9,594 observations) for set 1 and 707 individuals (8,484 observations) for set 2. Within each set these observation counts are distributed approximately evenly across communities.

Step One Estimation

For set 1 the farmland attribute vector $\mathbf{L}_{\mathrm{j}}$ includes indicators for "nursery," "food crop," "dairy or livestock," and "forest." The access vector $\mathbf{A}_j$ includes indicators for "walking" and "hunting." The $\mathbf{L}_j$ - components in set 2 are "food / field crop," "dairy or livestock," and "tree farm, nursery, or orchard," and $\mathbf{A}_j$ represents a single indicator for "access for passive recreation." For both sets we also include a constant term in $\mathbf{L}_j$ to capture the per-acre effect of the implicit baseline category "idle" and "no access." In all cases the parcel sizes include 20, 60, 100, and 200 acres, and preservation costs are $5, $15, $30, $50, $100 ($120 for set 2), and $200.

For set 1 the time horizons for probable development are given as "<10 years," "10-30 years," and "not likely in 30 years." Assuming that respondents envision a uniform distribution of development probabilities over these time horizons, we set π_j to 1/10 and 1/20, for the first two cases, respectively, and to zero for the third development scenario. The second set uses only an indicator for "development likely in less than 10 years," which we translate into a development probability of 1/10.

We model all regression coefficients other than the one for price as random. Given data limitations and to conserve on parameters, we set all hierarchical covariances to zero. While this breaks the cross-equation links in a given panel, it preserves the notion of unobserved heterogeneity for farmland attributes. We implement our Gibbs Sampler with standard vague but proper priors for all parameters (i.e. $\boldsymbol{\beta} \sim n\left(\mathbf{0}, 10 * \mathbf{I}\right)$ and $\Sigma_{jj} \sim ig\left(\frac{1}{2}, \frac{1}{2}\right)$ where $ig\left(a, b\right)$ denotes inverse-gamma density with shape parameter *a* and scale parameter *b*). All models are estimated using 10,000 burn-in draws and 10,000 retained draws flowing from the Gibbs Sampler. The decision on the appropriate amount of burn-ins was guided by Geweke's (1992) convergence diagnostic.

In the interest of brevity, first-step estimation results are omitted but provided in Moeltner, Johnston, and Rosenberger (2009). We simply report that all eight communities exhibit pronounced within-sample heterogeneity with respect to most farmland attributes, which lends support to our hierarchical modeling choice. Furthermore, the degree of heterogeneity in preferences varies across communities, supporting a case-by-case estimation approach. For most attributes and communities random parameter means lie in the negative domain, indicating a general preference for development as opposed to preservation for the prototypical respondent. Given the largely rural settings for most of these towns, this is not all that surprising.

To illustrate our approach to BT, we stipulate a single policy scenario (i.e. the preservation of one acre of idle farmland with access for passive recreation). The corresponding PPDs from all eight models are depicted in figure 1. As seen in the figure, all eight densities exhibit reasonably good distributional overlap. It is clear from the graph that some community pairs, such as Pomfret and Thompson, would be very well suited for cross-community transfers, but other single-study matches, such as Smyrna and Woodstock, would result in seriously misleading inferences.

Step Two Estimation

To assess the accuracy of our proposed methodology, we use, in sequence, each of the eight cases as the target site with a presumably unknown welfare distribution and the remaining seven densities to feed into the mixture model given in (10). As an indicator for transfer accuracy, we propose a novel metric, the "overlap in highest posterior density intervals (HPDIs) to full range of HPDIs," in short "OLR." The HPDI is the Bayesian analog to the classical confidence interval. A 95 percent HPDI, for example, delivers a lower and upper bound such that the resulting interval is the *smallest possible* to contain 95 percent of the density mass of a given distribution. Formally, the OLR between two distributions *s* and *z* is derived as:

$$(11) \quad OLR_{sz} = \left(\min\left(u_s, u_z\right) - \max\left(l_s, l_z\right)\right) / \left(\max\left(u_s, u_z\right) - \min\left(l_s, l_z\right)\right)$$

where $l_{(.)}$ and $u_{(.)}$ denote the lower and upper limits of the respective 95 percent HPDIs. The first half of table 1 shows this metric for all possible pairs of community-specific densities. As was evident from figure 1, Pomfret and Thompson exhibit close-to-perfect overlap, while the OLR drops to 51 percent for Smyrna and Woodstock.

We employ three different sets of mixture weights. The first two are, respectively, a set of uniform weights (i.e. $\psi_s = 1/\left(S-1\right) = 0.143, \forall s$); and (ii) a set of weights based on inverse distances (i.e. $\psi_s = D_{sp}^{-1} / \sum_{s=1}^{S-1} D_{sp}^{-1}$, where D_{sp} is the distance, in miles, between study *s* and the target site *p*). The third approach requires an additional estimation round. Specifically, we regress each of the $\binom{S-1}{2} = 21$ pair-wise OLR measures against distance and differences in aggregate community characteristics (i.e. population per acre and the share of urban households in the empirical sample). Since the OLR is naturally bounded by zero and one, we estimate this auxiliary model in a truncated regression framework via MLE.[3] We then combine the estimated parameters from this approach with the relative difference in community settings between each of the study sites and the *target site* and use the resulting *S-1* predicted values of OLR_{sp} to compute the mixture weights (i.e. $\psi_s = \hat{OLR}_{sp} / \sum_{s=1}^{S-1} \hat{OLR}_{sp}$). For specification details and estimation results see Moeltner, Johnston, and Rosenberger (2009).

We generally find that the regression-based weight do not differ by much from uniform weights, while the inverse-distance weights, with ranges between 0.01 to 0.3, exhibit stronger deviation from uniformity. However, the inverse distance approach is an imprecise tool, as it does not categorically assign higher weights to sites that have a better OLR with the target. The second half of table 1 shows the OLR values for all three mixture models with respect to all eight individual target sites. The key result captured in the table is that any of the three mixture models generates BT distributions that fit any of the target sites substantially better than the worst-case single-site transfer. Thus, at least for our application, the mixture model strategy appears to be a more robust approach than a single-site transfer.

Conclusion

We propose a novel approach to BT from multiple CE experiments that allows for a full recognition of heterogeneity in sub-population preferences and experimental designs, substantial flexibility in utility-theoretic modeling, and the use of secondary socio-demographic and geo-spatial information to refine BT functions. Our analysis can be extended in several dimensions. The critical next step for future research will be the identification of more pertinent community characteristics to further refine the step two mixture weights. Also, in a different policy context a different metric of transfer fit than the OLR may be required, such as the mean or median of welfare distributions. This can be easily incorporated in our methodological framework. Finally, it would be interesting to see this approach applied to a cluster of CEs with different resource focus and design heterogeneities.

[1] Our error specification is based on the recognition that since options within menus and menus within and across respondents change randomly by design, there is no rationale to allow for different error variances across equations. Since at least one variance has to be normalized in any case, we set all of them to one. By the same token there is no conceptual basis for specifying covariance terms.

[2] We opt for a classical estimation approach for this step as the truncated regression model would be extremely cumbersome to handle in a Bayesian framework, and the option to use informed priors does not present itself in this case.

Table 1. Posterior Predictive Fit of Compensating Surplus Distribution, Original Models, and Benefit Transfer Models

HPDI Relative Overlap: Original Models

	GT	MF	PR	SM	BR	PO	TH
MF	0.83	-	-	-	-	-	-
PR	0.80	0.83	-	-	-	-	-
SM	0.79	0.77	0.63	-	-	-	-
BR	0.84	0.89	0.73	0.86	-	-	-
PO	0.84	0.96	0.79	0.80	0.93	-	-
TH	0.88	0.94	0.80	0.79	0.92	0.95	-
WO	0.64	0.66	0.80	0.51	0.59	0.63	0.64

HPDI Relative Overlap: Transfer Models

	GT	MF	PR	SM	BR	PO	TH	WO
regression	0.89	0.90	0.86	0.69	0.80	0.89	0.88	0.66
distance	0.91	0.88	0.87	0.71	0.81	0.87	0.82	0.65
uniform	0.88	0.92	0.86	0.70	0.81	0.87	0.89	0.66

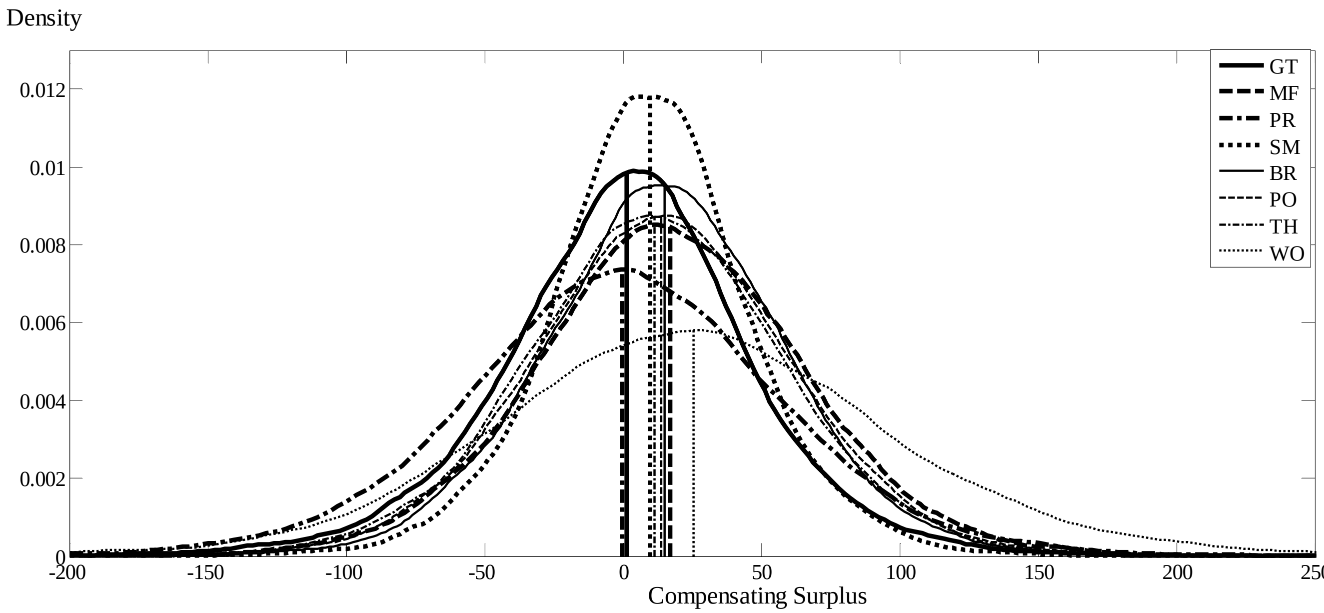

Figure 1. Posterior distribution of compensating surplus, individual models

Vertical lines indicate means

References

Geweke, J. 1992. " Evaluating the Accuracy of Sampling-based Approaches to the Calculation of Posterior Moments." In J. M. Bernardo, J. O. Berger, A. P. Dawid, A. F. M. Smith, eds. *Bayesian Statistics 4*. Oxford, UK: Oxford University Press, pp. 641-649.

Hausman, J.A., and D.A. Wise. 1978. "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences." *Econometrica* 48:403-426.

Johnston, R.J. 2007. "Choice Experiments, Site Similarity, and Benefit Transfer." *Environmental and Resource Economics* 38:331-351.

Johnston, R.J., and J.M. Duke. 2009. "Willingness to Pay for Land Preservation Across States and Jurisdictional Scale: Implications for Benefit Transfer." *Land Economics* 85:217-237.

Johnston, R.J., J.M. Duke, and J.G. Kukielka. 2008. " Optimizing Farmland Preservation Choices across Communities and Jurisdictional Scales: To what Extent are Amenity Values and Selection Criteria Transferable?" Paper presented at the Annual Meeting of the American Agricultural Economics Association, Orlando FL, 27-29 July.

Kukielka, J.G., R.J. Johnston, and J.M. Duke. 2008. "Systematic Variation in Willingness-to-Pay for Agricultural Land Preservation and Implications for Benefit Transfer: A Meta-Analysis." Paper presented at the Annual Meeting of the American Agricultural Economics Association, Orlando FL, 27-29 July.

Layton, D.F., and R.A. Levine. 2005. "Bayesian Approaches to Modeling Stated Preference Data." In R. Scarpa and A. Alberini, eds. *Applications of Simulation Methods in Environmental and Resource Economics*. Dordrecht, The Netherlands: Springer, pp. 187-205.

Layton, D.F., and R.A. Levine. 2003. "How Much Does the Far Future Matter? A Hierarchical Bayesian Analysis of the Public's Willingness to Mitigate Ecological Impacts of Climate Change." *Journal of the American Statistical Association* 98:533-544.

Moeltner, K., K. Boyle, and R. Paterson. 2007. "Meta-Analysis and Benefit Transfer for Resource Valuation - Addressing Classical Challenges with Bayesian Modeling." *Journal of Environmental Economics and Management* 53:250-269.

Moeltner, K., R.J. Johnston, and R. S. Rosenberger. 2009. "AJAE Appendix: Benefit Transfer from Multiple Contingent Experiments: A Flexible Two-Step Model Combining Individual Choice Data with Community Characteristics." Unpublished manuscript. Available at: http://ageconsearch.umn.edu/

Morey, E., and K.G. Rossmann. 2008. "Calculating, with Income Effects, the Compensating Variation for a State Change." *Environmental and Resource Economics* 39:83-90.

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

Morrison, M., and O. Bergland. 2006. "Prospects for the Use of Choice Modeling for

Benefit Transfer." *Ecological Economics* 60:420-428.