

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

Klaus Moeltner*

University of Nevada, Reno

Robert J. Johnston

Clark University

Randall S. Rosenberger

Oregon State University

Joshua M. Duke

University of Delaware

Please cite as

University of Nevada, Reno Joint Economics Working Paper Series

Working Paper No. 09-002

July, 2009

* Contact Information:

Klaus Moeltner

Department of Resource Economics / MS 204

University of Nevada, Reno

Reno, NV 89557-0105

phone: (775) 784-4803

fax: (775) 784-1342

e-mail: moeltner@cabnr.unr.edu

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.

Abstract:

This study proposes a new approach to utilize information from existing choice experiments to predict policy outcomes for a transfer setting. Recognizing the difficulties from pooling raw data from experiments with different designs and sub-populations we first re-estimate all underlying Random Utility Models individually, and then combine them in a second stage process to form a weighted mixture density for the generation of policy-relevant welfare estimates. Using data from recent choice experiments on farmland preservation we illustrate that our strategy is more robust to transfer inaccuracies than single-site approaches. The specification of "intelligent" mixture weights will be a fruitful ground for future research in the area of Benefit Transfer.

Introduction

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 feasible for researcher to calibrate CE designs to match past examples on similar topics and thus contribute to a "homogenization" of research 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 Kulioka (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 degrees of 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 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 unbalanced set 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 (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. However, the entire analysis could also be implemented in a classical estimation setting with a slightly different interpretation of predictive densities. 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 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. This also illustrates the above-mentioned utility-theoretic flexibility afforded to the BT analyst when working with raw choice data.

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

$$U_{(j)}^* = Q_j (\mathbf{L}'_j \boldsymbol{\gamma}_L + \mathbf{A}'_j \boldsymbol{\gamma}_A) + \delta (M - P_j) \quad (1)$$

where \mathbf{L}_j is a vector of indicators summing to one for land use (food production, idle, orchard, etc), \mathbf{A}_j is a vector of indicators summing to one for the level of public

accessibility (none, walking, hunting, etc.), M is annual income, and P_j is the stipulated annual preservation cost to the respondent. 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).¹

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 $(1-\pi_{k \neq j})$ probability that parcel $k \neq j$ remains undeveloped in the coming year as well.²

Thus, the choice of either option results in the following expected indirect utility:

$$U_j^* = Q_j (\mathbf{L}'_j \gamma_L + \mathbf{A}'_j \gamma_A) + (1-\pi_k) Q_k (\mathbf{L}'_k \gamma_L + \mathbf{A}'_k \gamma_A) + \delta(M - P_j) = \\ (Q_j \mathbf{L}_j + (1-\pi_k) Q_k \mathbf{L}_k)' \gamma_L + (Q_j \mathbf{A}_j + (1-\pi_k) Q_k \mathbf{A}_k)' \gamma_A + \delta(M - P_j), \quad j, k \in \{1, 2\}, k \neq j \quad (2)$$

Similarly, a decision to protect neither parcel results in

$$U_3^* = (1-\pi_j) Q_j (\mathbf{L}'_j \gamma_L + \mathbf{A}'_j \gamma_A) + (1-\pi_k) Q_k (\mathbf{L}'_k \gamma_L + \mathbf{A}'_k \gamma_A) + \delta M = \\ ((1-\pi_j) Q_j \mathbf{L}_j + (1-\pi_k) Q_k \mathbf{L}_k)' \gamma_L + ((1-\pi_j) Q_j \mathbf{A}_j + (1-\pi_k) Q_k \mathbf{A}_k)' \gamma_A + \delta M, \quad j, k \in \{1, 2\}, k \neq j \quad (3)$$

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 :

$$\begin{aligned} \pi_j Q_j (\mathbf{L}'_j \boldsymbol{\gamma}_L + \mathbf{A}'_j \boldsymbol{\gamma}_A) - \pi_k Q_k (\mathbf{L}'_k \boldsymbol{\gamma}_L + \mathbf{A}'_k \boldsymbol{\gamma}_A) + \delta (P_k - P_j) + (\varepsilon_j - \varepsilon_k) > 0 \quad \text{and} \\ \pi_j Q_j (\mathbf{L}'_j \boldsymbol{\gamma}_L + \mathbf{A}'_j \boldsymbol{\gamma}_A) - \delta P_j + (\varepsilon_j - \varepsilon_3) > 0 \end{aligned} \quad (4)$$

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. $\mathbf{L}'_j \boldsymbol{\gamma}_L + \mathbf{A}'_j \boldsymbol{\gamma}_A < 0$, $j=1,2$), 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. 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

$$\mathbf{U}_i^* = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{X}_{i1} \boldsymbol{\alpha}_i + \boldsymbol{\varepsilon}_i \quad \boldsymbol{\varepsilon}_i \sim n(\mathbf{0}, \mathbf{I}_{JxT_s}), \quad \boldsymbol{\alpha}_i \sim n(\mathbf{0}, \boldsymbol{\Sigma}) \quad (5)$$

where $\mathbf{U}_i^* = [U_{i11}^* \quad U_{i21}^* \quad \cdots \quad U_{iJT_s}^*]'$, \mathbf{X}_i includes all regressors in (2) and (3), \mathbf{X}_{i1} is a subset of \mathbf{X}_i that captures all regressors that are paired with random parameters, and $\boldsymbol{\beta} = [\boldsymbol{\gamma}'_L \quad \boldsymbol{\gamma}'_A \quad \delta]'$. Since we have no ex ante 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.³ 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

$$\mathbf{y}_i = [k_1 \quad k_2 \quad \cdots \quad k_{T_s}] \text{ if}$$

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

As illustrated in Layton and Levine (2003) this can be efficiently modeled by subtracting all other utilities from the winning utility within an individual menu via an appropriate

$(T^*(J-1)) \times (T^*J)$ 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

$$p(\mathbf{y}_i | \mathbf{X}_i, \boldsymbol{\beta}, \boldsymbol{\Sigma}) = \int_{\boldsymbol{\alpha}_i} \Phi(\mathbf{0}, \mathbf{D}_i \mathbf{V}_i \mathbf{D}_i'; \mathbf{R}_i) f(\boldsymbol{\alpha}_i) d\boldsymbol{\alpha}_i, \quad \mathbf{V}_i = \mathbf{X}_{ri} \boldsymbol{\Sigma} \mathbf{X}_{ri}' + \mathbf{I}_{J \times T}, \quad (6)$$

where (with slight abuse of notation) $\Phi(\cdot)$ denotes the *cdf* of the truncated multivariate normal density with mean $\mathbf{0}$, variance matrix $\mathbf{D}_i \mathbf{V}_i \mathbf{D}_i'$, and truncation region \mathbf{R}_i . This region will always be bounded by $-\mathbf{D}_i \mathbf{X}_i \boldsymbol{\beta}$ 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:

$$p(\boldsymbol{\beta} | \boldsymbol{\Sigma}, \mathbf{U}, \mathbf{X}), p(\boldsymbol{\alpha}_i | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{U}_i, \mathbf{X}_i), i = 1 \cdots N_s, p(\boldsymbol{\Sigma} | \boldsymbol{\alpha}_i), \text{ and } p(\mathbf{U}_i | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{y}_i, \mathbf{X}_i), i = 1 \cdots N_s \quad (7)$$

where $\mathbf{X} = [\mathbf{X}_1 \ \mathbf{X}_2 \ \dots \ \mathbf{X}_{N_s}]$, and $\mathbf{U} = [\mathbf{U}_1 \ \mathbf{U}_2 \ \dots \ \mathbf{U}_{N_s}]$. 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. $p(\boldsymbol{\beta}, \boldsymbol{\Sigma} | \mathbf{y}, \mathbf{X})$.⁴ The detailed step for this Gibbs Sampler are given in Appendix A.

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

$\mathbf{x}_p = Q_p * [\mathbf{L}'_p \ \mathbf{A}'_p]'$, 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 $\boldsymbol{\alpha}$:

$$C_s(\mathbf{x}_p) | \boldsymbol{\alpha}, \boldsymbol{\beta}_s = C_{sp} | \boldsymbol{\alpha}, \boldsymbol{\beta}_s = -\delta_s^{-1} (\mathbf{x}'_p \boldsymbol{\beta}_{-p,s} + \mathbf{x}'_{pr} \boldsymbol{\alpha}) \quad (8)$$

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

$$p(C_{sp} | \mathbf{y}_s, \mathbf{X}_s) = \int_{\boldsymbol{\theta}_s} \left(\int_{\boldsymbol{\alpha}} (C_{sp} | \boldsymbol{\alpha}, \boldsymbol{\beta}_s, \delta_s) f(\boldsymbol{\alpha} | \boldsymbol{\Sigma}_s) d\boldsymbol{\alpha} \right) p(\boldsymbol{\theta}_s | \mathbf{y}_s, \mathbf{X}_s) d\boldsymbol{\theta}_s \quad (9)$$

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 shown in Appendix B.

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 distributions with the S PPDs from step one as its continuous components, i.e.:

$$p(C_{pp}) = \sum_{s=1}^S \psi_s p(C_{sp} | \mathbf{y}_s, \mathbf{X}_s) \quad \text{with} \quad \sum_{s=1}^S \psi_s = 1 \quad (10)$$

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". Details on the first project can be found e.g. in Johnston and Duke (2009). The second project is described in Johnston, Duke, and Kulieka (2008). Respondents received three menus for set 1 and four menus for set 2. After eliminating observations with missing demographic information we retain 1066 individuals (9594 observations) for set 1 and 707 individuals (8484 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 L_j includes indicators for "nursery", "food crop", "dairy or livestock", and "forest". The access vector A_j includes indicators for "walking" and "hunting". The L_j - components in set 2 are "food / field crop", "dairy or livestock", and "tree farm, nursery, or orchard", and A_j represents a single indicator for "access for passive recreation ". For both sets we also include a constant term in 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 range from \$5 to \$200 in six increments.

For set 1 the time horizons for probably 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 also 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 within 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. $\beta \sim n(\mathbf{0}, 10 * \mathbf{I})$ and $\Sigma_{jj} \sim ig(\frac{1}{2}, \frac{1}{2})$ where $ig(a, b)$ denotes inverse-gamma density with shape parameter a and scale parameter b . All models are estimated using 10000 burn-in draws and 10000 retained draws flowing from the Gibbs Sampler. The decision on the appropriate amount of burn-ins was guided by Geweke's (1992) convergence diagnostic.

Tables 1 and 2 capture first-step estimation results. Clearly, 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 can be seen from 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% HPDI, for example, delivers a lower and upper bound such that the resulting interval is the *smallest possible* to contain 95% of the density mass of a given distribution. Formally, the OLR between two distributions s and z is derived as

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

where l and u denote the lower and upper limits of the respective 95% HPDIs. The first half of table 5 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% for Smyrna and Woodstock.

We employ three different sets of mixture weights. The first two are, respectively, a set of uniform weights, i.e. $\psi_s = 1 / (S - 1) = 0.143, \forall s$, 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.⁵ Table 3 depicts community characteristics and relative distances. We combine the estimated parameters from this auxiliary regression 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. $\psi_s = \hat{OLR}_{sp} / \sum_{s=1}^{S-1} \hat{OLR}_{sp}$. The details of the regression model are given in Appendix C. Appendix D describes the procedure to draw from the weighted mixture distribution.

The resulting weights from all three approaches are shown in table 4 for some selected sites. We generally find that the regression-based weights, ranging between 0.13 and 0.16 for most cases, do not differ by much from uniform weights, while the inverse-distance weights, which range between 0.01 to 0.3, exhibit much 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. For example, it is clear from figure 1 that Georgetown and Brooklyn have reasonably good overlap despite being almost 300 miles apart (table 3). In contrast, Georgetown and Smyrna are virtually neighbors, but exhibit a relatively poor OLR.

The second half of table 5 shows the OLR values for predictions from 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 is clearly a safer approach than a single-site transfer. Figure 2 depicts transfer results from the mixture models in graphical form. As can be seen from the figure, all three mixture distributions fit the target density reasonably well to extremely well for most target sites. However, there are cases (e.g. Woodstock) that leave room for predictive improvements. Within our proposed framework such improvements will require a more careful specification of mixture weights. This will be the subject of the next stage in this broader research project.

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, maximum 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, subject to 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.

Appendix A: Gibbs Sampler Details

For simplicity we will omit the "s" subscript for the sample size N and the number of menus per respondent, T .

Priors:

$$\begin{aligned}
 p(\boldsymbol{\beta}, \boldsymbol{\Sigma}) &= p(\boldsymbol{\beta}) p(\boldsymbol{\Sigma}) \quad \text{where} \\
 p(\boldsymbol{\beta}) &= (2\pi)^{-k_f/2} |\mathbf{V}_0|^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\mu}_0)' \mathbf{V}_0^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}_0)\right) \\
 p(\boldsymbol{\Sigma}) &= \prod_{j=1}^{k_r} p(\Sigma_{jj}) \quad \text{with} \\
 p(\Sigma_{jj}) &= \frac{\tau_0^{v_0}}{\Gamma(v_0)} (\Sigma_{jj})^{-(v_0+1)} \exp\left(-\frac{\tau_0}{\Sigma_{jj}}\right), \quad \text{where } E(\Sigma_{jj}) = \frac{\tau_0}{v_0 - 1}, \quad V(\Sigma_{jj}) = \frac{\tau_0^2}{(v_0 - 1)^2 (v_0 - 2)}
 \end{aligned} \tag{12}$$

We will use two layers of data augmentation: one for the $\boldsymbol{\alpha}_i$ terms and one for \mathbf{U} , the vector of latent utility differences. The fully augmented joint posterior takes the generic form of

$$\begin{aligned}
 p(\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}_i (i=1 \dots N), \mathbf{U} | \mathbf{y}, \mathbf{X}) &\propto \\
 p(\boldsymbol{\beta}) p(\boldsymbol{\Sigma}) p(\boldsymbol{\alpha}_i | \boldsymbol{\beta}, \boldsymbol{\Sigma}) p(\mathbf{U} | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{X}) p(\mathbf{y} | \mathbf{U})
 \end{aligned} \tag{13}$$

As suggested in Layton and Levine (2003) we do NOT condition draws of \mathbf{U} on $\boldsymbol{\alpha}_i$. We will also draw $\boldsymbol{\beta}$ without conditioning on $\boldsymbol{\alpha}_i$. However, we will need to draw the $\boldsymbol{\alpha}_i$ terms to feed into the draws of $\boldsymbol{\Sigma}$.

Let's take a closer look at $p(\mathbf{U} | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{X})$. For the full sample we can write

$$\begin{aligned}
 p(\mathbf{U} | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{X}) &= \prod_{i=1}^N (2\pi)^{-JT/2} |\mathbf{D}_i \mathbf{V}_i \mathbf{D}_i'|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{U}_i - \mathbf{D}_i \boldsymbol{\mu}_i)' (\mathbf{D}_i \mathbf{V}_i \mathbf{D}_i')^{-1} (\mathbf{U}_i - \mathbf{D}_i \boldsymbol{\mu}_i)\right) = \\
 (2\pi)^{-NJT/2} |\mathbf{D}_i \mathbf{V}_i \mathbf{D}_i'|^{-N/2} \exp\left(-\frac{1}{2} \sum_{i=1}^N (\mathbf{U}_i - \mathbf{D}_i \boldsymbol{\mu}_i)' (\mathbf{D}_i \mathbf{V}_i \mathbf{D}_i')^{-1} (\mathbf{U}_i - \mathbf{D}_i \boldsymbol{\mu}_i)\right) &\quad \text{with} \\
 \boldsymbol{\mu}_i = \mathbf{X}_i \boldsymbol{\beta}, \quad \mathbf{V}_i = \mathbf{X}_{ri} \boldsymbol{\Sigma} \mathbf{X}_{ri}' + \mathbf{I}_{J \times T}
 \end{aligned} \tag{14}$$

The last term of the augmented posterior kernel in (13) can be written as

$$p(\mathbf{y} | \mathbf{U}) = \prod_{i=1}^N \left(\prod_{t=1}^T \left(\sum_{k=1}^J I(y_{it} = k) I\left(\max \left\{ \{U_{ijt}\}_{j=1}^J \right\} = U_{ikt} \right) \right) \right)$$

The explicit form of the augmented posterior kernel can now be written as:

$$\begin{aligned} p(\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}_i (i=1 \cdots N), \mathbf{U} | \mathbf{y}, \mathbf{X}) &\propto \\ \exp\left(-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\mu}_0)' \mathbf{V}_0^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}_0)\right) &* \prod_{j=1}^{k_r} (\Sigma_{jj})^{-(v_0+1)} \exp\left(-\frac{\tau_0}{\Sigma_{jj}}\right) \\ |\boldsymbol{\Sigma}|^{-N/2} \exp\left(\sum_{i=1}^N -\frac{1}{2} \boldsymbol{\alpha}_i' \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha}_i\right) &* \\ (2\pi)^{-NJT/2} |\mathbf{D}_i \mathbf{V}_i \mathbf{D}_i'|^{-N/2} \exp\left(-\frac{1}{2} \sum_{i=1}^N (\mathbf{U}_i - \mathbf{D}_i \boldsymbol{\mu}_i)' (\mathbf{D}_i \mathbf{V}_i \mathbf{D}_i')^{-1} (\mathbf{U}_i - \mathbf{D}_i \boldsymbol{\mu}_i)\right) \\ \prod_{i=1}^N \left(\prod_{t=1}^T \left(\sum_{k=1}^J I(y_{it} = k) I\left(\max \left\{ \{U_{ijt}\}_{j=1}^J \right\} = U_{ikt} \right) \right) \right) \end{aligned} \quad (15)$$

This leads to the following conditional posterior for $\boldsymbol{\beta}$:

$$\begin{aligned} p(\boldsymbol{\beta} | \boldsymbol{\Sigma}, \mathbf{U}, \mathbf{X}) &\propto \\ \exp\left(-\frac{1}{2} \left((\boldsymbol{\beta} - \boldsymbol{\mu}_0)' \mathbf{V}_0^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}_0) + \sum_{i=1}^N (\mathbf{U}_i - \mathbf{D}_i \boldsymbol{\mu}_i)' (\mathbf{D}_i \mathbf{V}_i \mathbf{D}_i')^{-1} (\mathbf{U}_i - \mathbf{D}_i \boldsymbol{\mu}_i) \right)\right) \end{aligned} \quad (16)$$

This is equivalent to the conditional posterior for the generalized regression model. We can immediately derive the conditional posterior moments as:

$$\begin{aligned} \boldsymbol{\beta} | \boldsymbol{\Sigma}, \mathbf{U}, \mathbf{X} &\sim n(\boldsymbol{\mu}_1, \mathbf{V}_1) \quad \text{with} \\ \mathbf{V}_1 &= \left(\mathbf{V}_0^{-1} + \sum_{i=1}^N \mathbf{X}_i' (\mathbf{D}_i \mathbf{V}_i \mathbf{D}_i')^{-1} \mathbf{X}_i \right)^{-1} = \left(\mathbf{V}_0^{-1} + \sum_{i=1}^N \mathbf{X}_i' (\mathbf{D}_i (\mathbf{X}_{ri} \boldsymbol{\Sigma} \mathbf{X}_{ri}' + \mathbf{I}_{J \times T}) \mathbf{D}_i')^{-1} \mathbf{X}_i \right)^{-1} \\ \boldsymbol{\mu}_1 &= \mathbf{V}_1 \left(\mathbf{V}_0^{-1} \boldsymbol{\mu}_0 + \sum_{i=1}^N \mathbf{X}_i' (\mathbf{D}_i \mathbf{V}_i \mathbf{D}_i')^{-1} \mathbf{U}_i \right) = \left(\mathbf{V}_0^{-1} \boldsymbol{\mu}_0 + \sum_{i=1}^N \mathbf{X}_i' (\mathbf{D}_i (\mathbf{X}_{ri} \boldsymbol{\Sigma} \mathbf{X}_{ri}' + \mathbf{I}_{J \times T}) \mathbf{D}_i')^{-1} \mathbf{U}_i \right) \end{aligned} \quad (17)$$

Next, it is convenient to draw the individual random vectors $\boldsymbol{\alpha}_i$. We aim to draw $\boldsymbol{\alpha}_i$

from $p(\boldsymbol{\alpha}_i | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{U}_i, \mathbf{X}_i)$. Note that only data corresponding to individual “ i ” are relevant.

Letting $\tilde{\mathbf{U}}_i = \mathbf{U}_i - \mathbf{D}_i \mathbf{X}_i \boldsymbol{\beta} = \mathbf{D}_i \mathbf{X}_{ri} \boldsymbol{\alpha}_i + \mathbf{D}_i \boldsymbol{\varepsilon}_i$ yields:

$$p(\boldsymbol{\alpha}_i | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{U}_i, \mathbf{X}_i) \propto \exp\left(-\frac{1}{2}\left((\tilde{\mathbf{U}}_i - \mathbf{X}_{ri} \boldsymbol{\alpha}_i)' (\mathbf{D}_i \mathbf{D}_i')^{-1} (\tilde{\mathbf{U}}_i - \mathbf{X}_{ri} \boldsymbol{\alpha}_i) + \boldsymbol{\alpha}_i' \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha}_i\right)\right) \quad (18)$$

The conditional kernel then emerges as

$$\begin{aligned} \boldsymbol{\alpha}_i | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{U}_i, \mathbf{X}_i &\sim n(\boldsymbol{\mu}_1, \mathbf{V}_1) \quad \text{with} \\ \mathbf{V}_1 &= \left(\boldsymbol{\Sigma}^{-1} + \mathbf{X}_{ir}' (\mathbf{D}_i \mathbf{D}_i')^{-1} \mathbf{X}_{ir}\right)^{-1} \\ \boldsymbol{\mu}_1 &= \mathbf{V}_1 \left(\mathbf{X}_{ir}' (\mathbf{D}_i \mathbf{D}_i')^{-1} \tilde{\mathbf{U}}_i\right) = \mathbf{V}_1 \left(\mathbf{X}_{ir}' (\mathbf{D}_i \mathbf{D}_i')^{-1} (\mathbf{U}_i - \mathbf{X}_i \boldsymbol{\beta})\right) \end{aligned} \quad (19)$$

We have to repeat this N times, for each of the N $\boldsymbol{\alpha}_i$'s.

The next step in our GS is the draw of the diagonal terms of the hierarchical variance $\boldsymbol{\Sigma}$.

We obtain

$$\begin{aligned} p(\boldsymbol{\Sigma}_{ij} | \boldsymbol{\alpha}_{ij}) &= ig(v_1, \tau_1) \quad \text{where} \\ v_1 &= \frac{2v_0 + N}{2} \quad \text{and} \quad \tau_1 = \frac{2\tau_0 + \boldsymbol{\alpha}_{ij}' \boldsymbol{\alpha}_{ij}}{2} \end{aligned} \quad (20)$$

For a given individual and choice occasion the conditional posterior for latent utility takes the form of

$$\begin{aligned} p(\mathbf{U}_i | \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{y}_i, \mathbf{X}_{it}) &\propto \\ \exp\left(-\frac{1}{2}(\mathbf{U}_i - \mathbf{D}_i \boldsymbol{\mu}_i)' (\mathbf{D}_i \mathbf{V}_i \mathbf{D}_i')^{-1} (\mathbf{U}_i - \mathbf{D}_i \boldsymbol{\mu}_i)\right)^* & \\ \prod_{t=1}^T \left(\sum_{k=1}^J I(y_{it} = k) I\left(\max\left\{\{U_{ijt}\}_{j=1}^J\right\} = U_{ikt}\right) \right) & \end{aligned} \quad (21)$$

Draws from this density can be obtained via a "Gibbs-within-Gibbs" algorithm as described in Layton and Levine (2003 and 2005). We use 50 iterations for this sub-routine.

Appendix B: Posterior Predictive Draws

Draws from the PPDs of compensating surplus can be obtained as follows:

For each of the R draws of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ from the original Gibbs Sampler perform the following:

1) Draw a vector $\boldsymbol{\alpha}$ of random deviation terms from $f(\boldsymbol{\alpha} | \boldsymbol{\Sigma})$ and compute

$$C_{sp} | \boldsymbol{\alpha}, \boldsymbol{\beta} = -\delta^{-1} (\mathbf{x}'_p \boldsymbol{\beta}_{-p} + \mathbf{x}'_{pr} \boldsymbol{\alpha}).$$

2) Repeat the first step r_2 times for smoothness of the PPD (statistically, this is optional).

In our case, we first thin the original GS by retaining every 10th draw to weaken autocorrelation in the sequence. We then draw 25 $\boldsymbol{\alpha}$ -vectors per original parameter set, yielding a total of 25,000 draws from $p(C_{sp} | \mathbf{y}_s, \mathbf{X}_s)$.

Appendix C: Auxiliary regression model

This model produces the "regression weights" shown in the next table of this document, and ultimately the relative overlap results under the "regression" column in Table 1 of the main paper.

We declare one of the 8 communities as a "target site". We compute the OLR statistic (described in the main paper) for each of the remaining sites, based on their step 1 distributions for the policy-relevant welfare measure. Let y_{jk} be the OLR for a pair of sites j, k . We then compute $\tilde{y}_{jk} = 1 - y_{jk}$ as our dependent variable. This metric is bounded by 0 and 1 (since y_{jk} is bounded by 0 and 1). It takes a value of zero under perfect overlap (i.e. when $y_{jk} = 1$).

On the right hand side we use the distance between j and k , in 10-mile units (let's call it D_{jk}), plus a few aggregate community characteristics. For example, let x_j and x_k be the population density (residents / acre) in communities j and k . We then construct the following regressor:

$\tilde{x}_{jk} = 1 - \left(\min(x_j, x_k) / \max(x_j, x_k) \right)$. Thus, under perfect compatibility of the two sites with respect to this specific characteristic we have $\tilde{x}_{jk} = 0$. This transformation allows for a meaningful regression model without a constant term - the closer distance and \tilde{x}_{ij} are to zero, the closer \tilde{y}_{jk} should be to zero as well. In words: The more "similar" the two sites are based on our chosen community characteristics, the better they should overlap in step 1 welfare densities.

Formally, the doubly-truncated regression model is given as

$$\tilde{y}_{jk} = \tilde{x}_{p,jk} + \tilde{x}_{u,jk} + D_{jk} + \varepsilon_{jk}, I(0 < \tilde{y}_{jk} < 1) \quad \varepsilon_{jk} \sim n(0, \sigma^2)$$

where $\tilde{x}_{p,jk}$ is the aforementioned population density variable and $\tilde{x}_{u,jk}$ is the transformed ratio of the share of urban homes to suburban and rural homes (this was elicited in the survey but could probably be derived through secondary sources as well). We chose this specification after experimenting with numerous other (similar) models. With a sample size of 21 (all possible pairs from 7 sites) we can only include a few regressors.

We estimate this model, via MLE, sequentially for all S target site cases. So for example, in the first run, GT is the target site, and the remaining 7 town feed into the regression model. In the second run, MF is the target and GT returns to the regression sites, etc. For most of these runs, the overall fit is quite good but individual parameter significance is lacking in most cases. This is not unexpected, given the small sample size and the very general nature of our regressors.

The next step is key as it links the regression model to the target site. Assume GT is the target. Let GT be the " j " site. Compute $\tilde{x}_{p,jk}, \tilde{x}_{u,jk}, D_{jk}$ for GT with respect to all other sites. For each case, use the regression results to predict \tilde{y}_{jk} , and convert back to $y_{jk} = 1 - \tilde{y}_{jk}$. Thus we obtain 7 predictions of OLR of the regression sites with GT. Call these predictions $\hat{y}_{jk}, k \neq j$. The mixture weights are then computed as

$$\psi_k = \hat{y}_{jk} / \sum_{k=1}^{S-1} \hat{y}_{jk}, k \neq j. \text{ The intuition is that we would like to allocate more weight to sites}$$

that are expected to have better welfare overlap with the target, based solely on secondary attributes.

After drawing from the resulting mixture density, we compare the resulting distribution of compensating surplus to GT's actual distribution from step 1, based again on the OLR ratio. This is captured in the "regression" column of Table 1 in the main text.

Appendix D: How to draw from the mixture distribution

With empirical weight vector $\boldsymbol{\psi} = [\psi_1 \ \psi_2 \ \dots \ \psi_{S-1}]$ in hand, we can easily draw from

$$p(C_{pp}) = \sum_{s=1}^{S-1} \psi_s p(C_{sp} | \mathbf{y}_s, \mathbf{X}_s) \text{ as follows:}$$

1. Generate a vector $\boldsymbol{\psi}_{sum}$ containing the cumulative sum of $\boldsymbol{\psi}$. Thus, the first element of this vector will be ψ_1 and the last element will be 1.
2. Draw a random uniform term, say u , from the $[0,1]$ interval.
3. If $u < \psi_{sum(1)}$, take a draw from C_{1p} . If $\psi_{sum(1)} < u < \psi_{sum(2)}$, draw from C_{2p} , and so on.

We repeat this process 25,000 times to obtain the same number of draws as we have for the individual welfare distributions.

It is important to note that this is NOT equivalent to averaging draws from the $S-1$ underlying densities. We are not aiming to obtain a weighted average or expectation, but an entire weighted distribution.

Notes

¹ For ease of exposition we choose the same simple linear-in-acres-and-income utility-theoretic framework for all eight sources. This is without loss in generality as our approach could easily accommodate different RUM models for different sets of source studies, including models with non-linear components.

² In actuality, respondents were implicitly asked to commit to an *open-ended stream* of annual payments to preserve a parcel. Thus, a more complete theoretical model would contrast the discounted net present value of expected benefits to discounted costs, perhaps with development weights following some type of survival function. However, this would considerably complicate our analysis and require the stipulation of arbitrary discount rates and survival parameters. Generally, though, the question of how to deal with development risks in CE applications on land preservation posts a strong invitation for future research.

³ 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.

⁴ 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.

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., 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. 2007. "Choice Experiments, Site Similarity, and Benefit Transfer." *Environmental and Resource Economics* 38: 331 - 351.
- Johnston, R. J., J. M. Duke, and J. G. Kulieka. 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., July 27-29.
- Kukielka, J. B., 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., July 27-29.
- 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.
- 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. "Valuing New South Wales Rivers for Use in Benefit Transfer." *The Australian Journal of Agricultural and Resource Economics* 48 (2004): 591-611.
- Morrison, M., and O. Bergland. "Prospects for the Use of Choice Modeling for Benefit Transfer." *Ecological Economics* 60(2006): 420-428.

Table 1: First-step estimation results, Set 1

	Georgetown			Mansfield			Preston			Smyrna		
	mean	std	nse	mean	std	nse	mean	std	nse	mean	std	nse
<u>fixed</u>												
cost(\$10s)	-0.09	0.01	0.00	-0.06	0.01	0.00	-0.07	0.01	0.00	-0.07	0.01	0.00
<u>random means</u>												
acres (10s)	-1.41	0.56	0.04	-0.39	0.30	0.01	-0.86	0.43	0.02	-0.70	0.35	0.01
nursery*acres	-0.74	0.55	0.01	-0.51	0.40	0.01	-0.47	0.61	0.05	-0.46	0.59	0.04
food*acres	-0.32	0.56	0.02	0.22	0.40	0.01	-0.64	0.63	0.06	-0.07	0.66	0.04
dairy*acres	-0.52	0.59	0.03	0.01	0.38	0.01	-0.10	0.49	0.02	0.27	0.46	0.02
forest*acres	-0.60	0.75	0.06	0.16	0.49	0.02	0.07	0.46	0.01	-0.15	0.47	0.02
walking*acres	1.48	0.65	0.05	1.35	0.37	0.02	0.91	0.52	0.02	1.43	0.47	0.02
hunting*acres	1.29	0.57	0.02	0.16	0.38	0.01	0.29	0.49	0.03	0.47	0.40	0.01
<u>random stds</u>												
acres (10s)	2.73	0.61	0.06	2.34	0.37	0.04	2.76	0.50	0.05	1.85	0.46	0.04
nursery*acres	0.94	0.44	0.05	1.05	0.56	0.08	1.87	1.13	0.23	1.94	1.00	0.13
food*acres	1.22	0.77	0.12	1.55	0.82	0.18	2.02	1.09	0.18	3.34	1.65	0.27
dairy*acres	1.27	0.69	0.09	1.54	0.68	0.10	1.42	0.80	0.14	1.23	0.65	0.08
forest*acres	1.91	1.10	0.22	2.89	1.28	0.23	1.01	0.47	0.06	1.57	0.90	0.11
walking*acres	2.54	1.46	0.22	1.00	0.50	0.09	2.49	1.14	0.17	1.50	0.84	0.11
hunting*acres	1.47	0.96	0.14	1.23	0.64	0.10	1.95	0.90	0.11	1.08	0.52	0.06

*nse = numerical standard error

**stds = standard deviations

Table 2: First-step estimation results, Set 2

	Brooklyn			Pomfret			Thompson			Woodstock		
	mean	std	nse	mean	std	nse	mean	std	nse	mean	std	nse
<u>fixed</u>												
cost(\$10s)	-0.07	0.01	0.00	-0.06	0.01	0.00	-0.06	0.01	0.00	-0.04	0.01	0.00
<u>random means</u>												
acres (10s)	-0.74	1.62	0.02	-0.10	1.60	0.02	-1.12	1.60	0.02	-0.20	1.59	0.02
trees*acres	-0.13	1.62	0.02	-0.26	1.60	0.02	-0.46	1.60	0.02	-0.19	1.59	0.02
food*acres	-0.26	1.63	0.02	0.09	1.59	0.02	-0.21	1.59	0.02	-0.02	1.60	0.02
dairy*acres	-0.31	1.63	0.02	-0.04	1.59	0.02	-0.53	1.61	0.02	0.00	1.60	0.02
walking*acres	1.78	0.40	0.02	0.98	0.28	0.01	1.92	0.41	0.02	1.18	0.31	0.01
<u>random stds</u>												
acres (10s)	2.08	0.46	0.03	1.82	0.33	0.02	2.84	1.36	0.08	1.69	0.37	0.02
trees*acres	1.28	0.56	0.05	1.01	0.43	0.05	0.94	0.80	0.06	0.81	0.31	0.03
food*acres	1.33	0.56	0.06	0.92	0.37	0.04	0.82	0.72	0.06	0.85	0.35	0.04
dairy*acres	1.24	0.69	0.08	0.78	0.27	0.03	1.60	1.78	0.26	1.07	0.49	0.05
walking*acres	1.26	0.57	0.07	1.28	0.45	0.04	3.27	2.41	0.20	1.26	0.46	0.04

*nse = numerical standard error

**stds = standard deviations

Table 3: Community Characteristics and Distances*Community Characteristics*

	pop. / acre	homes / acre	fraction urban HHs	fraction rental HHs	average HH size	average yrs. of residence
Georgetown	0.194	0.066	0.099	0.109	2.560	18.660
Mansfield	0.709	0.188	0.085	0.091	2.510	20.110
Preston	0.231	0.094	0.025	0.065	2.690	20.170
Smyrna	0.224	0.080	0.179	0.046	2.750	15.750
Brooklyn	0.387	0.146	0.362	0.106	2.820	19.950
Pomfret	0.147	0.058	0.139	0.073	2.790	19.800
Thompson	0.284	0.119	0.409	0.042	2.691	22.150
Woodstock	0.183	0.077	0.155	0.066	2.730	19.630

Distances in Miles

	Georgetown	Mansfield	Preston	Smyrna	Brooklyn	Pomfret	Thompson
Mansfield	270	-	-	-	-	-	-
Preston	265	27	-	-	-	-	-
Smyrna	44	244	241	-	-	-	-
Brooklyn	281	22	21	258	-	-	-
Pomfret	286	20	28	262	6	-	-
Thompson	292	30	36	269	17	12	-
Woodstock	288	25	39	264	17	13	10

Table 4: Second-stage Weights for Benefit-Transfer Distributions for Selected Sites

	Target = GT			Target = MF	
	regression	distance		regression	distance
MF	0.151	0.084	GT	0.136	0.017
PR	0.136	0.085	PR	0.145	0.173
SM	0.158	0.516	SM	0.133	0.019
BR	0.143	0.080	BR	0.144	0.213
PO	0.140	0.079	PO	0.150	0.234
TH	0.138	0.077	TH	0.143	0.156
WO	0.135	0.078	WO	0.148	0.187
	Target = PO			Target = WO	
	regression	distance		regression	distance
GT	0.131	0.008	GT	0.129	0.011
MF	0.146	0.119	MF	0.152	0.130
PR	0.141	0.085	PR	0.133	0.083
SM	0.132	0.009	SM	0.137	0.012
BR	0.151	0.397	BR	0.148	0.191
TH	0.147	0.198	PO	0.154	0.249
WO	0.151	0.183	TH	0.147	0.324

Table 4: 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							
	regression		distance		uniform		
GT	0.89		0.91		0.88		
MF	0.90		0.88		0.92		
PR	0.86		0.87		0.86		
SM	0.69		0.71		0.70		
BR	0.80		0.81		0.81		
PO	0.89		0.87		0.87		
TH	0.88		0.82		0.89		
WO	0.66		0.65		0.66		

Figure 1: Posterior distribution of compensating surplus, individual models

Vertical lines indicate means

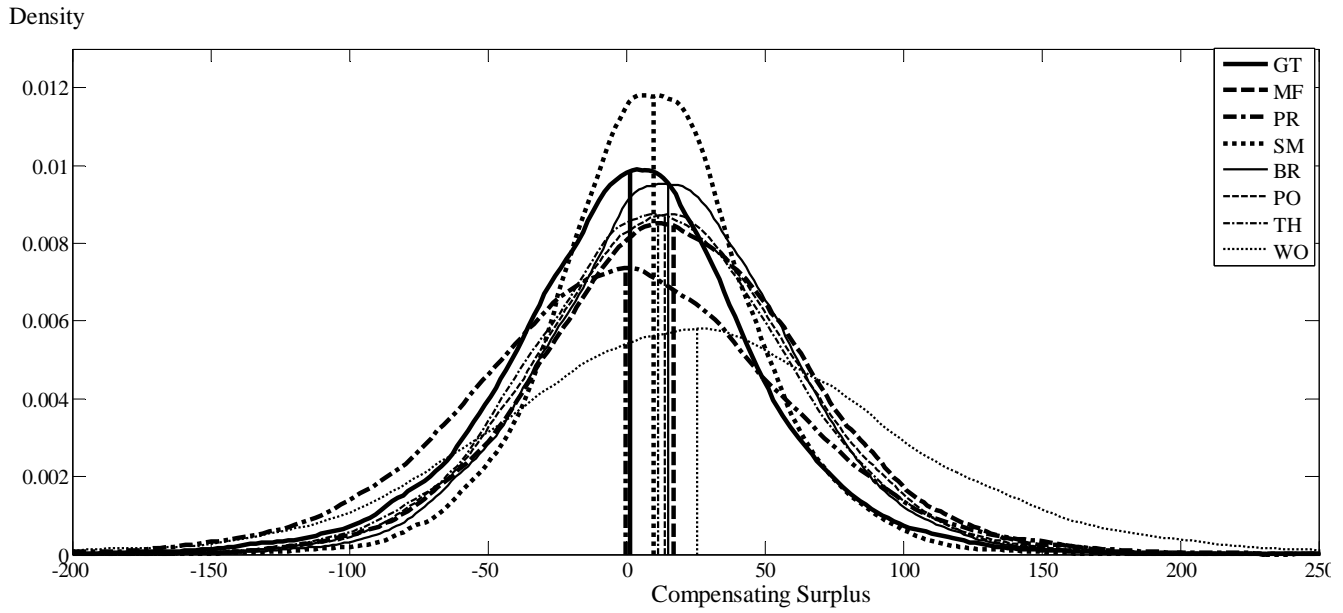
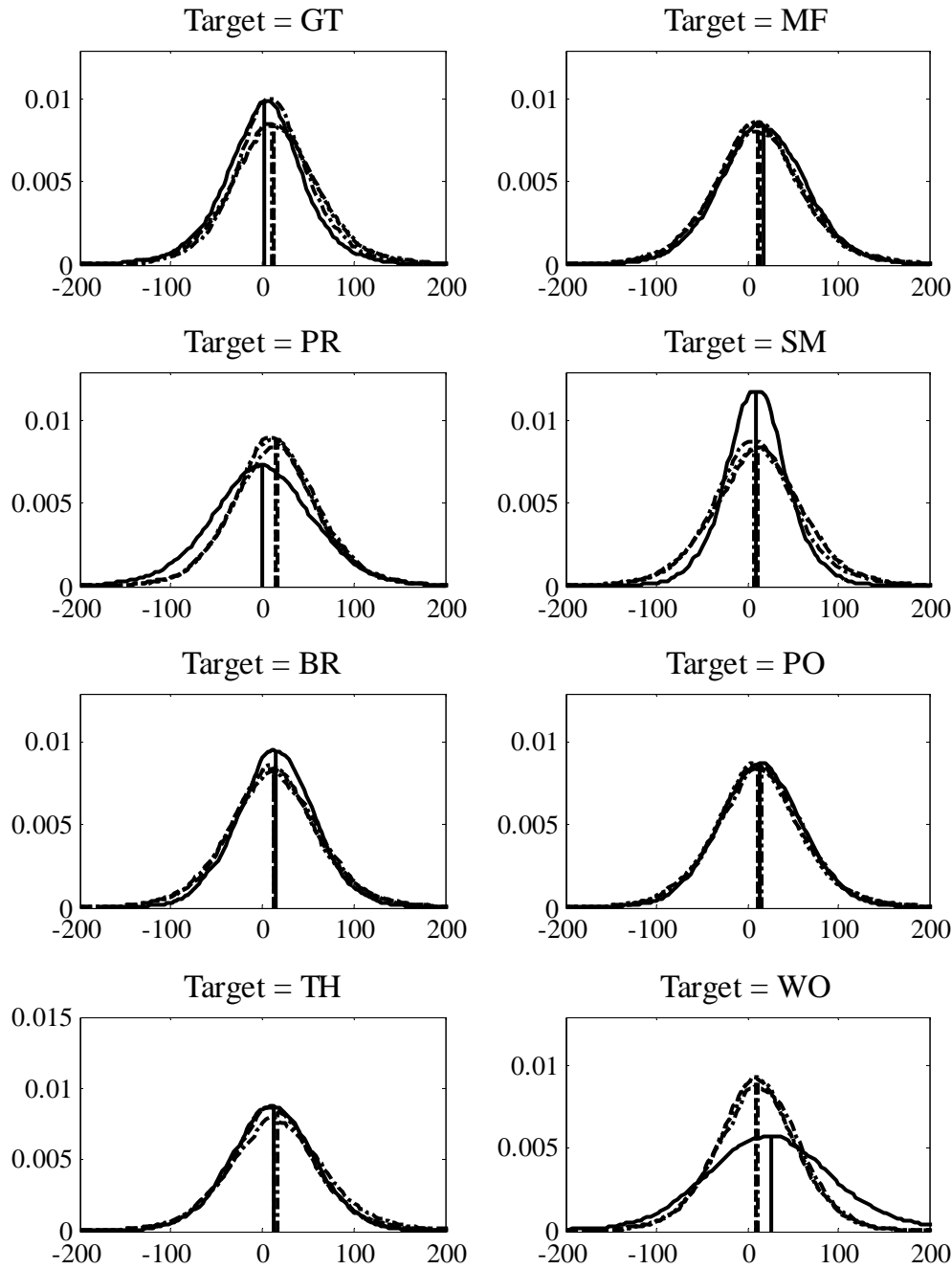


Figure 2: Posterior distribution of compensating surplus, original vs. benefit transfer models



Legend:

solid line:	original model
dashed:	BT via empirical weights
dashed-dotted:	BT via distance weights
dotted:	BT via uniform weights