Valuing Farmland Protection with Choice Experiments That Incorporate Preference Heterogeneity: Does Policy Guidance Depend On the Econometric Fine Print?

Robert J. Johnston
Department of Economics
Clark University

John C. Bergstrom
Department of Agricultural and Applied Economics
The University of Georgia

Contact Information:
Robert J. Johnston
George Perkins Marsh Institute
Clark University
950 Main St.
Worcester, MA 01610
Phone: (508) 751-4619
Email: rjohnston@clarku.edu

Valuing Farmland Protection with Choice Experiments That Incorporate Preference Heterogeneity: Does Policy Guidance Depend On the Econometric Fine Print?

Abstract

Although mixed logit models are common in stated preference applications, resulting welfare estimates can be sensitive to minor changes in specification. This can be of critical relevance for policy and welfare analysis, particularly if policymakers are unaware of practical implications. Drawing from an application to agricultural conservation in Georgia, this paper quantifies the sensitivity of welfare estimates to common variations in mixed logit specification and assesses practical implications for policy guidance. Results suggest that practitioners may wish to reevaluate modeling and reporting procedures to reflect the welfare and policy implications of common but often unnoticed variations in model specification.

I. INTRODUCTION

Stated preference (SP) welfare analysis of farmland preservation policies, including purchase of conservation easement (PACE) programs, can be complicated by substantial preference heterogeneity across respondents. Evidence suggests that values for identical preservation policies can vary both in magnitude and sign across individuals (e.g., Bergstrom and Ready 2009; Johnston et al. 2001).[1] Choice experiment analyses of farmland and rural landscape preservation in US (Johnston and Duke 2007, 2008, 2009) and non-US (Campbell 2007; Columbo et al. 2007; Columbo and Hanley 2008; Scarpa et al. 2007, 2009) contexts suggest similar preference variation. Such patterns, along with shortcomings of traditional multinomial logit models (Train 2009), have led to the rapid adoption of mixed logit (ML) models and other methods better able characterize preference heterogeneity in these and other policy contexts.[2]

Although ML models are now common in valuation applications, scrutiny of these models reveals a variety of often-overlooked concerns. These include potential sensitivity of welfare estimates to minor changes in model specification, including the assumed distribution of random parameters (Balcombe et al. 2009; Hensher and Greene 2003). Johnston and Duke (2007) identify such patterns in a case study of farmland preservation, although they do not quantify implications for welfare estimates.

Despite the potential sensitivity of welfare estimates to minor changes in model specification, "many of the ML applications in the environmental and resource economics

---

[1] Estimated WTP for farmland preservation in the US can vary according to numerous policy and methodological factors, including farmland type (Johnston and Duke 2007; Bergstrom and Ready 2009), the jurisdiction in which preservation occurs (Johnston and Duke 2009), the welfare measures considered or valuation methods applied (Johnston et al. 2001; Ready et al. 1997; Boyle and Ozdemir 2009), attributes of the policy process (Johnston and Duke 2007), and many other factors. As related welfare patterns have received considerable attention in the literature, they are not emphasized here.

[2] The two most common means of introducing individual preference heterogeneity to discrete choice modeling are mixed logit (ML) and latent class models, neither of which has been shown to provide broadly superior forecasts (Provencher and Bishop 2004). Both approaches overcome many of the behavioral limitations of traditional multinomial logit models (Hensher and Greene 2003; Train 2009).

literature have paid insufficient attention to the implications of model choice" (Balcombe et al. 2009, p. 226). As noted by Layton and Lee (2006), common applications calculate willingness to pay (WTP) for only a single or small number of model specifications chosen on the basis of various statistical tests and unpublished preliminary models, with little indication of the robustness concerns discussed by such authors as Hensher and Greene (2003), Balcombe et al. (2009), Scarpa et al. (2009) and others. This can be of critical relevance for policy analysis, particularly if welfare estimates and policy guidance are sensitive to the "econometric fine print," and if policymakers are unaware of this sensitivity. This issue is of greater relevance for ML models than for older forms of SP estimation, both because of the increased flexibility of these models, the potential sensitivity of results to small and perhaps unnoticed variations (Hensher and Greene 2003), and because of the rapidly increasing number and complexity of possibilities for model estimation (Layton and Lee 2006).[3]

The published literature addressing such concerns tends to emphasize methodological aspects of the problem, including novel approaches for model selection (Balcombe et al. 2009), averaging (Layton and Lee 2006), and estimation (Train and Weeks 2005). However, given the widespread and increasing use of ML results for policy guidance, applied policy implications are also crucial. Specifically, is the valuation literature conveying confidence in WTP point estimates that is unwarranted given the potential sensitivity of these estimates to minor changes in ML specifications? Lack of broader insight into such issues leaves practitioners with uncertainty regarding the confidence that should be placed in reported welfare estimates, particularly when based on results from a single model. Such issues may be particularly relevant for policy contexts such as farmland preservation, for which past work has shown significant

---

[3] Some of the challenges leading to WTP sensitivity can be ameliorated by solutions such as modeling preferences in WTP space (Train and Weeks 2005; Scarpa et al. 2007), but such approaches remain rare, in part due to "mixed results regarding the appropriateness of undertaking estimation in WTP space" (Balcombe et al. 2009, p. 228).

implications of preference heterogeneity for welfare estimates (Johnston et al. 2001; Johnston and Duke 2007, 2009; Campbell 2007; Columbo et al. 2007; Columbo and Hanley 2008).

This paper quantifies the sensitivity of welfare estimates and practical policy guidance to common and often minor variations in ML specification. The analysis purposefully emphasizes specifications common in the literature and often used for policy guidance. We emphasize that the purpose of this paper is *not* to advance methods for ML modeling or model selection. Rather, the goal is to address practical policy implications of increasingly ubiquitous welfare estimation methods. The analysis is based on a choice experiment application to agricultural conservation easement programs in Georgia, an application typical of the context for which preferences are likely heterogeneous and for which ML estimation is warranted. Results show that differences in welfare estimates across model specifications for the same data and site can be of similar magnitude to welfare differences found *across different sites* in the benefit transfer literature—highlighting the relevance of the challenge. These and other results suggest that practitioners may wish to reevaluate widespread modeling and reporting procedures to reflect the potential implications of common but often unnoticed variations in model specification. Results also show, however, that some aspects of policy guidance are more robust.

II. A RANDOM UTILITY MODEL OF CONSERVATION EASEMENT CHOICES

We begin with a simple choice model where a choice is made between alternative levels of a commodity, in this case either a PACE program, which we will refer to as Program j, or a baseline level of conservation (no program or the status quo), denoted by Program 0. The respondent chooses Program j over Program 0 if

$$U_j(\mathbf{z_j}, I - P_j) \ > \ U_0(\mathbf{z_0}, I), \qquad\qquad [1]$$

where $U_j(\cdot)$ and $U_0(\cdot)$ are utility levels attainable with (vector-valued) program characteristics $\mathbf{z_j} \neq 0$ and $\mathbf{z_0} = 0$, P_j is the price of $\mathbf{z_j}$, and I is income.[4] Within the familiar random utility context, $U_n(\mathbf{z_n}, I - P_n) = v_n(\mathbf{z_n}, I - P_n) + \varepsilon_n$ for $n = \{0, j\}$, where the observable component of utility is given by $v_n(\cdot)$, the unobservable component by ε_n, and the price of the status quo $P_0 = 0$. Following standard approaches, the ε_n are assumed to be stochastic utility shocks with a known distribution, modeled as random errors within the econometric model.

If the observable component of indirect utility can be approximated as a separable and linear function $v(\mathbf{z_n}, I - P_n) = \mathbf{z_j'}\boldsymbol{\alpha} + \gamma(I - P_n)$, the probability that one will choose Program j over Program 0 is given by

$$\Pr(U_j(\cdot) > U_0(\cdot)) = \Pr((\mathbf{z_j'}\boldsymbol{\alpha} - \gamma P_j) > (\varepsilon_0 - \varepsilon_j)) = F_{\varepsilon_0 - \varepsilon_j}(\mathbf{z_j'}\boldsymbol{\alpha} - \gamma P_j), \qquad [2]$$

where $F_{\varepsilon_0 - \varepsilon_j}$ is the distribution function of the difference between shocks ε_0 and ε_j. If the ε_n are assumed independently and identically drawn from a type I extreme value distribution, model coefficients $\boldsymbol{\beta} = \{\boldsymbol{\alpha}, \gamma\}$ may be estimated using the standard conditional logit (CL) model. Equating utility levels and solving for price yields a compensating surplus measure, or an individual's WTP for Program j:

$$\text{WTP} = \frac{\mathbf{z_j'}\boldsymbol{\alpha}}{\gamma} \qquad [3]$$

The simplicity of the CL model rests on the assumption that individuals both in the sample and population have identical characteristics; that is, for all individuals i, $\boldsymbol{\beta_i} = \boldsymbol{\beta} \; \forall i \in \Omega$. This assumption is unrealistic for many contexts, including that of farmland preservation. The imposition of a constant parameter vector across the population can also suppress variations in welfare relevant to policy development. Moreover, strong assumptions regarding the distribution

---

[4] We assume that the presence or absence of a farmland protection program is unlikely to significantly change one's consumption of other market goods or services, and that hence the modeled utility-optimizing behavior is local. Therefore consumption levels of other goods is considered constant.

of utility shocks within the CL model disallow correlations among multiple survey responses provided by the same individual, a common feature of SP surveys in the literature.

A mixed model relaxes the assumption of respondents being identical, replacing it with less restrictive assumption of respondents being identically distributed. This implies that $\beta_i \sim D(\theta) \ \forall \ i \in \Omega$, where $D(\theta)$ is a multivariate distribution of a known form. If the distribution is a generalized extreme value distribution, the model may be estimated as a ML model. The general theory and methods of ML modeling are well-established (Train 2009). The most common form is the random coefficients model (McFadden and Train 2000), which the decisionmaker's objective function is linear in coefficients and shocks are independently and identically distributed (IID). The framework is easily extended to accommodate panel data, i.e., multiple correlated responses per individual. Additional advantages include a less constrained utility structure that does not impose the often problematic independence from irrelevant alternatives (IIA) property on choice probabilities.

III. SPECIFICATION CHOICES IN MIXED LOGIT

ML models allow for coefficients on attributes to be distributed across sampled individuals according to a set of estimated coefficients and researcher-imposed restrictions. Even within the linear, additive, main effects utility specifications common in SP modeling, model estimation requires that researchers determine which coefficients in the utility functions should be randomized and the distributions that characterize these coefficients. This includes specification of bounds on coefficient distributions and the form of correlations, if any, between random coefficients. Although some specifications allow for randomization of the entire coefficient vector, in practice researchers often randomize only a subset of coefficients. While statistical

tests can in some cases help researchers determine initial candidate sets of random coefficients (McFadden and Train 2000), these tests are of low power with difficult-to-retrieve critical values (Hu et al. 2006). Moreover, it is possible that competing model specifications could have similar statistical fit, yet generate divergent welfare estimates (Layton and Lee 2006).

As a result, determining specifications for ML models involves both quantitative assessments and researcher judgment. Still additional choices are made when simulating welfare estimates from model results (Hensher and Greene 2003), a challenge reported in a farmland preservation context by Johnston and Duke (2007). Although the practical relevance of these and other choices are often overlooked by the applied valuation literature, effects on welfare estimates can be substantial (Balcombe et al. 2009; Hensher and Greene 2003). While some authors have promoted various forms of model averaging as a solution, these are not commonly found in applied analysis, perhaps because of fear among practitioners that such approaches will lead to a perception of results as "fragile" or "arbitrary" (Layton and Lee 2006, p. 80).

Compounding these challenges, some of the most common specification choices can contribute to substantial challenges and uncertainties for welfare estimation (Hensher and Greene 2003). An example is the specification of the coefficient on program cost within choice experiments (Hensher and Greene 2003; Train and Weeks 2005; Scarpa et al. 2007). Although the magnitude of this coefficient may vary across individuals, it is nonetheless expected to have an unambiguously negative sign, reflecting a positive marginal utility of income. To ensure an appropriate sign for this coefficient within ML models, a common solution is to specify a lognormal distribution on the sign-reversed cost coefficient. This solution, however, leads to well-known ambiguities for WTP estimation related to the long right-hand tail of the lognormal distribution, and often unrealistic mean WTP estimates over the unconstrained distribution

6

(Hensher and Greene 2003; Train and Weeks 2005; Scarpa et al. 2007). Given these challenges, one may truncate the distribution (e.g., Hensher and Greene 2003) estimate median (instead of mean) WTP over the lognormal distribution (Hu et al. 2005; Johnston and Duke 2007), estimate models in WTP-space (Train and Weeks 2005; Scarpa et al. 2007), use alternative distributions for the cost coefficient Hensher and Greene (2003)[5], or model this parameter as non-random (e.g., Layton and Brown 2000), but each of these options has potential shortcomings. Even though this choice is only one of many facing ML modelers, it alone can result in substantial variation in welfare estimates (Hensher and Greene 2003; Layton and Lee 2006).

Despite these and other challenges, most works in the literature display only one, or a very small number of ML model specifications, with little or no discussion of the robustness of model results to changes in assumptions regarding coefficient distributions, bounds or correlations. While such considerations might be unimportant when the primary purpose of the analysis is methodological, or for policy contexts in which preference heterogeneity is inconsequential, they could be of crucial relevance in contexts such as agricultural land preservation, particularly when applied welfare estimates are used for policy guidance.

IV. THE DATA

[5] An alternative recommended by Hensher and Greene (2003, p. 146-148) is the bounded triangular distribution. The triangular distribution is characterized as follows: "the … density function looks like a tent: a peak in the centre and dropping off linearly on both sides of the centre. Let c be the centre and s the spread. The density starts at $c - s$, rises linearly to c, and then drops linearly to $c + s$. It is zero below $c - s$ and above $c + s$. The mean and mode are c. The standard deviation is $s/\sqrt{6}$; hence the spread is the standard deviation times $\sqrt{6}$. The height of the tent at c is $1/s$ [and] the slope is $1/s^2$." If one specifies the distribution of a positive ML parameter with a triangular distribution bounded such that the mean (m) is equal to the spread (s), then "the density starts at zero, rises linearly to the mean, and then declines to zero again at twice the mean. It is peaked, as one would expect. It is bounded below at zero, bounded above at a reasonable value that is estimated, and is symmetric such that the mean is easy to interpret." For negative parameters the sign is reversed, but the shape of the distribution and interpretation remain identical.

The empirical model seeks to quantify the practical implications of ML model specification for a welfare assessment involving government-sponsored PACE programs. The goal of the analysis is the provision of heretofore unavailable welfare guidance for agricultural policy, emphasizing the role and relevance of preference heterogeneity as quantified using various modeling options. A secondary and broader goal is to illustrate the types and magnitude of welfare impacts that may be associated with seemingly minor and often unnoticed choices in ML model specification.

The data are drawn from the SP survey entitled *Purchasing Conservation Easements to Agricultural Land in Georgia,* funded through a multi-state USDA National Research Initiative grant.[6] Details of the survey and research effort can be found in Boyle and Özdemir (2009), Paterson et al. (2004), Özdemir (2003) and Özdemir et al (2004). Details of the multi-state PACE survey development and testing, including the 12 focus groups used in survey design, are summarized by Boyle and Ozdemir (2009). Insight from focus groups helped to ensure that the survey language and format could be easily understood by respondents, that respondents shared interpretations of survey terminology and scenarios, and that the scenarios captured policy attributes viewed as relevant and realistic by respondents. Focus groups also led to a self-administered mail survey, following a choice experiment framework (Adamowicz et al. 1998).

Prior to the administration of choice questions, respondents were instructed to consult an information booklet that provided background information on PACE programs in Georgia, as well as detailed instructions included in the survey booklet. Each choice question then asked respondents to consider two hypothetical statewide PACE programs—Program A and Program B—that would each preserve specified quantities of land with varying programmatic priorities for farmland use, location, and land quality. Each question also included a status quo response

---

[6] The overall USDA National Research Initiative (NRI) grant project was entitled, "Improved Information in Support of a National Strategy for Open Land Policies", Kevin J. Boyle (Project Leader) with Mary Ahearn, Anna Alberini, John C. Bergstrom, Lawrence W. Libby, Michael P. Welsh (Project Cooperators).

8

option, neither program, that involved no additional preservation. Each respondent was provided

with four independent choice questions. Table 1 shows design attributes and levels that

distinguished the presented conservation easement programs. As discussed by Boyle and

Ozdemir (2009), choice questions were constructed using random assignment of attribute levels.[7]

The Georgia PACE survey was mailed to a random sample of 1,000 Georgia households in the

spring of 2002, with follow-up mailings following Dillman (2000). A total of 213 completed

questionnaires were returned. After adjusting for undeliverable questionnaires due to bad

mailing addresses (175) and people who returned the questionnaire blank (19), the effective

response rate to the survey was 26.4%. This response rate was similar to parallel PACE surveys

conducted in Maine and Ohio under the original NRI grant referenced above.

V. THE EMPIRICAL MODEL

As the purpose of this analysis is to characterize the potential impact of common variations in

ML specification for model estimates and policy guidance, we follow traditions of the literature

and illustrate a range of specifications consistent with the types of models which often appear.

Our purpose is not to identify the correct or best model, but rather to characterize variations in

results that can result from commonly applied specifications. As a foundation for subsequent

modeling, we begin with a baseline empirical specification in which utility is given by

$$v_j(\cdot) = \delta(ASC_0) + \boldsymbol{\alpha}(\mathbf{z_j}) + \gamma(P_j) + \boldsymbol{\mu}(\mathbf{D})(ASC_0), \qquad [4]$$

where ASC_0 is the alternative specific constant for the no program status quo, $\mathbf{z_j}$ is a vector of

---

[7] This experimental design strategy was selected given the emphasis placed by the funding agency on estimating interaction effects between attributes. There has been substantial emphasis, and some controversy, in the more recent literature over the suitability of random designs (Lusk and Norwood 2005), with some researchers pointing out possible risks in such designs (Carson et al. 2009). Lusk and Norwood (2009) argue that there are instances in which random designs may be more appropriate than conventional fractional factorial design strategies. Readers should consider results of the present choice experiment in light of the random assignment design that was applied.

non-price program characteristics such as the number of acres and type of land preserved, P_j is the unavoidable price or household cost of the program entered with sign-reversal (Hensher and Greene 2003)[8], $\mathbf{D}$ is a vector of socioeconomic attributes, and the conforming vector of coefficients to be estimated is given by $\boldsymbol{\rho} = [\delta, \boldsymbol{\alpha}, \gamma, \boldsymbol{\mu}]$. The subscript j references unique preservation options. Following standard practice, the inclusion of interactions with $\mathbf{D}$ allows baseline utility under the status quo to vary according to socioeconomic attributes.

From [4] a variety of ML models may be derived, each with different distributions of random coefficients over individual respondents i. For the full set of coefficients in the model, $\boldsymbol{\rho}$, these may be characterized by the general form $\boldsymbol{\rho_i} = \boldsymbol{\rho} + \boldsymbol{\Gamma v_i}$, where $\boldsymbol{\rho}$ is the vector of population means, $\mathbf{v_i}$ is a vector of individual heterogeneity components with a mean of zero and standard deviation of one, and $\boldsymbol{\Gamma}$ is a triangular matrix with coefficient standard deviations σ_k on the diagonal. Off-diagonal elements of $\boldsymbol{\Gamma}$ are zero when random coefficients are uncorrelated; with free correlation $\boldsymbol{\Gamma}$ becomes an unrestricted triangular matrix (Greene 2007). This specification encompasses numerous variants. Of these, seven specifications are chosen for estimation here, selected to exemplify models common in the literature:

Model I:
$\delta_i = \delta;$
$\alpha_i = \alpha;$
$\gamma_i = \gamma;$
$\mu_i = 0;$
$\boldsymbol{\Gamma} = 0;$

Model II:
$\delta_i = \delta + \sigma_\delta v_i,\ v_i \sim N[0,1];$
$\alpha_i = \alpha + \sigma_\alpha v_i,\ v_i \sim N[0,1];$
$\gamma_i = \gamma;$
$\mu_i = 0;$
$\boldsymbol{\Gamma} = \mathrm{diag}(\sigma_1, \sigma_{2\,...,}, \sigma_{K)};$

Model III:
$\delta_i = \delta + \sigma_\delta v_i,\ v_i \sim N[0,1];$

$$\alpha_i = \alpha + \sigma_\alpha v_i, \; v_i \sim N[0,1];$$
$$\gamma_i = \gamma;$$
$$\mu_i = 0;$$
Γ unrestricted (i.e., correlated random coefficients);

Model IV:
$$\delta_i = \delta + \sigma_\delta v_i, \; v_i \sim N[0,1];$$
$$\alpha_i = \alpha + \sigma_\alpha v_i, \; v_i \sim N[0,1];$$
$$\gamma_i = \gamma + \gamma v_i, \; v_i \sim triangle[-1,1];$$
$$\mu_i = 0;$$
$$\Gamma = \text{diag}(\sigma_1, \sigma_2 \dots, \sigma_K);$$

Model V:
$$\delta_i = \delta + \sigma_\delta v_i, \; v_i \sim N[0,1];$$
$$\alpha_i = \alpha + \sigma_\alpha v_i, \; v_i \sim N[0,1];$$
$$\gamma_i = \gamma + \gamma v_i, \; v_i \sim triangle[-1,1];$$
$$\mu_i = \mu;$$
$$\Gamma = \text{diag}(\sigma_1, \sigma_2 \dots, \sigma_K);$$

Model VI:
$$\delta_i = \delta + \sigma_\delta v_i, \; v_i \sim N[0,1];$$
$$\alpha_i = \alpha + \sigma_\alpha v_i, \; v_i \sim N[0,1];$$
$$\gamma_i = exp(\gamma + \sigma_\gamma v_i), \; v_i \sim N[0,1];$$
$$\mu_{hi} = 0;$$
$$\Gamma = \text{diag}(\sigma_1, \sigma_2 \dots, \sigma_K);$$

Model VII:
$$\delta_i = \delta;$$
$$\alpha_i = \alpha;$$
$$\gamma_i = exp(\gamma + \sigma_\gamma v_i), \; v_i \sim N[0,1];$$
$$\mu_i = 0;$$
$$\Gamma = \text{diag}(\sigma_1, \sigma_2 \dots, \sigma_K);$$

Differences between the seven models may be described succinctly. Model I is a standard CL model. Models II and III are ML models with parameters on the ASC and on all non-price program attributes specified as random with a normal distribution. Model II allows no correlation among random parameters, while Model III allows free correlation. All of these models include a non-random coefficient on program cost and no interactions between the ASC and socioeconomic attributes. Model IV is identical to II, save that program cost is random with a bounded triangular distribution following Hensher and Greene (2003)[9] Model V is similar to IV, but also includes interactions between socioeconomic attributes and the ASC, thereby modeling

---

[9] Additional applications of this distribution are cited by Balcombe et al. (2009). The illustrated specification constrains the parameter to lie between 0 and 2γ.

preference heterogeneity using both fixed and random effects. Model VI is also similar to IV,

except that program cost is assumed to have a lognormal distribution rather than triangular.

Finally, Model VII is identical to I, except that program cost is considered to be random with a

lognormal distribution. These features are summarized in table 2.

All specifications are consistent with theory. For example, all ML models constrain the

sign of the parameter on sign-reversed program cost (γ_i) to the positive domain, as suggested by

economic theory. In addition, all models reflect specifications common in the applied valuation

literature. All ML models are estimated using simulated maximum likelihood with 100 Halton

draws per likelihood simulation; the CL model is estimated using maximum likelihood.

VI. RESULTS AND IMPLICATIONS

Table 3 defines independent variables included in the estimated models. Table 4a,b presents

model results. All models are significant at $p<0.0001$ based on likelihood ratio tests. Measures of

model fit suggest reasonable performance of each model relative to prior models found in the

literature, with significant coefficients and value surfaces conforming to prior expectations.

General patterns also comport with CL results of the parallel PACE survey conducted in Maine

under the same NRI grant (Boyle and Ozdemir 2009). Although some models outperform others

according to standard statistical tests, all provide reasonable estimates of the type that are

commonly used for policy guidance.

Following traditions in the recent literature, we emphasize implications for both implicit

prices and compensating surplus, each calculated using standard approaches (cf., Morrison et al.

2002; Hanley et al. 2006).[10] For the CL model, welfare measures are calculated using standard

---

[10] Comparison of raw parameter estimates across models provides limited insight, as these are confounded with
scale parameters. Although there are established methods for testing the equivalence of parameters in logit models

approaches, with empirical confidence intervals and significance levels simulated following Poe et al. (2005). Because ML models include random coefficients, we estimate these measures using the welfare simulation of Johnston and Duke (2007), following standard methods illustrated by Hensher and Greene (2003) and Hu et al. (2005), among others. The procedure begins with a parameter simulation following the parametric bootstrap of Krinsky and Robb (1986), with R=500 draws taken from the mean parameter vector and associated covariance matrix. For each draw, resulting parameters are used to characterize asymptotically normal empirical densities for fixed and random parameters. For each of these R draws, a coefficient simulation is then conducted for each random coefficient, with S=200 draws taken from simulated empirical densities.[11] Here, all coefficient simulations draw from a normal distribution except that on *cost*, which draws from either a lognormal or bounded triangular distribution. Welfare measures are calculated for each draw, resulting in a combined empirical distribution of $R \times S$ observations from which summary statistics are derived. These distributions accommodate both the sampling variance of parameter estimates and the distribution of random parameters.

For models with a non-random or bounded triangular distribution on the cost coefficient, we simulate welfare estimates as the mean over the parameter simulation of mean WTP calculated over the coefficient simulation (mean of mean WTP). Models with the coefficient on cost distributed lognormal, with WTP simulated as the mean over the parameter simulation of *mean* WTP over the coefficient simulation, generated implausible WTP estimates and are hence excluded. Similar findings have been reported by numerous authors, including Johnston and Duke (2007) and Hensher and Greene (2003). Hence, for these models, we follow Johnston and

---

(Swait and Louviere 1993), these are not applied here given our emphasis on implications for policy guidance, which tends to rely on welfare estimates.

[11] One may also conduct these simulations with larger numbers of draws (Johnston and Duke 2007, 2009), but affects on model results in the present case are trivial.

Duke (2007) and Hu et al. (2005) and simulate welfare estimates as the mean over the parameter simulation of *median* WTP calculated over the coefficient simulation (mean of median WTP).

Sensitivity of Implicit Prices

Table 5a,b illustrates resulting implicit price estimates from each model together with significance levels and 90% confidence bounds. Table 5b also reports the percentage difference between the lowest and highest implicit price for each attribute in any of the seven models. As shown by these results, implicit prices are at least moderately sensitive to changes in model specifications and in some cases vary to a substantial degree. For statistically significant implicit prices, the largest differences tend to involve models with lognormal distributions on the cost parameter (Models VI and VII). This is of particular note given the ubiquity of such specifications in the valuation literature. In contrast, implicit price estimates are often robust across Models I through IV.

There are exceptions, however, and most models include at least one idiosyncratic result. Some of these are highly relevant for policy guidance. For example, implicit price distributions from Model III are wider than those found in other models, leading to reduced significance levels. In two cases, implicit prices significant at $p<0.05$ in all other models (i.e., *Vegetable, Hay*) are not statistically significant in Model III. Similar patterns are found for other attributes and models. For example, the implicit price for *Acres* is robust across Models I through V at values between 0.122 and 0.147. This value, however, drops to 0.060 in Model VI and rises to 0.255 in Model VII (a change of 325%). Over all attributes and models, percentage differences between the lowest and highest implicit price range from 62.23% to 782.12%, a range similar to errors found in the benefit transfer literature when considering transfers across *different sites* (cf.

14

Rosenberger and Stanley 2006). Here we find similar variations across different models used to estimate implicit prices from the same site and survey data.

Taken together, these results suggest that while implicit price estimates can be robust across CL and ML model specifications, large differences also occur. These differences can result from modest changes to model specification. For example, the sole difference between Models IV and VI is the assumed distribution of the parameter on program cost (bounded triangular vs. lognormal), together with associated differences in the welfare simulation. As a result of this change, the implicit prices on *Acres* and *Location* both decrease by more than 50%, while the implicit price on *Vegetable* increases by approximately 30%. Such results are not entirely surprising given prior findings that a lognormal distribution on program cost can have substantial implications for WTP (e.g., Hensher and Greene 2003). Nonetheless, it is notable that these two ML model specifications, both common in the valuation literature, lead to such divergent conclusions. These differences can lead to substantial increases in some estimated implicit prices, while decreasing others, still further complicating potential policy implications.

The illustrated implicit price variations are likely to be most relevant when policy deliberations focus on the cardinal costs and benefits associated with marginal changes in PACE program design, e.g., whether the additional benefits of emphasizing prime farmland justify the costs. In contrast, the ordinal ranking of implicit prices is largely (although not universally) robust across specifications. For example, results over a wide range of specifications agree that a priority on land used for growing vegetables, berries, fruits and nuts (*Vegetable*) is among the most important positive attributes of a PACE program, while a priority on hay fields (*Hay*) is the most important negative aspect. Hence, for the case of implicit prices, the practical relevance of the illustrated welfare sensitivity depends largely on the intended use of model results.

15

Sensitivity of Compensating Surplus

Notwithstanding the considerable insight available from implicit prices, compensating surplus (CS) assessments arguably provide a more relevant perspective for policy purposes, as benefit cost analyses commonly rely on these measures (Morrison et al. 2002). Here, we assess CS differences for cases in which each of the seven models is treated as the "baseline" from which absolute value percentage differences are calculated, compared to analogous estimates from other "comparison" models. Given attribute levels in the experimental design (table 1), there are 96 unique combinations of conservation easement attributes for which welfare measures may be estimated.[12] Following convergent validity testing methods from the benefit transfer literature, we report the average absolute value difference between parallel CS measures over all 96 possible attribute combinations (Rosenberger and Stanley 2006). Differences are presented as trimmed mean (5%) absolute value percentages over all 96 policy options.

Results are shown in Table 6. Illustrated CS differences provide a mixed message for the sensitivity of welfare estimates to changes in model specification. Perhaps most noticeable is the large divergence of CS estimates derived from the CL model compared to those from ML specifications. These large differences, combined with the more restrictive and sometimes unrealistic assumptions imposed by the former (Train 2009), suggest that researchers may wish to exercise caution when drawing policy conclusions from CL models. Here, the average absolute value percentage difference between CS estimates across models is 126.72% when including all CL and ML model pairs. When one excludes the CL model from comparisons, this difference drops to 27.67%. That is, CS estimates from different ML models are much more

---

[12] This number is derived by including all combinations of the acreages (4), locations (2), quality levels (2) and land types (6) possible given attribute levels in the experimental design.

similar, on average, than those from CL and ML models.

Whether the more modest average differences between CS estimates from competing ML specifications (ranging from 9.77% to 60.51%, and averaging 27.67%) are of concern for policy applications is context dependent, as the required reliability and/or robustness for welfare measures depends on the level of accuracy deemed necessary in a given policy context (Ben-Akiva, 1981; Bergstrom and De Civita, 1999; Johnston and Rosenberger 2010). It has been suggested that acceptable accuracy is a function of the precision necessary for different types of decisions (Navrud and Pruckner 1997). For example, higher degrees of precision may be needed as one moves from broad cost-benefit analyses for information gathering or project screening to calculation of compensatory amounts in negotiated settlements and litigation (Rosenberger and Johnston 2009). Given the variations in average CS estimates shown in Table 6, it would seem that these estimates would be most suitable for use in cases where somewhat broader welfare guidance—as opposed to highly precise point estimates—are required. Again, the percentage differences between models often rival those found in the benefit transfer literature when comparing welfare estimates from different studies at different sites—highlighting the potential magnitude and policy relevance of implications for welfare estimates.

Implications for Policy Prioritization

In cases where decision-makers seek monetized benefit measures, comparative welfare assessments such as those above can provide guidance concerning expected bounds of sensitivity analysis. In other cases, however, including many relevant to farmland preservation, policymakers may desire only a welfare grounded means to prioritize preservation options (Duke and Aull-Hyde 2002; Jiang, Swallow and McGonagle 2005; Messer 2006). In such cases, the

emphasis is not on the convergent validity of welfare estimates but rather on the transferability of resultant policy rankings—results that may remain robust even if welfare estimates fail convergent validity tests or are associated with large variation across models. Such issues are of particular relevance for PACE programs, as ranking approaches are common in the decision making of farmland preservation planners (Duke and Johnston 2009).

To assess the robustness of welfare-based rankings farmland easement policies, the $N=96$ policy options addressed above are ranked (1 to 96) based on average per household CS for each option, with unique rankings calculated for each of the seven models. These model-specific rankings are compared to those from other models, providing 21 cross-model comparisons of policy rankings. Correlations among rankings for each possible model pair are calculated using Spearman's rank correlation coefficient (ρ). Results are shown in table 7.

As shown by Table 7, welfare based policy rankings across sites are highly correlated. The average correlation (ρ) across all model pairs is 0.94, with $\rho > 0.9$ for 19 out of 21 pairs. In contrast to results of tables 5 and 6, which provide mixed results regarding the robustness of welfare estimates across model specifications, results in Table 7 suggest clear robustness when welfare estimates are used solely to prioritize policy options. Across many model pairs there is a near perfect correlation among welfare based policy rankings for PACE programs. The strength of these correlations is particularly notable given the magnitude of average CS differences across the same models, which ranges from 9.77% to 884.26%. The contrast of these empirical outcomes demonstrates that, at least for the PACE policy context, the intended use of benefit estimates is a critical factor when assessing the confidence that one should place in model results. That is, even though CS estimates can vary to a large degree depending on CL and ML model specification choices, these variations may have little or no impact on associated policy

prioritization. Hence, as concluded above in the assessment of implicit prices, the practical

relevance of welfare sensitivity depends critically on the intended use of welfare estimates.

VI. CONCLUSIONS

As noted by Layton and Lee (2006, p. 52), "the growing sophistication of our econometric

literature serves to highlight an uncomfortable fact – we still do not know the true model." This

has not prevented common policy analyses from drawing upon results of individual ML models,

with little attention to the possible sensitivity of welfare estimates to a range of reasonable

functional forms. This paper provides insight regarding the sensitivity of welfare estimates and

practical policy guidance to common variations in ML specification—the econometric fine print.

As in most empirical assessments, there are a variety of analyses that are omitted and many

topics that remain unexplored. For example, the specific empirical results reported here apply

only to our particular case study addressing agricultural PACE programs in Georgia. In addition,

as in all works in the applied valuation literature, the model specifications illustrated here are at

least to some degree *ad hoc*, chosen primarily to represent specifications common in the

literature. These and other limitations aside, the analysis provides a variety of insights relevant

to the estimation and practical use of welfare estimates from discrete choice experiments.

Model results suggest that greater attention should be paid to practical policy implications

of even seemingly-minor changes in model specification, particularly given the flexibility of the

ML model and the large number of different specifications that appear in the literature. This

econometric fine print can have substantial implications for welfare estimates. Here we find the

potential for particularly large impacts on implicit prices. While CS estimates are somewhat

more robust on average, even these can vary to a degree that may be unacceptable for many

19

types of policy analysis. Indeed, variations from often unnoticed ML specification changes can lead to welfare consequences of a similar magnitude to those found in benefit transfers between different sites. Yet, while policy implications of errors in benefit transfer are a primary focus in that literature (Rosenberger and Stanley 2006), similar concerns for applied policy analysis related to ML model specification are often overlooked. Policy rankings, in contrast, appear robust across different CL and ML model specifications—providing encouraging news for those who seek to use results to prioritize policies rather than to quantify welfare change.

At a minimum, results suggest that additional emphasis should be given to sensitivity analysis in ML welfare analysis. The common practice of providing results for one or two ML choice experiment models, can mask large variations in welfare estimates across a range of seemingly reasonable model specifications. Attention to robustness is particularly relevant in discrete choice experiments given almost unavoidable uncertainty concerning the "true" or "best" model (Layton and Lee 2006; Louviere 2006).

From a broader perspective, results highlight the importance of research into the practical implications of welfare estimation methods. These implications tend to be overlooked by a valuation literature that tends to emphasize methodological advances, often at the expense of high quality, well-documented, and well-tested empirical estimates (Johnston and Rosenberger 2010). As stated by McComb et al. (2006, p.471), "[p]ressure from publications to create novel methods or formulations has resulted in an abundance of studies that are distant from the day-to-day needs of policy makers…" Among these needs is more reliable information on the confidence that may be placed on welfare point estimates and associated policy guidance, given the potential sensitivity of these outputs to often-overlooked statistical variations.

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.

Balcombe, K. A. Chalak and I. Fraser. Model Selection for the Mixed Logit with Bayesian
Estimation. Journal of Environmental Economics and Management 57(2): 226-237.

Ben-Akiva, M. 1981. Issues in transferring and updating travel-behavior models. In P.R.
Stopher, A.H. Meyburg and W. Borg (eds.) *New Horizons in Travel-Behavior Research*
(pp. 665-686). Lexington, MA: Lexington Books.

Bergstrom, J.C. and DeCivita, P. 1999 Status of benefit transfer in the United States and Canada:
a review. *Canadian Journal of Agricultural Economics* 47: 79-87.

Bergstrom, J.C., Ready, R.C., 2009. What have we learned from over 20 years of farmland
amenity valuation research in North America? *Review of Agricultural Economics* 31, 21-
49.

Boyle, K.J. and S. Özdemir. 2009. Convergent Validity of Attribute-Based, Choice Questions in
Stated-Preference Studies. *Environmental and Resource Economics* 42:247–264.

Campbell, D. 2007. Willingness to Pay for Rural Landscape Improvements: Combining Mixed
Logit and Random Effects Models. *Journal of Agricultural Economics* 58(3): 467-483.

Carson, R.T., J. J. Louviere and N. Wasi. 2009. A Cautionary Note on Design of Discrete Choice
Experiments: A Comment on Lust and Norwood's "Effect of Experiment Design on
Choice-Based Conjoint Valuation Estimates." *American Journal of Agricultural
Economics* 91(4): 1056-1063.

Columbo, S. and Hanley, N., 2008. How can we reduce the errors from benefits transfer? An

investigation using the choice experiment method. *Land Economics* 84, 128-147.

Columbo, S., Calatrava-Requena, J. and Hanley, N. 2007. Testing choice experiment for benefit transfer with preference heterogeneity. American Journal of Agricultural Economics 89: 135-151.

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

Duke, J.M. and R. Aull-Hyde. 2002. Identifying public preferences for land preservation using the analytic hierarchy process. *Ecological Economics* 42(1-2):131-45.

Duke, J.M. and R.J. Johnston. 2009. Nonmarket valuation of multifunctional farm and forest preservation, in: Goetz, S. and Brouwer, F., (Eds.) *New Perspectives on Agri-Environmental Policies: A Multidisciplinary and Transatlantic Approach*. Oxford, UK, Routledge Publishers.

Greene, W.H., 2003. *Econometric Analysis*, 5th ed. Upper Saddle River, NJ: Prentice Hall.

Greene, W.H., 2007. NLOGIT Version 4.0 Reference Guide. Plainview, NY: Econometric Software, Inc.

Hensher, D.A. and Greene, W.H., 2003. The mixed logit model: The state of practice. *Transportation* 30, 133–176.

Hu, W., W.L. Adamowicz and M.M. Veeman. 2006. Labeling Context and Reference Point Effects in Models of Food Attribute Demand. *American Journal of Agricultural Economics* 88:1034-1049.

Hu, W., M.M. Veeman and W.L. Adamowicz. 2005. Labeling Genetically Modified Food: Heterogeneous Consumer Preferences and the Value of Information. *Canadian Journal of Agricultural Economics* 53:83-102.

22

Jiang, Y., Swallow, S.K. and McGonagle, M., 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. and Duke, J.M., 2007. Willingness to pay for agricultural land preservation and policy process attributes: Does the method matter? *American Journal of Agricultural Economics* 89, 1098-1115.

Johnston, R.J. and Duke, J.M., 2008. Benefit transfer equivalence tests with non-normal distributions. *Environmental and Resource Economics* 41, 1-23.

Johnston, R.J. and Duke, J.M., 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.J. Opaluch, T.A. Grigalunas, and M.J. Mazzotta. 2001. Estimating Amenity Benefits of Coastal Farmland. *Growth and Change* 32(summer): 305-325.

Johnston, R.J. and R.S. Rosenberger. 2010. Methods, Trends and Controversies in Contemporary Benefit Transfer. Journal of Economic Surveys, available online, doi:10.1111/j.1467-6419.2009.00592.x. Print version in press.

Krinsky, I. and A.L. Robb. 1986. On Approximating the Statistical Properties of Elasticities. Review of Economics and Statistics 68:715-719.

Layton, D.F. 2000. Random Coefficient Models for Stated Preference Surveys. *Journal of Environmental Economics and Management* 40:21-36.

Layton, D.F. and S.T. Lee. 2006. Embracing Model Uncertainty: Strategies for Response Pooling and Model Averaging. *Environmental and Resource Economics* 34(1): 51-85.

Layton, D. F. and G. Brown. 2000. Heterogeneous Preferences Regarding Global Climate Change. *The Review of Economics and Statistics* 82(4): 616–624.

23

Louviere, J.J. 2006. What You Don't Know Might Hurt You: Some Unresolved Issues in The

    Design and Analysis of Discrete Choice Experiments. *Environmental and Resource*

    Economics 34(1): 173-188.

Lusk, J.L. and F.B. Norwood. 2005. Effect of Experimental Design on Choice-Based Conjoint

    Valuation Estimates. American Journal of Agricultural Economics 87:771-785.

Lusk, J.L. and F.B. Norwood. 2009. A Cautionary Note on the Design of Discrete Choice

    Experiments: Reply. *American Journal of Agricultural Economics* 91(4): 1064-1066.

McComb, G., Lantz, V., Nash, K. and Rittmaster, R. (2006) International valuation databases:

    overview, methods and operational issues. *Ecological Economics* 60: 461-472.

McFadden, D. and K. Train. 2000. Mixed MNL Models for Discrete Response. Journal of

    Applied Econometrics 15:447-70.

Messer, K.D. 2006. The conservation benefits of cost-effective land acquisition: A case study of

    Maryland. *Journal of Environmental Management* 79:305-315.

Morrison, M., Bennett, J., Blamey, R., and Louviere, J., 2002. Choice modeling and tests of

    benefit transfer. *American Journal of Agricultural Economics* 84, 161-170.

Navrud, S. and Pruckner, G.J. (1997) Environmental valuation – to use or not to use? A

    comparative study of the United States and Europe. *Environmental and Resource*

    Economics 10: 1-26.

Özdemir, S. 2003. Convergent validity of conjoint values for farmland conservation easement

    programs. Master's thesis, the University of Maine.

Özdemir. S. K.J. Boyle, M. Ahearn, A. Alberini, J. Bergstrom, L. Libby and M.P. Welsh. 2004.

    "Preliminary Report: Farmland Conservation Easement Study for the United States,

    Georgia, Ohio and Maine Samples." Working Paper, Department of Resource

Economics and Policy, University of Maine, Orono.

Paterson, R., K. Boyle, M. Ahearn, A. Alberini, J. Bergstrom, L. Libby, and M. Welsh. 2004. Public preferences for farmland attributes in conservation easement programs. In S. Goetz, J. Shortle, and J. Bergstrom (Eds.), Land Use Problems and Conflicts: Causes, Consequences and Solutions. Routledge.

Poe, G.L., K.L. Giraud and J.B. Loomis. 2005. "Computational Methods for Measuring the Difference in Empirical Distributions." American Journal of Agricultural Economics 87:353-365.

Provencher, B. and R.C. Bishop. 2004. Does Accounting for Preference Heterogeneity Improve the Forecasting of the Random Utility Model? A Case Study. *Journal of Environmental Economics and Management* 48: 793-810.

Ready, R.C., M.C. Berger, and G.C. Blomquist. 1997. Measuring Amenity Benefits from Farmland: Hedonic Pricing vs. Contingent Valuation. Growth and Change 28 (Fall): 438-458.

Rosenberger, R.S. and Stanley, T.D., 2006. Measurement, generalization and publication: Sources of error in benefit transfers and their management. *Ecological Economics* 60, 372-378.

Scarpa, R., M. Thiene and K. Train. 2007. Utility in WTP Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps. Working Paper in Economics 15/06, Department of Economics, University of Waikato.

Scarpa, R., T.J. Gilbride, D. Campbell and D.A. Hensher. 2009. Modelling Attribute Non-attendance in Choice Experiments for Rural Landscape Valuation. European Review of Agricultural Economics 36: 151-174.

Swait, J. and J. Louviere. 1993. The Role of the Scale Parameter in the Estimation and

      Comparison of Multinomial Logit Models. Journal of Marketing Research 30:305-314.

Train, K.E. 1998. Recreation Demand Models with Taste Differences over People. *Land*

      Economics 74: 230-239.

Train, K.E. 2009. *Discrete Choice Methods with Simulation.* Cambridge, UK: Cambridge

      University Press.

Train, K. E. and Weeks, M. (2005). Discrete Choice Models in Preference Space and Willing-to-

      Pay Space. In: R., Scarpa and A., Alberini, (eds), Applications of Simulation Methods in

      Environmental and Resource Economics. Dordrecht: Springer, 1–16.

Table 1. Attributes and Attribute Levels Included in the Georgia PACE Survey

Attributes	Attribute Levels
Commodity Produced	1. Grain Crops
	2. Vegetables, Berries, Fruit and Nut Crops
	3. Hay
	4. Pasture Grass for Livestock
	5. Timber
	6. no priority
Location	1. Near Urban Areas
	2. no priority
Quality	1. Prime agricultural soil
	2. no priority
Total Acres Preserved	1. 100,000 acres
	2. 500,000 acres
	3. 1,000,000 acres
	4. 2,000,000 acres
One-time Cost to Household	1. $3
	2. $5
	3. $7
	4. $10
	5. $25
	6. $50

Table 2. Summary of Model Specifications

	Distribution of Parameters (F=fixed, N=normal, L=lognormal, T=bounded triangular)						
	Model Specification						
Parameters	I	II	III	IV	V	VI	VII
Neither (ASC)	F	N	N	N	N	N	F
Acres	F	N	N	N	N	N	F
Location	F	N	N	N	N	N	F
Quality	F	N	N	N	N	N	F
Grain	F	N	N	N	N	N	F
Hay	F	N	N	N	N	N	F
Pasture	F	N	N	N	N	N	F
Vegetable	F	N	N	N	N	N	F
Forest	F	N	N	N	N	N	F
Cost (negative)	F	F	F	T	T	L	L
Demographic Fixed Effects Included[a]	No	No	No	No	Yes	No	No
Correlation Among Random Parameters	N/A	No	Yes	No	No	No	No

a   Included as interactions with the ASC.

Table 3. Model Variable Definitions and Summary Statistics

Variable	Description	Mean (Std. Dev.)[a]
Neither (ASC)	Alternative specific constant identifying the status quo (no policy).	0.333 (0.471)
Acres	Total acres of easements purchased in Georgia under each conservation easement program, in 10,000 acre increments.	59.689 (72.024)
Location	Binary (dummy) variable identifying policies that prioritize easement purchases near urban areas (versus the default of no priority).	0.333 (0.471)
Quality	Binary (dummy) variable identifying policies that prioritize easement purchases of prime farmland (versus the default of no priority).	0.325 (0.469)
Grain	Binary (dummy) variable identifying policies that prioritize easement purchases of land used for grain crops (versus the default of no priority).	0.115 (0.319)
Hay	Binary (dummy) variable identifying policies that prioritize easement purchases of land used for hay (versus the default of no priority).	(0.117) (0.322)
Pasture	Binary (dummy) variable identifying policies that prioritize easement purchases of land used for livestock pasture (versus the default of no priority).	0.095 (0.293)
Vegetable	Binary (dummy) variable identifying policies that prioritize easement purchases of land used for growing vegetables, berries, fruits and nuts (versus the default of no priority).	(0.110) (0.312)
Forest	Binary (dummy) variable identifying policies that prioritize easement purchases of land used for timber (versus the default of no priority).	(0.112) (0.316)
Birth_Date	The year during which the survey respondent was born.	1952.35 (14.633)
College	Binary (dummy) variable identifying college graduates.	(0.522) (0.500)
Hi_Income	Binary (dummy) variable identifying respondents with a household income of $60,000 or greater.	(0.447) (0.497)
Cost (negative)	Unavoidable one-time household cost (2002 dollars), with sign reversal.	(-10.991) (15.344)

[a] Includes zeros for the 'status quo' option.

Table 4a. Estimation Results—Conditional and Mixed Logit Models I-IV[a]

Attribute	Conditional Logit I	Mixed Logit II	Mixed Logit III[b]	Mixed Logit IV
Neither (ASC)	0.2684	-3.2830***	-8.0049***	-2.0390**
	(0.1877)	(0.8617)	(2.4783)	(0.8481)
Acres	0.0023***	0.0039**	0.0175*	0.0034**
	(0.0008)	(0.0017)	(0.0093)	(0.0016)
Location	0.3517***	0.5239**	2.8485**	0.4962**
	(0.1100)	(0.2348)	(1.1802)	(0.2172)
Quality	0.4137***	0.8405***	3.5731**	0.7582***
	(0.1074)	(0.2203)	(1.4186)	(0.2072)
Grain	0.2291	0.2997	-0.4258	0.2654
	(0.1806)	(0.2984)	(1.1561)	(0.2802)
Hay	-0.4402**	-0.8235***	-2.4146*	-0.7676**
	(0.1925)	(0.3147)	(1.4545)	(0.3096)
Pasture	0.1453	0.0259	-0.3649	0.0557
	(0.1909)	(0.3750)	(1.8495)	(0.3320)
Vegetable	0.6677***	1.0857***	3.2785	1.0792***
	(0.1832)	(0.3291)	(2.0800)	(0.3217)
Forest	0.2031	0.3682	2.1136	0.3351
	(0.1824)	(0.3767)	(1.3809)	(0.3252)
Cost (negative)	0.0163***	0.0340***	0.1263***	0.0339***
	(0.0036)	(0.0075)	(0.0393)	(0.0080)
Standard Deviations of Parameter Distributions				
Std Dev Neither	--	7.8395***	20.2633***	7.3166***
		(1.2553)	(6.2478)	(1.1785)
Std Dev Acres	--	0.0102***	0.0472***	0.0094***
		(0.0027)	(0.0134)	(0.0032)
Std Dev Location	--	1.5167***	6.0764***	1.2002***
		(0.4892)	(1.8158)	(0.3996)
Std Dev Quality	--	0.9801**	6.6565***	0.7341
		(0.4288)	(1.6927)	(0.5162)
Std Dev Grain	--	0.5486	5.9926***	0.2949
		(0.8402)	(1.4168)	(1.0690)
Std Dev Hay	--	0.3953	5.3652***	0.1216
		(0.6273)	(1.9364)	(0.8464)
Std Dev Pasture	--	1.6900***	7.9595***	1.1034*
		(0.5497)	(1.9912)	(0.6578)
Std Dev Vegetable	--	0.0339	8.9500***	0.8147
		(0.9204)	(2.9610)	(0.9281)
Std Dev Forest	--	2.2359**	6.8547***	1.5498**
		(0.8681)	(2.3158)	(0.6423)
Std Dev Cost	--	--	--	0.0339*** [c]
				(0.0080)
Log-Likelihood χ^2	102.5526***	533.1811***	577.0305***	526.2655***
(df)	(10)	(19)	(55)	(19)
Pseudo-R^2	0.05	0.30	0.33	0.30
N	803	803	803	803

Table 4b. Estimation Results—Mixed Logit Models V-VII[a]

Attribute	Mixed Logit V	Mixed Logit VI	Mixed Logit VII
Neither (ASC)	-33.9418	-3.6716***	-1.6351***
	(70.4278)	(0.8773)	(0.2413)
Acres	0.0035**	0.0052**	0.0037***
	(0.0016)	(0.0023)	(0.0009)
Location	0.4656**	0.5881**	0.3683***
	(0.2208)	(0.2762)	(0.1221)
Quality	0.6934***	0.9123***	0.6151***
	(0.2024)	(0.2766)	(0.1278)
Grain	0.1997	0.3406	0.3573
	(0.3004)	(0.3722)	(0.2182)
Hay	-0.7856***	-0.9421**	-0.5103**
	(0.3012)	(0.4444)	(0.2197)
Pasture	-0.0203	0.1699	0.1714
	(0.3502)	(0.4070)	(0.2310)
Vegetable	1.0686***	1.6214***	0.9482***
	(0.3386)	(0.5102)	(0.2296)
Forest	0.2715	0.5820	0.3630*
	(0.3411)	(0.4981)	(0.2166)
Cost (negative)	0.0357***[b]	-4.0311***[c]	-3.9817***[c]
	(0.0087)	(0.6605)	(0.6420)
Demographic Interactions			
*Neither*Birth_date*	0.0162	--	--
	(0.0360)		
*Neither*College*	-1.8378	--	--
	(1.2919)		
*Neiher*Hi_Income*	2.8554**	--	--
	(1.1464)		
Standard Deviations of Parameter Distributions			
Std Dev Neither	7.8228***	3.3389***	--
	(1.4659)	(0.9941)	
Std Dev Acres	0.0083***	0.0125***	--
	(0.0027)	(0.0033)	
Std Dev Location	1.3345***	1.7022***	--
	(0.4031)	(0.5962)	
Std Dev Quality	0.6452	0.5749	--
	(0.4960)	(0.5158)	
Std Dev Grain	0.6891	0.9681	--
	(0.7245)	(0.7362)	
Std Dev Hay	0.1344	1.0025	--
	(0.6641)	(0.9345)	
Std Dev Pasture	1.2530*	1.4811*	--
	(0.6824)	(0.8148)	
Std Dev Vegetable	0.8533	1.3913*	--
	(0.8377)	(0.7350)	
Std Dev Forest	1.7024***	2.6040***	--

	(0.6639)	(0.8020)	
Std Dev Cost	0.0357***[b]	5.1101***[c]	4.9595***[c]
	(0.0087)	(0.8087)	(0.8355)
Log-Likelihood χ^2	486.41***	536.9539***	501.0977***
(df)	(22)	(20)	(11)
Pseudo-R^2	0.29	0.30	0.28
N	752	803	803

a   Standard errors in parentheses. Single (*), double (**) and triple (***) asterisks denote significance levels of 0.10, 0.05 and 0.01, respectively.

b   Spread of the triangular distribution. Model is constrained such that the parameter estimate is equal to the spread, to ensure a the parameter distribution lies in the positive domain.

c   Lognormal distribution on sign-reversed cost parameter.

Table 5a. Implicit Prices: Model Specifications I-IV[a]

Attribute	Conditional Logit I[b]	Mixed Logit II[c]	Mixed Logit III[c]	Mixed Logit IV[c]
Acres	0.147***	0.122***	0.140**	0.145**
	(0.060, 0.261)	(0.024, 0.225)	(0.019, 0.264)	(<0.001, 0.322)
Location	22.653***	15.872**	22.690**	12.759**
	(10.254, 39.675)	(2.404, 30.698)	(7.656, 37.419)	(2.929, 46.299)
Quality	26.522***	25.436***	28.456**	32.880***
	(13.577, 44.821)	(13.447, 40.303)	(10.441, 47.247)	(15.195, 57.507)
Grain	14.672	8.329	-6.159	11.933
	(-4.006, 36.420)	(-3.434, 23.823)	(-26.594, 12.649)	(-8.741, 35.538)
Hay	-28.457***	-25.388***	-21.641	-32.602***
	(-54.584, -7.515)	(-45.260, -7.934)	(-48.166, 1.215)	(-60.414, -10.208)
Pasture	9.449	1.108	-2.817	3.285
	(-10.555, 31.791)	(-20.790, 21.311)	(-35.068, 26.680)	(-22.548, 32.864)
Vegetable	42.765***	33.694***	26.658	47.784***
	(21.093, 74.037)	(15.620, 58.236)	(-1.522, 54.037)	(21.841, 89.402)
Forest	13.046	11.592	16.784	15.822
	(-6.214, 35.486)	(-9.765, 35.946)	(-3.642, 37.021)	(-9.258, 45.222)

Table 5b. Implicit Prices: Model Specifications V-VII [a]

Attribute	Mixed Logit V[c]	Mixed Logit VI[d]	Mixed Logit VII[d]	Percent Change, Lowest to Highest[e]
Acres	0.143**	0.060**	0.255***	325.00%
	(0.015, 0.299)	(<0.001, 0.253)	(0.057, 0.615)	
Location	19.013**	5.534**	26.358***	376.29%
	(1.011, 42.102)	(0.001, 23.909)	(3.833, 70.153)	
Quality	28.099***	44.516***	42.312***	75.01%
	(12.098, 48.613)	(0.556, 144.514)	(9.281, 108.587)	
Grain	8.170	5.696	26.399	528.62%
	(-11.125, 31.946)	(-3.917, 36.816)	(-0.769, 78.062)	
Hay	-33.046***	-30.693**	-35.108**	62.23%
	(-60.763, -10.794)	(-123.822, -0.132)	(-101.157, -3.336)	
Pasture	-1.249	2.998	14.288	782.12%
	(-28.801, 25.620)	(3.005, 26.524)	(-12.357, 64.410)	
Vegetable	43.957***	62.257***	66.906***	150.98%
	(19.742, 76.934)	(3.039, 206.079)	(14.418, 170.902)	
Forest	11.535	5.786	26.540**	358.69%
	(-14.922, 36.589)	(-0.092, 28.681)	(1.413, 77.718)	

a   Bounds on empirical 90% confidence interval in parentheses. * p< 0.10, ** p<0.05, *** p<0.01.
b   Mean WTP.
c   Mean of Mean WTP (see text).
d   Mean of Median WTP (see text).
e   Absolute value percentage difference between lowest and highest implicit prices over Models I – VII.

Table 6. Compensating Surplus Estimates: Average Absolute Value Percentage Differences Across Models I - VII[a]

Baseline Model	Comparison Model						
	Model I	Model II	Model III	Model IV	Model V	Model VI	Model VII
Model I	--	673.93%	489.15%	596.20%	529.67%	826.13%	884.26%
Model II	82.04%	--	23.40%	9.77%	16.41%	21.55%	32.42%
Model III	77.41%	31.58%	--	21.02%	10.79%	60.51%	73.37%
Model IV	81.02%	11.52%	16.47%	--	8.79%	34.24%	46.19%
Model V	80.00%	20.40%	9.39%	10.24%	--	46.31%	58.92%
Model VI	85.18%	17.22%	36.63%	25.04%	30.92%	--	12.14%
Model VII	87.14%	24.18%	41.91%	31.14%	36.95%	10.70%	--

	All Models	All Mixed Logit Models
Mean Percentage Difference	126.72%	27.67%

[a] 5% trimmed mean in absolute percentage difference calculated over 512 possible conservation easement attribute combinations for each model.

34

Table 7. Welfare-Based Policy Rankings: Spearman Rank Correlations across Models I-VII

	Model I	Model II	Model III	Model IV	Model V	Model VI	Model VII
Model I	1.000	--	--	--	--	--	--
Model II	0.988	1.000	--	--	--	--	--
Model III	0.918	0.947	1.000	--	--	--	--
Model IV	0.905	0.937	0.840	1.000	--	--	--
Model V	0.986	0.995	0.934	0.935	1.000	--	--
Model VI	0.904	0.928	0.856	0.945	0.933	1.000	--
Model VII	0.988	0.992	0.924	0.940	0.983	0.922	1.000

[a] Correlation of rankings among 96 possible conservation easement attribute combinations for each model.