COMBINING REVEALED AND STATED PREFERENCE DATA TO ESTIMATE THE NONMARKET VALUE OF ECOLOGICAL SERVICES: AN ASSESSMENT OF THE STATE OF THE SCIENCE

John C. Whitehead[1]

Department of Economics, Appalachian State University

Subhrendu K. Pattanayak

Environmental and Natural Resource Economics Program, Research Triangle Institute

George L. Van Houtven

Environmental and Natural Resource Economics Program, Research Triangle Institute

Brett R. Gelso

Department of Public Administration and Policy, American University

August 16, 2005 (revised December 11, 2007)

[1] Corresponding author e-mail: whiteheadjc@appstate.edu.

Abstract. This paper reviews the marketing, transportation, and environmental economics literature on the joint estimation of revealed and stated preference data. The revealed preference and stated preference approaches are first described with a focus on the strengths and weaknesses of each. Recognizing these strengths and weaknesses, the potential gains from combining data are described. A classification system for combined data that emphasizes the type of data combination and the econometric models used is proposed. A methodological review of the literature is pursued based on this classification system. Examples from the environmental economics literature are highlighted. A discussion of the advantages and disadvantages of each type of jointly estimated model is then presented. Suggestions for future research, in particular opportunities for application of these models to environmental quality valuation, are presented.

Keywords: Environmental valuation revealed preference, stated preference, data combination, joint estimation, calibration

JEL Code: Q51

1. Introduction

There are numerous valuation methodologies available to estimate the benefits of public goods (Freeman, 1993). These methodologies can be classified as revealed preference or stated preference approaches. Revealed preference approaches use behavioral data to estimate the *ex-post* willingness to pay for various commodities. The travel cost method, averting behavior method, the hedonic price method, and the production function approach are examples of revealed preference approaches. Stated preference approaches use hypothetical data to estimate the *ex-ante* willingness to pay for various commodities. The contingent valuation, contingent behavior, and conjoint analysis (i.e., discrete choice experiments) methods are examples of stated preference approaches.

In the midst of the heated debate sparked by the Exxon Valdez oil spill over the contingent valuation of nonuse values (Carson et al., 2003; Portney, 1994), Cameron (1992) and Adamowicz, Louviere and Williams (1994) introduced an exciting new approach to the environmental valuation literature – the combination and joint estimation of revealed preference and stated preference data (i.e., the data enrichment paradigm; see Louviere, Hensher and Swait, 2000). In part, these seminal papers led to the organization of the 1996 Association of Environmental and Resource Economists (AERE) Workshop held in Lake Tahoe, California: "Combining Stated Preference Data and Revealed Preference Data to Estimate the Demand for and/or Benefits from Environmental Amenities."

The excitement over the promise of joint estimation methods to "solve" the problems of contingent valuation method has dimmed over the years. Revealed behavior is required for joint estimation and the current methods for "calibration" of nonuse values were developed in the experimental economics literature. But, a new literature was launched, fueled by the AERE workshop, that addresses hypothetical bias of contingent behavior data, valuation of quality change beyond the range of historical experience, and development of new econometric and survey methods, among other advances.

Much of this new literature has roots in the marketing and transportation fields where stated preference data has been found especially useful. In marketing, analysis of demand change with the introduction of new brands of products is essential (Ben-Akiva et al. 1994). In transportation, a particular problem in the public sector is the analysis of new public transportation modes and routes. Revealed preference data does not exist for new products and transportation alternatives and stated preference data can be used to fill in the gaps in observed behavior (Hensher, 1994). More recently, significant advances have been made in the field of health economics where new treatments or drugs do not have any related demand and stated preference data is useful (Mark and Swait, 2004). In environmental economics, stated preference data is useful for analysis of policies leading to changes in behavior (Phaneuf and Smith, 2005) as well as policies that may not lead to changes in behavior (Carson and Hanemann, 2005). Stated preference data could also be useful in many other fields where analysis of new policies or products is limited by the quality of revealed preference data. Given the disadvantages of stated preference data relative to revealed preference data, data combination and joint estimation is an important methodological option in all of these fields.

The purpose of this paper is to review the marketing, transportation, and environmental economics literature on the joint estimation of revealed and stated preference data with a focus on application in environmental economics. The revealed preference and stated preference approaches are first described with a focus on the strengths and weaknesses of each. Recognizing these strengths and weaknesses, the potential gains from combining data are described. A classification system for combined data that emphasizes the type of data combination and the econometric models used is proposed. A methodological review of the literature is pursued based on this classification system. Examples from the environmental economics literature are highlighted. A discussion of the advantages and disadvantages of each type of jointly estimated model is then presented. Suggestions for future research, in particular opportunities for application of these models to environmental quality valuation, are presented. Conclusions follow.

1.1 *Revealed Preference Methods*

The travel cost method is a revealed preference method that is most often used to estimate the benefits of outdoor recreation (Herriges and Kling, 1999; Parsons, 2003); for example, the improved fishing opportunities following water quality improvement. The travel cost method begins with the realization that the major cost of outdoor recreation is the travel and time costs incurred to get to the recreation site. Since individuals reside at varying distances from the recreation site, the variation in distance and the number of trips taken are used to trace out a demand curve. The recreation demand curve is used to derive the benefits of the recreation site. With the appropriate demand shifters (i.e., independent variables such as measures of water quality) the benefits of changes in policy variables can be derived. A variation of the travel cost method is the random utility model. The choice underlying the random utility model is the recreation site choice. Individuals make tradeoffs among trip costs and site characteristics (e.g., expected catch). Analyses of these choices allow the estimation of the benefits of the site characteristics.

The averting behavior method begins with the recognition that individuals seek to protect themselves when faced with environmental risk such as contaminated drinking water (Smith, 1991; Dickie, 2003). Defensive behavior requires expenditures that would not normally be made. For example, the purchase of bottled water or water filters may only be made when faced with the risk of contaminated drinking water. Generally speaking, these increased expenditures are positively correlated with the economic benefits of policy that reduces the drinking water risk.

The hedonic price method exploits the relationship between characteristics of land and labor markets, including water quality, and housing prices and wages (Palmquist, 1991; Taylor, 2003). For example, improved water quality is an amenity for some landowners. Land parcels in close proximity to water bodies with high quality water command higher prices in land markets relative to others (Leggett and Bockstael, 2000). Job markets with greater locational amenities are associated with lower wages as the supply of labor is higher relative to other locations (Clark and Kahn, 1989). The housing price and wage differentials are measures of the implicit price of locational amenities such as water quality. Housing and labor market differences can be used to trace out the demand for water quality and used to measure economic benefits.

The major strength of the revealed preference approaches is that they are based on actual choices. With revealed preference data, individuals consider the internal costs and benefits of their actions and experience the consequences of their actions. Choices based on the perceived costs and benefits better reflect the values of the population and allow more valid estimates of willingness to pay. The major weakness of revealed preference approaches is their reliance on historical data. New government policies and new products are often beyond the range of historical experience. For example, few residents of an urban area nearby a long-degraded stream have experience with a fishable stream. Behavior in response to policies designed to restore the stream is nonexistent. Analysis of the economic benefits of policy is often difficult with revealed preference data.

1.2 *Stated Preference Methods*

The contingent valuation method is a stated preference approach that directly elicits willingness (and ability) to pay statements from survey respondents (Mitchell and Carson, 1989; Bateman et al., 2003). In other words, respondents are directly asked about their willingness to pay. The method involves the development of a hypothetical market in the context of in-person, telephone, mail, or other types of surveys. In the hypothetical market respondents are informed about the current problem and the policy designed to mitigate the problem. The state of the environment before and after the policy is described. Other contextual details about the policy are provided such as the policy implementation rule (e.g., majority rule) and the payment vehicle (e.g., increased taxes or utility bills). Finally, a hypothetical question is presented that confronts respondents with a choice about improved environmental quality and increased costs versus the status quo. Respondents can be presented with multiple scenarios and make multiple choices.

The contingent behavior approach is similar to the contingent valuation method in that it involves hypothetical questions. In contrast, the questions involve hypothetical behavior instead of hypothetical willingness to pay. For example, respondents can be asked about hypothetical recreation trips with and without water quality improvements. The dependent variable is typically continuous but can also be a discrete choice. Conjoint analysis, or choice experiments, is another type of contingent behavior or choice approach that asks about hypothetical recreation site and other preferences over a range of fixed alternatives (Louviere, Hensher, and Swait, 2000). Respondents can be presented with multiple scenarios and make multiple choices.

A strength of the stated preference approaches is flexibility. Stated preference approaches can be used to construct realistic policy scenarios for most new policies. Oftentimes, hypothetical choices are the only way to gain policy relevant information. Another strength of the stated preference approaches, especially contingent valuation and conjoint analysis, is the ability to measure non-use values. The major weakness of the stated preference approaches is their hypothetical nature. Respondents are sometimes placed in unfamiliar situations in which complete information is not available. At best, respondents give truthful answers that are limited by their unfamiliarity. At worst, respondents give trivial answers due to the hypothetical nature of the scenario.

2. Gains from Combining Data

Historically, researchers have seen revealed and stated preference approaches as substitutes when considering the choice of valuation methods. Comparisons of revealed and stated preference approaches were conducted to determine the validity of stated preference methods (Cummings, Brookshire, and Schulze, 1986). Yet, the strengths of the revealed preference approaches are the weaknesses of the stated preference approaches. Recognizing this complementarity, a new paradigm emerged in the marketing and transportation literatures that sought to combine and jointly estimate revealed and stated preference data (Ben-Akiva and Morikawa, 1990; Morikawa, Ben-Akiva, and Yamada, 1991; Hensher and Bradley, 1993; Swait and Louviere, 1993). These methods were then used in the field of environmental economics (Cameron, 1992; Adamowicz, Louviere, and Williams, 1994). The combination and joint estimation of revealed and stated preference data seeks to exploit the contrasting strengths of the various approaches while minimizing their weaknesses. This combined data approach is therefore described in Louviere (2000) as following a “data enrichment” paradigm.

Revealed preference data can be enhanced by stated preference data. Some revealed preference data are limited to analyzing a range of behavior in response to a limited range of market or environmental change. For example, there is no revealed preference data for new market products, a new transportation mode, or behavior in response to future environmental change at individual sites. Stated preference surveys can be designed to collect data on this hypothetical behavior and the stated preference data allows estimation of behavior beyond the range of historical experience. Combining revealed preference data with stated preference data allows an extension of the behavioral model beyond the limited range of historical experience.

A related issue is the size of the market or ‘endogenous stratification’. Revealed preference data are often collected with on-site surveys of current consumers, truncating the dependent variable at one. The market size with a revealed preference survey is limited to current consumers. General population surveys can be used to survey non-users of the commodity. This allows analysis of the decision to participate in the market but these data are limited when trying to understand changes in participation in response to a new product, policy or quality change. Combining revealed preference data with stated preference data from surveys of the general population can be used to understand changes in participation and the market size with new products or environmental change.

Even when there is experience with the product or policy under consideration, there is often high correlation between variables of interest and other variables (Cameron, et al. 1996). For example, various pollution measures tend to be highly correlated. Fuel efficiency is correlated with the type of automobile. Multicollinearity among characteristics leads to statistically insignificant coefficient estimates which make it impossible to estimate the value of changes in the variables of interest. A related problem is endogeneity. Catch rates are correlated with fishing experience and both variables are related to fishing trip frequency. Analysis of a policy to value an increase in catch rates with revealed preference data can be confounded by fishing experience. Berry, Levinsohn, and Pakes (2004) and Murdock (2006) have proposed two-stage econometric strategies for addressing endogeneity when only revealed preference data are available. However, an alternative strategy is to combine revealed preference data with stated preference data. For example, Von Haefen and Phaneuf (2007) have shown that, when revealed and stated preference date are collected under an appropriate experimental design, they can be used to break the multicollinearity and avoid endogeneity problems.

Another limitation of revealed preference data is the inefficiency of data collection. Oftentimes, a revealed preference cross-section survey will collect only one data point. Cross-section, time series (i.e., panel) surveys collect more data from each respondent under possibly different conditions when allowing time for changes in behavior. However, this significantly increases the costs of data collection and panel data are prone to a significant loss of respondents potentially leading to sample selection problems. Supplementing the single data point from a cross-section survey with one, or several, stated preference questions can significantly increase the sample size. These additional choice occasions can be treated as panel data to provide more information about the preferences of each individual in the sample. More information from each respondent can lead to increased econometric efficiency. With stated preference data and increased econometric efficiency, smaller sample sizes can be used to achieve confidence intervals that required larger sample sizes without stated preference data. Reduced sample sizes reduce survey costs which reduces the costs of research.

Stated preference data can be enhanced by revealed preference data. Hypothetical bias can be a major problem with stated preference data. In many cases, hypothetical choices may not reflect budget, and other, constraints on behavior. For example, in a stated preference survey consumers may prefer sports cars over sedans and environmentalists may express a desire to pay for a large number of environmental projects. Yet, when the actual choice must be made, constraints are confronted and the sedan is purchased and some environmental projects go unfunded. Combining stated preference data with revealed preference data grounds hypothetical choices with real choice behavior.

Combining revealed preference and stated preference data can also be used to test the validity of both the revealed and stated preference methods. For example, the travel cost method has been criticized for the measurement error and endogeneity of the trip cost (Randall, 1994). As a result, Randall argues, welfare estimates resulting from the travel cost method are ordinal values based on subjective costs instead of cardinal values based on objective costs. If this problem is severe, invalid results from the travel cost method analysis could be used in policy analysis without detection. Combination with stated preference data can isolate the effect of changes in trip costs (Englin and Cameron, 1996; Earnhart, 2004). As mentioned previously, hypothetical bias is the key problem with stated preference methods. Combination with revealed preference data can be used to detect and mitigate hypothetical bias and validate stated preference methods.

Researchers who solely focus on revealed preference or stated preference methods have spent a considerable amount of time dealing with these issues in isolation from each other.

Another approach to dealing with these, and other problems that may arise in data collection and analysis, is combining data to test for convergent validity. Convergent validity exists if two methods for measuring an unknown construct (i.e., willingness to pay) yields measures that are not statistically different. Joint estimation can be used to restrict parameter estimates in theoretically appropriate ways to test for convergent validity. Combination and joint estimation of revealed and stated preference data can be used to validate both types of data.

3. Types of Combined Data Studies and Models

There are several versions of combined revealed preference (RP) and stated preference (SP) data in the literature (Figure 1). For the purposes of this delineation, combined revealed and stated preference (RP-SP) data includes any research effort that employs both types of data. Two dimensions for which the use of combined RP-SP data differs are whether the data are stacked and whether the error term is assumed to be identically and independently distributed (IID). Data can be stacked when revealed and stated preference observations have similar dependent and independent variables. Researchers treat the stated preference data as additional observations and increase the sample size by including the observations with the revealed preference data. The parameter estimates are generally constrained to be equal between the revealed and stated preference data. The typical assumption made in regression analysis is that the error terms are independent and identically distributed. When revealed and stated preference data are combined this assumption is usually violated because the error terms across respondents are correlated (i.e., not independent).

Three of the four types of RP-SP studies in Figure 1 are jointly estimated. Joint estimation occurs when the relationships between the independent variables and the dependent variables are estimated in a one or two or more equation model at the same time. The exception is the comparison study which does not stack data and assumes independent and identically distributed errors. The comparison study typically employs a revealed preference method (e.g., travel cost) and a stated preference method (e.g., contingent valuation) to separately estimate willingness to pay for the same policy change, often with the same sample of respondents (e.g., Laughland et al., 1996). The willingness to pay estimates are compared to determine convergent validity. Willingness to pay estimates that do not converge can be used to provide guidance about the bias generated by one or both of the methods. Carson, et al. (1996) conduct a meta-analysis of willingness to pay comparisons from the contingent valuation method and revealed preference methods. They find that there is a positive correlation in the willingness to pay estimates. This result suggests that both types of data reflect common preferences and that jointly estimating the data is potentially valid (see also Wardman, 1988 for an example with transportation data). On the other hand, studies that compare willingness to pay from hypothetical and actual referendum votes have mixed results regarding convergent validity (Vossler and Kerkvliet, 2003; Schläpfer, Roschewitz and Hanley, 2004).

Convergent validity refers to the case when both the RP and SP parameter estimates or willingness to pay estimates are statistically equivalent. If the estimates are statistically equivalent, that indicates that bias exists in one data that is different than the other data; however; there is no way to tell which data is correct. Consider the case of two non-market valuation surveys. The SP data are plagued with hypothetical market bias while corresponding RP data does not have a lot of variation and hence most of the coefficients in the model are statistically insignificant. If one were to attempt to combine the SP and RP data, the research would reject the hypothesis that the data were generated from the same preferences due to the problems inherent in each method. Other the other hand the data enrichment paradigm recognizes the limitations of both types of data and allows that data combination might improve both types, even if tests of convergent validity fail. Biases in both SP and RP data can be mitigated by combination and joint estimation.

Pooled data studies use revealed and stated preference data that are stacked with errors assumed to be independent and identically distributed. Pooled data studies typically constrain the coefficients to be equal across data types but ignore the correlation in behavior by the same individual within data sources and from one data source to another. Panel data studies use revealed and stated preference data that is stacked with errors assumed to be correlated (*i.e.*, not independently and identically distributed). A variety of approaches have been used to induce error correlation across observations including heteroskedastic, fixed effects, random effects, and autoregressive models. Mixed data studies use revealed and stated preference data that is not from a common framework and unable to be stacked for joint estimation.

Revealed and stated preference data are combined and analyzed with three major classes of econometric models: discrete choice models (e.g., logit, probit), continuous choice models (e.g., Tobit, Poisson), and mixed choice models (Figure 2). Discrete choice models typically involve combination of multiple choice revealed preference data with conjoint analysis stated preference data in which respondents choose from a menu of consumer goods, transportation choices, environmental programs or recreational sites. Since the data has the same structure in terms of the form of the dependent variable and independent variables it can be stacked and jointly estimated.

When data are simply pooled in multiple discrete choice models, the multinomial logit model, which assumes the errors are independently and identically distributed, is typically used (Adamowicz, Louviere, and Williams, 1994). Following Hensher, Rose and Greene (2005), we use multinomial logit model to refer to a class of discrete choice models including conditional, nested and mixed logit models. The coefficients across revealed and stated preference data are typically constrained except when various coefficients suffer from opposite signs. Then, dummy variables are used to allow the coefficients to be unconstrained. When panel data approaches are used, they can also be considered with and without constrained coefficients. The econometric models used in this type of research with constrained coefficients include heteroskedastic (Hensher, Louviere, and Swait, 1999) and other extensions to the multinomial logit, foremost being the mixed logit (Train, 1999). The bivariate dichotomous choice (e.g., logit, probit) models could be used to estimate unconstrained coefficient models with two choices. But, this type of joint estimation has not appeared in the literature.

Continuous choice models typically involve revealed preference data on recreation trips (Englin and Cameron, 1996) or other continuous variable such as miles driven (e.g., Golob, Bunch, and Brownstone, 1997). These data are combined with stated preference data in which respondents make similar choices under hypothetical conditions. As in discrete choice models, the data from the two sources have the same structure (in terms of the continuous form of the dependent variable and similar independent variables), and they can be stacked and jointly estimated.

When the continuous choice revealed and stated preference data are simply pooled, ordinary least squares (Layman, Boyce and Criddle, 1996), Poisson/Negative Binomial (Eisworth, et al., 2000), or Tobit models which assume the errors are independently and identically distributed, can be used. In general, most of the coefficients across revealed and stated preference data are constrained to be equal. When panel data approaches are used, the econometric models used in this type of research with constrained coefficients include random effects Tobit (Azevedo, Herriges, and Kling, 2003) and random effects Poisson/Negative Binomial models (Whitehead, Huang, and Haab, 2000). The bivariate Poisson has been used to estimate unconstrained coefficient models with continuous choices (Grijalva, Bohara, and Berrens, 2003). The bivariate Tobit model could also be used to jointly estimate RP-SP models but has not appeared in the literature for this purpose.

Mixed choice models typically involve a discrete choice contingent valuation hypothetical choice (e.g., a referendum) jointly estimated with a related demand model. These data can not be stacked, typically, but can be jointly estimated by allowing for error correlation (Cameron, 1992; Huang, Haab, and Whitehead, 1997). Cameron (1992) uses data from a discrete choice on the decision to continue taking fishing trips and continuous fishing demand. The model is estimated with correlated errors and trips predicted from the demand model as an exogenous variable in the willingness to pay decision. Some mixed choice models involve discrete/discrete and continuous/continuous dependent variables. For example, Loomis (1997) stacks discrete choice recreation demand and discrete choice contingent valuation data. Whitehead (2005) uses a simultaneous ordinary least squares and Tobit model to jointly estimate the continuous change in revealed and stated preference recreation trips and continuous willingness to pay for a water quality improvement by allowing correlated errors. Use of bivariate probit and Tobit models with mixed choices is also feasible.

3.1 *Discrete Choice Models*

Discrete choice models are most commonly based on random utility theory. With a RP choice, respondents face numerous alternatives with different characteristics. Many of these characteristics are often correlated. With a SP choice experiment (using conjoint analysis), respondents are presented with a similar choice among, typically, two or three options. Each option has a set of two or more characteristics in an experimental design that breaks the multicollinearity in the revealed preference data (Louviere, Hensher and Swait, 2000). In the case of an automobile purchase, the characteristics of the automobiles might include style, color, engine type, type of fuel, fuel efficiency, and cost, among other possible characteristics. A typical environmental economics application is the recreational site choice random utility model. The characteristics of each site might include bag or catch rates, scenery, water features, and cost. The discrete choice models have traditionally been estimated using the multinomial logit model. The random parameters (i.e., mixed) logit model is increasingly being used.

3.1.1 *Multinomial Logit*

Random utility theory is the basis for RP-SP research involving choice experiments. The basis for random utility theory is the following utility model

$$\begin{aligned} U_{ij} &= V_{ij} + \varepsilon_{ij} \\ &= \beta' X_{ij} + \varepsilon_{ij} \end{aligned} \tag{1}$$

where U_{ij} is the utility consumer i receives from alternative j, i = 1, …, I, j = 1, …, J, $V_{ij} = \beta' X_{ij}$ is the systematic portion of utility, β is a vector of parameters and X_{ij} is a vector of variables specific to the choice and individual characteristics, and ε is the random error. When the random errors are independent and identically distributed (IID) Gumbel errors, the multinomial logit

(MNL) model results (Hensher, Louviere, and Swait, 1999)

$$P_{ij} = \frac{\exp(\mu\beta' X_{ij})}{\sum_{j=1}^{J} \exp(\mu\beta' X_{ij})} \tag{2}$$

where P_{ij} is the probability that individual i chooses site j, and μ is the scale factor which is typically assumed to be equal to one.

One drawback to the MNL model is the restrictive independence of irrelevant alternatives (IIA) restriction. Some techniques used to relax this restriction are inclusion of alternative specific constants, an opt-out choice and the use of the nested logit model (Louviere, Hensher, and Swait, 2000). Alternative specific constants in a conditional logit capture unobserved heterogeneity about the choices themselves. In a two alternative choice experiment an opt-out choice is the decision not to choose either alternative. The nested logit models the choice as a decision tree. Respondents may first choose to participate in water-based recreation relative to other types of recreation and then choose the recreation site. The mixed logit completely relaxes the assumption of IID and IIA, allowing for preferences to vary over individuals and pairs of alternatives to be correlated. One limitation of these types of models is the very large computational algorithms. Since closed form solutions are not possible with models that allow free variance, simulated maximum likelihood is necessary.

Since both revealed and stated preference data follows the theoretical choice framework above, the MNL model can be used to combine and jointly estimate RP-SP data. Suppose the MNL is used to estimate the probability of the RP choice of alternative j, P_{Rij},

$$P_{Rij} = \frac{\exp(\mu_R \beta_R' X_{ij})}{\sum_{j=1}^{J} \exp(\mu_R \beta_R' X_{ij})} \tag{3}$$

where β_R is the vector of RP parameters and μ_R is the RP scale parameter which is inversely related to the variance of the error term (Swait and Louiviere, 1993). With identical elements in the X_{ij} vector, a similar model results when the MNL is used to estimate the SP choice probabilities

$$P_{Sij} = \frac{\exp(\mu_S \beta_S' X_{ij})}{\sum_{j=1}^{J} \exp(\mu_S \beta_S' X_{ij})} \tag{4}$$

When RP or SP data are estimated separately the scale parameter is arbitrarily set equal to one, as in all discrete choice modeling (Maddala, 1983). When RP and SP data are stacked and estimated jointly, it is common for the error terms that result from the different data to have unequal variance leading to unequal scale parameters. When $\mu_R = 1$ it is typical for the SP data to have a higher variance due to the unfamiliarity of the choice task and $0 < \mu_S < 1$. The difference in the scale parameter will cause the MNL coefficients to differ. In this case the RP coefficient estimates will be larger than the SP coefficient estimates, $\beta_R > \beta_S$, indicating that the characteristics of the SP choices have an unduly large effect on each choice, relative to the RP data. The relative scale factor in a stacked data set can be estimated as described by Swait and Louviere (1993) and Hensher, Rose and Greene (2005).

It is not necessary for the RP and SP coefficients vectors to contain the same elements. In fact, it is typical for RP data sets to have some variables that do not change across individuals. In this case, the SP data must be used to identify the effect of the variable on the choices (Adamowicz, Louviere, and Williams, 1994). Also, researchers often encounter RP and SP data sets that exhibit different signs on coefficients for the same variables. In these cases it is possible to jointly estimate the coefficients that have the same signs, estimate the opposite signed coefficients separately, and test the restrictions after the appropriate scaling.

In the transportation literature the typical application is to the choice of transportation modes. The RP data involves the current choice and the SP data involves a future choice among current and new (e.g., new train lines, electric cars) transportation alternatives. Ben-Akiva and Morikawa (1990) present the first application of the jointly estimated RP-SP model. Their results are typical to most of the later applications. First, when the RP model is estimated separately, key parameters are not significantly different from zero due to multicollinearity. When the SP model is estimated separately, choice of the new alternative is overstated leading to an upward bias on the associated alternative specific constant. When the data are combined and jointly estimated, the scale parameter indicates higher variance in the SP data. When the scale parameter is used to adjust the SP parameters, the SP parameter estimates are similar to those of the RP model.

The conjoint analysis RP-SP model is increasingly being applied in environmental economics. Adamowicz, Louviere, and Williams (1994) present the first application for water based recreation site choice. Their results are typical of the literature. The RP data are limited due to collinearity and the SP data are prone to hypothetical bias. The RP and SP data can be combined once the scale factor is used to adjust the SP coefficients. Adamowicz, et al. (1997) present an application to moose hunting with a focus on a comparison of RP-SP models with objective or subjective characteristics of the choices. They conclude that the RP-SP model with subjective characteristics outperforms the other models.

Earnhart (2001, 2002) presents an application of the discrete choice RP-SP model to housing choice. The RP component is a discrete choice hedonic model of recent home sales for a sample of recent home buyers. The SP choice set includes randomly drawn houses that were available at the same time as the purchase. The SP choice experiment attempts to simulate the same decision for the households in the recent home sales sample. The focus in Earnhart (2001) is the valuation of water-based (e.g., wetlands) and land-based amenities (e.g., forests). Amenity values are widely divergent for RP and SP models. The joint RP-SP models improves estimation of the amenity values.

3.1.2 *Serial Correlation*

Respondents typically provide multiple SP observations. It is possible that an individual specific error term arises that is correlated across these multiple observations (Hensher, 2001)

$$U_{ijt} = \beta' X_{ijt} + \varepsilon_{ijt} \tag{5}$$

where t is the number of RP and SP observations for each respondent i, t = 1, …, T. In this model the error term is non-IID. The problem with pooling these data is that the correlation in error terms across respondents is ignored. Each observation in the data is assumed to be a separate individual with their own error term instead of each individual having a common error term across observations. Simply pooling data can lead to inefficient and inconsistent coefficient estimates. In other words, standard errors can be inflated and parameter estimates will not approach the true parameters. Swait, Adamowicz, and van Bueren (2004) show that a jointly estimated RP-SP time-series model produces significantly different welfare estimates when compared to a jointly estimated RP-SP cross-section model using the same data.

The correlation in error terms could be treated as a random effect, serial correlation, or a combination of a random effect and serial correlation. For example, with a random effects specification the error term in the equation is decomposed into two components

$$\varepsilon_{ijt} = e_{ij} + v_{ijt} \tag{6}$$

where e_{ij} is the individual specific error term that is carried across each question the respondent faces (or concept that the analyst is measuring – e.g., WTP – using individual data) and v_{ijt} is the random error term that is independently and identically distributed. With a serial correlation error structure the error term is also decomposed into two components

$$\varepsilon_{ijt} = \rho\varepsilon_{ijt-1} + v_{ijt} \tag{7}$$

where ρ is the first-order autocorrelation coefficient and ε_{ijt-1} is the lagged error term. The magnitude of the estimate of ρ measures the strength of the serial correlation. With both random effects and serial correlation the error term is decomposed into three components

$$\begin{aligned} \varepsilon_{ijt} &= e_{ijt} + \eta_{ijt} \\ \eta_{ijt} &= \rho\eta_{ijt-1} + v_{ijt} \end{aligned} \tag{8}$$

In an empirical application of transportation choices, Hensher (2001) shows that the unrestricted error models are superior to the standard MNL model. The major benefit comes from the random effects specification. Serial correlation is statistically significant without random effects but not statistically significant with random effects. The policy implications are significant. The value of travel time savings is biased downward by almost 50% in the standard MNL model.

3.1.3 *Heteroskedasticity*

The jointly estimated RP-SP model adopts the IID assumption. Louviere, Hensher and Swait (2000) describe heteroskedasticity in a discrete choice model as emerging from error variance differences between individual decision makers, resulting in biased and inconsistent parameter estimates when IID is imposed. As such, heteroskedastic models have emerged in recent years to not only correct for non-constant error variance in discrete choice models but also to capture it as a behavioural phenomenon to predict choice.

Louviere, Hensher and Swait (2000) identified several modeling approaches that allow all unobserved components of utility to predict choice. First of all, the Heteroskedastic Extreme Value Model allows for free variance of all alternatives in the given choice set; the main contribution of the HEV is the relaxing of IIA and allowing for pairs of alternatives to be correlated (Bhat, 1994 and Allenby and Ginter, 1995). Second, the Covariance Heterogeneity model allows for covariance heterogeneity to be a useful source of behavioral information, with cross-substitutions of alternatives being allowed to vary. Next, the Random Parameters Logit relaxes both the IID and IIA assumption in the completely unrestricted model; conceptually the RPL is similar to the random effects model in classical linear regressions. Lastly, the Multinomial Probit model allows for free variance, however, it assumes normally distributed parameters in the utility function.

In an empirical application Hensher, Louviere, and Swait (1999) test for homoskedastic errors in RP, SP and RP-SP models. They find that the RP data exhibits heteroskedasticity with the parameterized heteroskedastic model. The errors in the SP data are homoskedastic. When the data are combined the authors reject the hypothesis of parameter equality even when accounting for heteroskedasticity within the sources of data. The major problem is contrasting signs on coefficients in the RP and SP data. When the constraint on equality of signs on the troublesome coefficients is relaxed the likelihood ratio test indicates that the data can be combined.

3.1.4 *Mixed Logit Model*

The mixed logit model has become popular with RP-SP researchers. The mixed logit was developed in response to recognition of some of the limitations of the MNL including the IIA and IID errors assumptions. The mixed logit is significantly more difficult to estimate relative to the conditional and nested logit but has become more practical with increasing computing power and the development of simulated maximum likelihood estimation (Train, 1999). Still, the mixed logit model is much more complex due to specification issues such as choosing (1) the variables which have random parameters, (2) the distribution of random parameters, and (3) the simulation method and number of random draws (Hensher and Greene, 2003).

The mixed logit employs the same random utility theoretical framework as the multinomial logit model. The divergence in the two models is the decomposition of the error term into two components: a mean zero random error term that depends on independent variable and parameter vectors and a mean zero random error term that is IID. The random parameters logit model is a special case of the mixed logit model in which the distribution of the non-IID error term depends on a vector of parameters and individual specific variables, attribute specific variables, or both.

Specification of the error terms and the error variance can be complex. First, it is likely that the variance between RP and SP models will differ. Therefore, it is necessary to include a scale parameter in jointly estimated RP-SP models. The scale parameter can be estimated iteratively or jointly with the parameters. Brownstone, Bunch and Train (2000) set the RP scale parameter equal to one and estimate the SP scale parameter using full information maximum likelihood. They find that the scale parameters are not statistically different.

Mixed logit models, through the specification of the random error term, allow choices to be correlated. In addition to scale, Bhat and Castelar (2002) raise three issues for RP-SP studies: inter-alternative error structure, unobserved heterogeneity, and state-dependence. Inter-alternative error structure is simply heteroskedasticity across choice alternatives. Unobserved heterogeneity is the contribution to the error term that is common across individuals in choice experiment panel data. State-dependence is the influence of previous choices on the current choice in choice experiment panel data. Previous research has considered these issues separately but Bhat and Castelar (2002) argue that it is important to consider them together to account for interactions and avoid specification error. All of these issues can be addressed with the mixed logit model. In a comparison of a pooled logit model that assumes the data are a cross section and a mixed logit that acknowledges the panel-nature of the data, the authors find that the logit model overstates the effects of time and cost on the transportation choice. Scale differences are not observed in the mixed logit with panel data suggesting that scale difference is a function of an inappropriate assumption about the error terms.

The only application of a jointly estimated RP-SP mixed logit model in the environmental economics literature is Boxall, Englin, and Adamowicz (2003) who apply the model to recreation site choice. They adopt a state dependence error structure in which subsequent choices depend on previous choices. State dependence is empirically important suggesting that the MNL comparison models are mis-specified. As in Bhat and Castelar (2002) when state dependence is controlled, there are no scale differences between RP and SP data.

3.2 *Continuous Choice Models*

The continuous choice models are typically used to combine the travel cost method and the contingent behavior method. The revealed preference component of the survey might ask for the number of recreation trips taken during the past month, season, or year. The stated preference component of these surveys would then ask for the hypothetical number of trips that would be taken (or would have been taken) in the next month, season, or year. Hypothetical scenarios are typically constructed for changes in prices and/or quality. Continuous choice RP-SP studies tend to focus on recreation but other applications appear in the literature, such as household disposal of solid waste (Nestor, 1998) and seafood consumption (Huang, Haab, and Whitehead, 2004).

According to the utility theory underlying choice models, the following indirect utility function is obtained from the utility maximization problem

$$v(p, y, q) = \max U(x, y, q) \text{ s.t. } y = px \tag{9}$$

where x is the number of trips, p is a vector of travel costs, y is income, and q is a vector of quality and other site characteristics. The Marshallian demand function is obtained by Roy's identity

$$x(p, y, q) = \frac{-\partial v / \partial p}{\partial v / \partial y} \tag{10}$$

Following the basic framework in Azevedo, Herriges, and Kling (2003), consider the revealed preference demand for *j* = 1

$$x_{i1t}^{R} = x(p_{ijt}^{R}, y_{it}^{R}, q_{jt}^{R}; \beta^{R}) + \varepsilon_{it}^{R} \tag{11}$$

where *i* = 1, …, I individuals, *j* = 1, …, J recreation sites, *t* is the time period, *t* = 1, β^{R} is a vector of parameters, and $\varepsilon_{it}^{R} \sim N(0, \sigma_{R}^{2})$. The stated preference demand for site *j* = 1 with a hypothetical change in trip costs or quality is

$$x_{i1t}^{S} = x(p_{ijt}^{S}, y_{it}^{S}, q_{jt}^{S}; \beta^{S}) + \varepsilon_{it}^{S} \tag{12}$$

where, *t* = 2, …, T is the number of hypothetical scenarios presented to the respondent, β^{S} is a vector of parameters, and $\varepsilon_{it}^{S} \sim N(0, \sigma_{S}^{2})$.

The stated preference price vector may include the hypothetical price change for site *j* = 1. The price change is the sum of the revealed preference price and the hypothetical change in price, $p_{i1t}^{S} = p_{i1t}^{R} + \Delta p_{i1t}$, where $\Delta p_{i1t} \overset{>}{\underset{<}{-}} 0$ is the hypothetical change in price. The stated preference quality vector may include a hypothetical quality change for site *j* = 1. The quality change is the sum of the revealed preference quality and the hypothetical change in quality, $q_{1t}^{S} = q_{1t}^{R} + \Delta q_{1t}$, where $\Delta q_{1t} \overset{>}{\underset{<}{-}} 0$ is the hypothetical change in quality. Income under the stated preference scenarios may differ from income under the revealed preference scenario if the hypothetical question explicitly asks for future trips. Future trip decisions may be based on expected income which can differ from current income, $y_{it}^{S} \overset{>}{\underset{<}{-}} y_{i1}^{R}$. The hypothetical changes in price, quality, or expected income may lead to a change in the number of trips at site j = 1, $x_{i1t}^{S} = x_{i1t}^{R} + \Delta x_{i1t}$ and $\Delta x_{i1t} \underset{<}{\overset{>}{-}} 0$. Trips are negatively related to price changes, positively related to quality changes, and positively related to income changes for normal goods.

Since the data has a common structure it can be stacked and jointly estimated. Several RP-SP studies have used pooled data with constrained coefficients to jointly estimate the model (e.g., Bergstrom et al., 1996; Layman, Boyce and Criddle, 1996; Eiswerth, et al., 2000). The problem with this approach is similar to that of the MNL model; the correlation in error terms across respondents is ignored. Each observation in the data is assumed to be a separate individual with their own error term instead of each individual having a common error term across observations. Pooling data can lead to inefficient and inconsistent coefficient estimates.

Two approaches have been used to deal with pooling data in the continuous choice RP-SP studies: fixed effects and random effects panel data models. Englin and Cameron (1996) compare the pooled data model with the fixed effects model. They find few differences. However, this should not be considered a general result. Loomis (1997), in a discrete choice RP-SP combination, illustrates the divergence between the random effects probit and the pooled probit. The own-price parameter is not significantly different from zero in the pooled probit models. The parameter is significantly different in the random effects models, allowing welfare analysis.

As in the discrete choice RP-SP models, a major question is the consistency of preferences across RP and SP data and hypothetical bias. Consistency is typically tested with SP demand and slope shifters. Consider a semi-log RP-SP panel demand model (with subscripts suppressed)

$$\ln x = \beta_o + \beta_p p + \beta_y y + \beta_q q + \varepsilon \tag{13}$$

A model that tests for a demand and slope shifts from the SP treatment would first define an SP dummy variable as

$$SP = \begin{cases} 0 & \text{if } t = 1 \\ 1 & \text{if } t > 1 \end{cases} \tag{14}$$

The "differentiated" model (Englin and Cameron, 1996) is

$$\ln x = \beta_o^R + \beta_o^* SP + \beta_p^R p + \beta_p^* (SP \times p) + \beta_y^R y + \beta_y^* (SP \times y) + \beta_q^R q + \beta_q^* (SP \times q) + \varepsilon \tag{15}$$

The tests for data consistency are individual t-tests and joint likelihood ratio tests for the interaction parameter vector, $\beta^* = (\beta_o^*, \beta_p^*, \beta_y^*, \beta_q^*)$. If equality of parameters is rejected the SP parameters are equal to the sum of the RP and interaction parameters, $\beta_k^S = \beta_k^R + \beta_k^*$, where k = o, p, y, and q.

Most continuous choice studies find that RP and SP preferences are not consistent. For example, several studies find that the constant and own-price parameters differ between RP and SP data (Englin and Cameron, 1996; Whitehead, Haab, and Huang, 2000; Azevedo, Herriges, and Kling, 2003). Stylized results are that, $\beta_o^* > 0$ and $\beta_p^* > 0$ indicating that the SP quantity intercept is greater than the RP quantity intercept (i.e., $x^R < x^S$) and the SP demand is more elastic than the RP demand. One interpretation of the lack of consistency is hypothetical bias. Respondents are allowed to report trips that do not fully reflect their actual behavior because they are not faced with the discipline of budget and time constraints. Under this interpretation, a useful result from combining RP and SP data is the "calibration" of SP data by counterfactually eliminating the hypothetical bias. Calibration is achieved by setting the SP dummy variable equal to zero, $SP = 0$, and proceeding with the analysis.

Another interpretation of the lack of consistency between the RP and SP data is due to differences in errors between RP and SP data (Azevedo, Herriges, and Kling, 2003). Errors in RP data may be due to recall errors, digit bias, and measurement error in the price variables. Errors in SP data may be due to lack of familiarity with the quality change or uncertainty about future behavior. Because the RP and SP contexts are inherently different, the differences in the data may be expected. These different sources of error can lead to heteroskedasticity between the RP and SP data which can be corrected in conceptually similar ways to the scale parameter adjustment in MNL models. Azevedo, Herriges, and Kling (2003) test for differences in the variance of the error terms between RP and SP data and find no difference. This should not be considered a general result because no other study, to date, has examined the issue.

A lack of consistency between RP and SP data will affect the welfare estimates. Since many studies find that SP demands are more elastic, a common result is that the SP consumer surplus estimates will exceed the RP consumer surplus estimates (Englin and Cameron, 1996; Eisworth et al. 2000). If the lack of consistency in RP and SP data is due to hypothetical bias, the consumer surplus estimates will converge when the SP dummy variable is set to zero.

Many continuous choice RP-SP studies employ only one RP observation and one SP observation for each respondent, $t = 2$ (Rosenberger and Loomis, 1999; Eisworth, et al., 2000; Hanley, Bell, and Alvarez-Farizo, 2003; Azevedo, Herriges, and Kling, 2003). In addition to the limited efficiency gains of a single hypothetical scenario, when studies include only one SP question in a survey to measure the effects of quality change, $t = 2$, and quality change is measured by a dummy variable, it is not possible to test for consistency between the RP and SP data (e.g., Rosenberger and Loomis, 1999; Hanley, Bell, and Alvarez-Farizo, 2003). The SP dummy variable measures the same behavior as the quality change dummy variable. The implicit assumption is that the RP and SP data exhibits the same underlying preferences.

In the context of recreation demand, one approach to deal with this problem is to ask for the number of future trips that are planned with the same price and quality (e.g., Whitehead, Huang, and Haab, 2000; Egan and Herriges, 2006). RP and SP trips under identically observed conditions are not necessarily equal. Several variables may lead to a difference in *ex-post* and *ex-ante* trips including changes in expected income and family size. Asking this additional SP question and including the observation allows for the isolation of the effect of the SP treatment on the number of trips. If the demand intercept is positively affected by the SP dummy variable then the consumer surplus estimate based on the difference between SP trips with improved quality and RP trips with current quality will be greater than a similar difference in SP trips. Again, setting the SP dummy variable equal to zero will lead to equality in the two measures of consumer surplus and a mitigation of the hypothetical bias.

3.3 *Mixed Choice Models*

Mixed choice models have been used to combine the contingent valuation method and the travel cost method. These models typically employ utility-theoretic specifications of the willingness to pay function and specify constraints to be placed on parameters in the jointly estimated model. There is much theoretical research linking willingness to pay and demand for market commodities linked to environmental attributes (McConnell, 1990; Bockstael and McConnell, 1993; Whitehead, 1995; Cunha-e-Sá and Ducla-Soares, 1999).

For example, consider the willingness to pay for site access

$$v(\overline{p}, y, q) = v(p, y - WTP, q) \tag{16}$$

where $\overline{p}$ is the choke price. When respondents are faced with a randomly assigned price increase, A, the probability that they will respond yes to a discrete choice question (i.e., they are willing to pay A and continue taking trips) is

$$\Pr(yes) = \Pr[v(\overline{p}, y, q) - v(p, y - A, q) \geq \varepsilon_o - \varepsilon_1] \tag{17}$$

where ε_o and ε_1 are the random errors attached to the utility functions due to the unobserved heterogeneity across respondents. A utility function linear in income yields the linear functional form for the difference in utilities

$$\Delta v = \alpha_o + \alpha_1 p + \alpha_2 y + \alpha_3 q + \alpha_4 A \tag{18}$$

McConnell (1990) shows that the quantity demanded can be recovered from the utility difference function

$$\frac{\partial \Delta v}{\partial p} = \frac{\alpha_1}{\alpha_4} = x(p, y, q) \tag{19}$$

The link between the recreation demand and utility difference can be used to combine the RP and SP data. Whitehead (1995) provides a similar result for the willingness to pay for a quality improvement.

3.4 *Applications*

In the first empirical application, Cameron (1992) combines the contingent valuation method and the travel cost method in a recreational fishing application. The recreation demand model is based on RP data (ex-post trips). The contingent valuation model is a discrete choice (i.e., yes/no) of the willingness to continue taking trips when faced with a randomly assigned increased total cost of trips. The combined data allows the demand function to be more fully understood. The RP data allows estimation of only a portion of demand. The portion of the demand function at prices between the current price and the choke price is measured with the SP data.

The discrete and continuous choices are jointly estimated with the predicted trips included as an independent variable on the right hand side of the discrete choice model. In order to derive this model, an explicit functional form for the utility function is assigned. The functional forms for the difference in utility and the recreation demand are derived from the functional form of utility

$$\begin{aligned} \Pr(yes) &= \Pr\left[\Delta U(p, y, \hat{x}(p, y), A) \geq e_1\right] \\ x &= x(p, y) + e_2 \\ \rho &= corr(e_1, e_2) \end{aligned} \tag{20}$$

where $\hat{x}(p, y)$ is predicted trips and ρ is the correlation in the error terms. The error terms are distributed bivariate normal. The link between the two decisions is that the discrete choice is a function of recreation trips, common parameters between the utility difference and demand function are constrained to be equal, and the errors are correlated. Statistical tests for the consistency of the RP and SP data reject consistency.

Kling (1997) presents a simulation exercise illustrating the efficiency gains from combining contingent valuation and travel cost methods data. Instead of specifying the utility function, Kling specifies a semi-log demand function and integrates back to the corresponding utility function. The rest of the model is similar to the Cameron (1992) model. The willingness to pay for access is defined with this utility function. The simulation experiment indicates that jointly estimating the demand and willingness to pay models leads to significant efficiency gains and significant reductions in bias. The gains increase as the sample size falls and as the correlation between demand and willingness to pay falls. Gillig et al. (2003) extend Kling's model to allow for truncation of the recreation demand and apply the model with survey data. They find that the precision of willingness to pay is improved and that willingness to pay is bound from above by the contingent valuation method estimate and from below by the travel cost method estimate.

Niklitschek and León (1996) develop a model that links SP recreation trips with the willingness to pay for quality improvement. This application is implicitly a combination of RP and SP data because the current number of trips taken with degraded quality is zero. Therefore the SP trips are a measure of the change in trips from degraded quality (RP) to improved quality (SP). Their theoretical approach is identical to Kling (1997) but the willingness to pay is for a quality improvement

$$v(p, y, q^o) = v(p, y - A, q') \tag{21}$$

The probability of willingness to pay a randomly assigned increase in the water bill, A, is based on the utility difference. The joint model is based on the utility difference function

$$
\begin{aligned}
\Pr(yes) &= \Pr\left[\Delta v(p, \Delta q, y, x) - A \geq +e_1\right] \\
x &= x(p, q, y) + e_2 \\
\rho &= corr(e_1, e_2)
\end{aligned}
\tag{22}
$$

where $\Delta q = q' - q^o$. The model is tested with utility difference and demand parameters constrained as theory suggests. The constrained model is superior to the unconstrained model indicating that preferences are consistent between the RP and SP data. The welfare estimates from the jointly estimated model are between the welfare estimates from the independently estimated willingness to pay and recreation demand models.

Huang, Haab, and Whitehead (1997) diverges from the other jointly estimated mixed models by combining discrete choice willingness to pay data and the change in trips that result from a quality improvement. The change in trips is measured in two ways: (1) the difference in trips measured with RP data with current quality and SP data with improved quality and (2) the difference in trips measured with SP data with current quality and SP trips with improved quality. Also, the variation function is used to model the willingness to pay for the quality improvement. The variation function is based on the expenditure function

$$e(p, q, u) = \text{Min } px \text{ s.t.} \bar{u} = u(x, q) \tag{23}$$

Willingness to pay is the difference in expenditure functions

$$WTP = e(p, q^o, \bar{u}) - e(p, q', \bar{u}) \tag{24}$$

Substitution of the indirect utility function into the willingness to pay function yields the variation function

$$
\begin{aligned}
s &= e(p,q,v(p,q',y)) - y \\
&= s(p,\Delta q,y)
\end{aligned} \tag{25}
$$

Two models are jointly estimated with and without constraints on parameters that theory suggests should be equal. The first model considers the consistency of RP and SP behavior and the second considers the consistency of SP and SP behavior.

$$
\begin{aligned}
\Pr(yes) &= \Pr\left[s(p,\Delta q,y) - A \geq \varepsilon_1\right] \\
x^m(q^o) &= f[x^S(q')] + \varepsilon_2 \\
\rho &= corr(\varepsilon_1,\varepsilon_2)
\end{aligned} \tag{26}
$$

where $m = R, S$. The link between the models is that, with a linear variation function, the parameter on the own-price coefficient should be equal to the change in trips with the quality improvement adjusted by the marginal utility of income (Whitehead, 1995). The results indicate that the RP and SP behavior is not consistent but the SP and SP behavior is consistent. This is a result similar to that found by jointly estimating the same continuous choice RP-SP data (Whitehead, Haab, and Huang, 2000) and suggests that calibration is necessary to mitigate hypothetical bias. Cunha-e-Sá et al. (2004) re-analyze these data within the Cunha-e-Sá and Ducla-Soares (1999) mixed demand framework and come to similar conclusions.

Pattanayak (2001) presents an exception to the mixed choice applications described. He uses production theory to consider water quantity (from protected ‘upstream watersheds’) to be an ecosystem input into agricultural production on downstream farms. Contingent valuation data on farmers’ willingness to pay for this environmental input is combined with farm profit data, with parameter restrictions across WTP and profit equations generated by household production theory, which is a special case of the utility theory presented above. Estimated coefficients are consistent with theory and are used to calculate marginal profits of water and WTP for the ecosystem service. Although parameter restriction across the combined profit and CV models is rejected, joint estimation achieves theoretical consistency and allows the estimation of marginal and non-marginal values, as well as household specific perception of the ecosystem service.

Each of the above models fully integrates the RP and SP data in a jointly estimated model with constraints on parameters. There are a number of less restrictive empirical models that can be used to jointly estimate the data from the contingent valuation method and the travel cost method (and other RP methods). The least constraining empirical approach is to estimate the RP and SP models "side by side" without constraints on parameters but allowing the error terms to be correlated. Whitehead (2005a) pursues this strategy with a linear willingness to pay model estimated with a demand change model. The demand change is the difference in SP trips with a quality improvement and RP trips with current quality. The models are linked by including the predicted demand change, to avoid endogeneity, as an independent variable in the willingness to pay model and by allowing the errors to be correlated

$$\begin{aligned} WTP &= s(\Delta q, \Delta\hat{x}(\Delta q, y, D), y) + \varepsilon_1 \\ \Delta x &= f(\Delta q, y, D) + \varepsilon_2 \\ \rho &= corr(\varepsilon_1, \varepsilon_2) \end{aligned} \tag{27}$$

where $\Delta x = x^S(q') - x^R(q^o)$. Note that the own-price of access does not appear in either equation. This allows inclusion of trips as a proxy for the theoretically preferred own-price (McConnell, 1990). Theory does not suggest restrictions across equations so consistency tests can not be conducted. Some gains in efficiency are achieved by joint estimation.

Most recently, Eom and Larson (2006) jointly estimate recreation trip demand and willingness to pay for a water quality improvement, Δq, and decompose total value into use and nonuse value. Beginning with an explicit functional form for demand, they integrate back to the expenditure function and derive the willingness to pay function that includes nonuse value (*NUV*). The estimated model is

$$\begin{aligned} WTP &= f(\Delta x(\Delta q), NUV(\Delta q)) + \varepsilon \\ x &= f(p, q, y) + \eta \\ \rho &= corr(\varepsilon, \eta) \end{aligned} \tag{28}$$

Eom and Larson improve upon the previous mixed models by estimating nonuse values in a theoretically consistent model.

4. Predictive Validity

A crucial issue is the predictive validity of the jointly estimated RP-SP models. In this context, predictive validity of jointly estimated RP-SP data is the ability of the joint models to outperform independently estimated RP and SP models in predicting actual behavior. The limitation of independently estimated RP models is that they may have difficulty forecasting behavior beyond the range of historical experience. The limitation of independently estimated SP models is that hypothetical bias may overstate or understate different choices. The expectation is that the gains from combining RP and SP data leads to an improvement in forecast accuracy and greater predictive validity. There are two types of predictive validity tests: (1) within sample tests and (2) out-of-sample tests.

Within sample tests examine the extent to which a model predicts behavior of individuals in the model. A within sample test is the standard "percentage of correct predictions" in discrete choice models. For example, Adamowicz, et al. (1997) find that the within sample predictive validity of jointly estimated RP-SP models is superior to independently estimated RP and SP models. Brownstone, Bunch, and Train (2000) find that the within sample predictive validity of the jointly estimated RP-SP model is superior to the independently estimated SP model. The percent correct predictions test is criticized because it uses an individual's own information to predict their own choices. Also, the percentage of correct predictions statistic can appear to be relatively high when it performs no better than a model that predicts at random. Recognizing this, Louviere, Hensher, and Swait (2000) develop a statistic that adjusts the percentage of correct predictions statistic for an improvement over chance. Nevertheless, for purposes of comparing RP, SP, and RP-SP models, the percentage of correct predictions provides valuable information on the relative predictive validity of each model.

Out-of-sample tests predict the extent to which a model predicts behavior of individuals outside the model. Out-of-sample tests are more powerful than within sample tests because the model does not use a respondent's own information in predicting their own behavior. Out-of-sample tests can be conducted with panel data by partitioning the sample into early, $t = 1, \ldots, r$, and late, $t = r, \ldots, T$, time periods (or any sub-groups of the sample). The early time period RP, SP and RP-SP models can be used to predict RP behavior in the late time period.

There have been several applications of out-of-sample predictive validity tests in the RP-SP literature. Ben-Akiva and Morikawa (1990) examine choices with a new transportation mode choice available to respondents. Data on the actual choices was collected after the transportation mode change was realized. After adjusting for respondent confusion relative to the actual choice question, they find that the RP, SP and joint RP-SP models perform similarly in predicting actual choices but only after the hypothetical bias in the SP data is corrected.

Grijalva et al. (2002) conduct an out-of-sample predictive validity test of rock climbing trip behavior using panel data. Respondents are surveyed about their RP trip behavior and SP behavior under future access conditions. Following the realization of the hypothetical scenarios (closure of rock climbing areas), respondents are surveyed again to determine if their SP behavior is able to predict their future RP behavior. With hypothetical closure of rock climbing areas, SP rock climbing trips fall. When the areas are actually closed, RP trips differ in the expected direction to the SP trips. A conclusion from a combined RP-SP model is that the model is predictive valid. An independently estimated RP model could not be used to make these predictions. Unfortunately, no predictive validity comparison to the independently estimated SP model is made.

Whitehead (2005b) compares the within sample and out-of-sample predictive validity of hurricane evacuation behavior with panel data. Respondents are surveyed about their RP evacuation behavior after low-intensity storms, a discrete choice, and SP behavior after hypothetical low-intensity and high-intensity storms. Two hurricanes followed the survey and respondents are surveyed again to determine their RP behavior. Results are mixed, but the jointly estimated RP-SP model seems to perform adequately in forecasting future behavior with prediction error of less than 20%. As in Grijalva et al. (2002), an independently estimated RP model could not be used to make these predictions and, unfortunately, no comparison to the independently estimated SP model is made.

Out-of-sample tests can also be conducted with cross-section data if holdout samples or cross validation tests are used. The use of holdout samples involves using only a portion of the entire sample in estimation, $n' = n - h$, where n is the sample size and h is the holdout sample.

Behavior is predicted for the holdout sample with the *n'* observations and compared to their actual behavior. Haener, Boxall, and Adamowicz (2001) conduct an out-of-sample predictive validity test with cross-section data. They compare the performance of RP, SP and RP-SP conjoint analysis models in predicting similar recreation site choice decisions of different respondents in a different recreation region. They find that the SP models predict choices as accurately as RP models. The combined RP-SP model predicts best. Gelso (2002) finds that the SP model predicts better than the RP and combined RP-SP model.

A cross validation test is similar to the holdout sample test except the holdout sample is equal to one respondent, $h = 1$. Cross validation models are estimated n times (Layton, 2000) and n predictions are made. No cross validation tests appear in the RP-SP literature. To date there has also been no predictive validity test of the mixed RP-SP models. The major difficulty is that micro panel data on actual willingness to pay is difficult to collect. However, holdout sample and cross validation tests are possible.

In the contingent valuation literature, a large number of studies compare actual willingness to pay obtained from laboratory experiments and field surveys with hypothetical willingness to pay (and willingness to accept) obtained from contingent valuation surveys. Divergence in actual and hypothetical willingness to pay is evidence of hypothetical bias. List and Gallet (2001) perform a meta-analysis of these studies by regressing study characteristics on the hypothetical to actual willingness to pay ratio. Some of their results are relevant to the RP and SP literature. They find that private goods generate less hypothetical bias than public goods. Questions based on familiar behavior (i.e., behavior that leads to use value) will generate less hypothetical bias. These results suggest that SP behavior data should have greater predictive validity than SP willingness to pay responses since willingness to pay data may contain nonuse values.

5. Discussion

A preoccupation with each of the models is whether the data can or should be combined. The issue of whether RP and SP data can be combined is statistical and flows from the paradigm that RP data are superior to SP data. Combining data requires the imposition of constraints across models. If RP and SP parameters are significantly different then the data sources can not be combined without biasing the RP parameters. Another paradigm is a compromise and dominates the literature among practitioners of joint estimation. This compromise view recognizes the limitations of both data sources and the gains from combining data. Under these conditions it should be expected that parameters will differ across data sources. When parameters differ the strategies are to (a) constrain SP parameters to be equal to the RP parameters and avoid hypothetical bias or (b) allow the parameters to vary across each data source. This perspective is also echoed in research on combining SP data with experimental data (Harrison, 2006 and Fox et al.).

A significant amount of effort has been devoted to understanding the statistical issues surrounding data combination. Testing for equality of parameters in the discrete choice models is complicated by the scale parameter since the estimated parameters is the product of the underlying parameters and the scale parameter. The scale parameter is inversely related to the variance of the underlying data source. If the variance of the data source differs (e.g., heteroskedasticity) then the RP and SP parameters will differ. It is possible to estimate the scale parameter in combined RP-SP discrete choice models, as long as there is at least one common explanatory variable between the RP and SP data and one is willing to assert that the corresponding preference parameter(s) only differ by the scale factor.. The dominant approach in the continuous choice literature is to test for parameter equality while ignoring the potential problem of heteroskedasticity across data sources. In the mixed choice literature, the issue of heteroskedasticity has been ignored because the data are not stacked. Future studies that ignore heteroskedasticity in combined RP-SP data should be viewed with caution.

A significant amount of effort has also been devoted to addressing the statistical problem associated with multiple observations from each respondent. The early discrete choice literature pooled the RP and SP data, essentially assuming that the RP and SP data were collected from different samples of the population. The panel nature of the data was ignored for two unavoidable reasons: (1) the preoccupation with the scale parameter and (b) limited computing power. These factors, combined with the complications of the MNL model, restricted the attention that could be devoted to the error term. Recent advances in computing power have allowed the testing of alternative theories of the error term (e.g., random effects, autoregressive models, and the use of the mixed logit model). The treatment of the panel data in the continuous choice literature is not uniform. While statistical software is available that easily allows stacked data to be treated as a panel, incorporation of these models is not prevalent. Future studies that fail to recognize the panel nature of combined RP-SP data should be viewed with caution.

Making an appropriate choice from among the different types of RP-SP models is not always straightforward. There are advantages and disadvantages to each of the types of jointly estimated RP-SP models. The choice of a discrete choice RP-SP model over a continuous or mixed choice RP-SP model must consider the following tradeoffs. The major advantage of the discrete choice models is the ability to analyze complex tradeoffs among the characteristics of choices. Continuous choice and mixed choice models require a much larger number of hypothetical scenarios to achieve the same level of complexity. An increase in the number of hypothetical scenarios can lead to respondent fatigue and related problems. Discrete choice RP-SP models are preferred over the continuous and mixed choice models when the policy context requires consideration of a large number of incongruent or mutually exclusive characteristics. Also, discrete choice models are best suited for addressing substitution between options, whereas continuous models emphasize the choice intensity of specific options. In mixed models, the linkage between these two margins of choice can become very complex; however, this is the case even when one is not mixing RP and SP data.

A limitation of the discrete choice models is that there is little connection between the one-time, discrete choice willingness to pay and the aggregation of the one-time willingness to pay across the continuous choices that is often needed for policy analysis. Parsons, Jakus, and Tomasi (1999) have compared different approaches for combining the discrete and continuous decisions but none addresses the issue of how willingness to pay estimated from a discrete decision would change with repetition. For example, is the willingness to pay for an increase in water quality for a single recreation trip constant over the course of the recreation season? The discrete choice RP-SP models are more appropriate for analysis of one-time decisions (e.g., a car purchase, a one-time trip). The continuous and mixed choice models may be more appropriate for analysis of repeated decisions.

Willingness to pay estimated from discrete choice, continuous choice and mixed choice models is the quotient of parameter estimates, complicating the construction of confidence intervals. Numerical and simulation methods have been developed for constructing confidence intervals for continuous choice recreation demand models and discrete choice willingness to pay. These are easily applied to continuous and mixed choice RP-SP models. A limitation of the discrete choice models is the relative difficulty of constructing confidence intervals for willingness to pay estimates. The uncertainty of the discrete choice willingness to pay estimates must be expressed in a nontraditional and suboptimal manner. Policy analysis which requires the subtlety of a confidence interval, relative to a point estimate, of willingness to pay requires the information available from the continuous and mixed choice RP-SP models.

A final consideration is the issue of the type of value estimate that is elicited from respondents. Continuous choice models are limited to the elicitation of use values for recreation or other types of behavior. Conjoint analysis choice experiments and the CVM can be designed to elicit nonuse values. When nonuse values are expected to be a significant component of the total value of the policy, the discrete choice or mixed RP-SP models should be used (e.g., Eom and Larson, 2006).

5.1 *Challenges and Opportunities for Future Research*

We have focused on the advantages of combining RP and SP data. Yet, RP and SP data may resist combination when the underlying regression parameters are not consistent. Our view is that joint estimation still presents an improvement over independent use of RP and SP data, except in extreme cases of poor quality data. Of course, judgment is needed when modeling and making predictions (see chapter 8 of Louviere, Hensher and Swait, 2000)

A major area for future research is assessment of the relative performance of discrete and continuous stated preference data. Only a few studies consider the predictive validity of stated preference data. These include simple discrete choice - hurricane evacuation (Whitehead, 2005b), complex discrete choices - recreation site selection (Haener, Boxall and Adamowicz, 2001), and continuous choice - recreation visits (Grijalva, et al., 2002). Joint estimation will be more challenging as the complexity of the stated preference choices increase. Consideration of if and when joint estimation is inappropriate, in other words, when the stated preference data is too inaccurate to improve upon the revealed preference data, should be considered in future research.

Research is also needed to statistically assess improve each of the three types of jointly estimated RP-SP models. For example, the statistical issue of confidence intervals for willingness to pay from discrete choice models should be resolved so that practitioners can present information about the uncertainty of willingness to pay estimates to policy makers with relative ease. This is not an intractable problem and one which can be resolved with sufficient attention. Researchers should also continue to provide comparisons between the conditional and nested logit models and the mixed logit model to determine the conditions under which the restrictions of the conditional and nested models are problematic. All three types of jointly estimated models should also pay attention to the econometric issues associated with data collection methods. For example, on-site surveys are prone to endogenous stratification (Egan and Herriges, 2006).

Past research using the continuous choice models has often been limited by the shallowness of the panel data. This limitation may be addressed through combinations with discrete choice preference data; however, an unresolved issue is the optimal number of stated preference scenarios that should be presented to respondents. One of the benefits of stated preference questions is the increased efficiency of using more information from each individual.

Applications of the discrete choice models exploit this advantage while applications of the continuous choice models do not. On the other hand, there is a limit to the number of scenarios that respondents can assimilate before they tire of the survey task. Research is needed to determine the number of additional observations necessary to maximize the econometric efficiency of the panel data models. Also, the dominant approach in the continuous choice literature is to test for parameter equality while ignoring the potential problem of heteroskedasticity across data sources. Future research needs to address the issue of heteroskedasticity across RP and SP data.

Even when deep SP panel data (i.e., with multiple observations per individual) are available for combination with RP data, they typically provide preference data for a single point in time. It may be the case that intertemporal RP data, with multiple observations from different time periods, offer greater opportunities for explaining and predicting demand than static RP-SP combinations. For example Swait, Adamowicz, and van Beuren (2004) estimate a time series discrete choice model of recreation site choices over a 5 month period, and find that it provides a statistically different and more complete representation of behavior than a static version of the model. As suggested in their paper, a potentially even more promising approach might be to combine static SP panel with time series RP panel data. In general, an important challenge for future combined RP-SP research is to more explicitly incorporate intertemporal considerations into these models.

Other varied combinations of RP and SP data are also worth pursuing. Thus far, mixed models have primarily focused on the joint estimation of RP behavioral data and SP willingness to pay data. Future research should pursue joint estimation of continuous choice RP-SP behavioral models and willingness to pay models. For example, an extension of the Cameron (1992) approach is to include information on SP trips after higher travel costs are incurred. In the mixed choice literature, the issue of heteroskedasticity has been ignored because the data are not stacked. Mixed choice applications that estimate data “side by side” may also need to consider the problem of heteroskedasticity.

A wide range of applications also offer potential opportunities for combining RP and SP data. The standard application of RP-SP models is to consumer choices among products, transportation mode, or recreation alternatives with a purpose of developing welfare estimates useful for policy. Application of these methods to other types of household choices (e.g., home production, labor supply, recreation participation) could provide insights into the effects of environmental quality on other decisions relevant to policy analysis (e.g., labor productivity, migration, and market size). For example, Dosman and Adamowicz (2006) use a jointly estimated RP-SP model to examine intra-household bargaining in the context of household vacation choice. Whitehead (2003) uses a jointly estimated RP-SP model to estimate the aggregate costs of hurricane evacuation. Application to the estimation of costs imposed by environmental regulation seems feasible.

In addition to the standard consumer applications, RP-SP data has been combined to assess the decisions of firms (Swait, Louviere, and Williams, 1994; Cooper, 1997; Hubbell, Marra, and Carlson, 2000). Further application of these methods to business firms facing new production choices can be used to assess the extent to which new technologies can be expected to lead to environmental quality improvement. Also, most of the existing RP-SP research has been conducted in the context of developed countries. Applications in which the environment plays a much more critical role in the lives of people (e.g., Pattanayak, 2001) could present valuable insights by focusing on more critical elements of the budget constrained utility maximization process.

Most of the RP valuation methods summarized at the beginning of this paper have been combined with the SP valuation methods in the transportation and recreation fields. A glaring omission is the averting behavior method for health valuation. Combining averting behavior and willingness to pay seems relatively straightforward. For example, Laughland, et al. (1996) present models of averting expenditures and willingness to pay with identical independent variables side by side. Joint estimation, by allowing a correlated error term, might increase the econometric efficiency of both models. Pattanayak and Van Houtven (2003) provide a conceptual model for the joint estimation of willingness to pay and averting behavior related to drinking water quality. Rosada et al. (2006) pursue an application of joint estimation of averting behavior and willingness to pay related to drinking water quality with the bivariate probit regression model. More applications such as this are needed.

The value of risk reductions has been independently estimated with the averting behavior method. Choices such as seat belt use, motorcycle and bicycle helmet use can be used to measure the costs of reducing risk. When combined with the magnitude of risk avoided the value of statistical life can be estimated (Blomquist, 2004). The contingent valuation method has also been used to estimate the value of reducing risk. Future research should explore joint estimation of averting behavior and willingness to pay for risk reductions. Joint estimation may lead to improvements in the controversial estimation of the value of statistical life.

Although combined RP-SP methods offer important advantages and opportunities for future environmental valuation research, these methods also continue to face important challenges. For example, although combining RP and SP data can be effective for addressing endogeneity in RP data, available SP data may not necessarily provide all of the exogenous explanatory variables that are needed. In these instances, methods that combine joint RP-SP estimation with the two stage approaches proposed by Berry, Lehvinson, and Pakes (1999) and others may need to be developed and tested.

Another challenge is what to do when the estimated RP and SP preferences are not the same (e.g., when cross-equation restrictions on parameter equality are rejected). One approach is to impose parameter equality and treat the resulting model as a compromise between the RP and SP data. Another approach is to relax the cross-equation restrictions on preference parameters and allow different portions of the data to inform different portions of preferences. For example, Swait, Louviere and Williams (1994), Pattanayak (2001) and Von Haefen and Phaneuf (2007) propose and examine different strategies for restricting a subset of the parameters to match either the RP or SP data, then using the other data to identify the remaining parameters. Accepting the models with parameter restrictions (even though restrictions are rejected) would suggest that the analyst is more interested in reducing bias at the cost of efficiency in conducting policy analysis. More research is needed to determine how sensitive the welfare and policy implications are to the selected approaches for combining results.

More out-of-sample tests of predictive validity are also needed to better understand the implications of combined RP-SP methods. One of the purported advantages of SP data is the ability to forecast beyond the range of historical experience. Yet, this is largely an assertion. Comparison of behavioral forecasts between RP and RP-SP models are needed. Long-term panel data sets are needed that allow the collection of RP and SP data with time for respondents to experience gradual (e.g., climate change) and rapid (e.g., oil spills) environmental change. These data will provide the raw material needed for predictive validity tests and also the ideal information for policy analysis of these important issues.

Mitigation of hypothetical bias in the contingent valuation method has proceeded with joint estimation with data collected from laboratory experiments (Blackburn, Harrison, and Rutström, 1994; Fox et al., 1998; Harrison, 2006) or with survey design advances (Champ et al., 1997; Cummings and Taylor, 1999). Hypothetical bias in the contingent valuation method may also be mitigated by joint estimation with RP, SP or joint RP-SP models. In the hypothetical bias context, out-of-sample tests of predictive validity are needed for the CVM portion of the mixed models. The key to collection of these data is an experimental design in which the money represented by hypothetical willingness to pay statements can actually be collected in a natural field experiment. Opportunities exist with voluntary donations, fishing and hunting license purchase, or participation in “green” energy programs.

Insights from the emerging field of behavioral economics could also play a role in understanding hypothetical bias. Behavioral economics acknowledges that consumers and firms do not always follow the dictates of neoclassical economic theory. Insights from psychology and sociology, among other social science disciplines, are used to understand aberrant economic behavior. For example, Fujii and Garling (2003) employ attitude theory from social-psychology to improve the predictive accuracy of jointly estimated models. Green et al. (1998) use psychological insights to better understand anchoring bias. In similar ways, hypothetical bias could be better understood, mitigated, and avoided when using RP-SP models to predict future behavior.

Related to behavioral economics, stated preference data beyond that typically considered by economists and joint estimation of SP-SP data should be further explored (Hensher, Louviere and Swait, 1999). Risk perceptions, environmental attitudes and other types of value data are used by non-economists. These are typically included as explanatory variables in behavior or willingness to pay models but their joint estimation could improve economic valuation. There are some examples of this type of joint estimation in environmental economics. Whitehead (2006) considers the role that quality perceptions play in the willingness to pay for quality and Morey, Thacher and Breffle (2006) jointly consider attitudes and recreation site choice. Harris and Keane (1999) provide an example of combination of attitudes and revealed preference data from health economics.

Finally, combination and joint estimation of RP and SP data beyond the confines of the three types of models reviewed here is needed. For example, Pattanayak, Smith, and Van Houtven (2003) propose a variation of the typical meta-analysis in which micro-data (e.g., willingness to pay) is combined and jointly estimated with aggregate data on the same variables (e.g., willingness to pay summaries as reported in meta-analyses). This “micro-macro” joint estimation approach could significantly improve the practice of benefit transfer in many of the same ways that RP-SP (i.e., “micro-micro”) joint estimation leads to improved welfare estimates (Smith, Pattanayak and Van Houtven, 2006). Micro-macro joint estimation could lead to (1) improved statistical efficiency of meta-analyses by increasing sample sizes, (2) reductions in bias of welfare estimates used in benefit transfers by imposing restrictions from theory, and (3) supplementing meta-analyses that suffer from omitted variables with policy relevant variables from micro-data. Further research in micro-macro joint estimation is needed in order to determine the feasibility of realizing these potential gains.

5.2 *Conclusions*

Combination of RP and SP data can exploit the advantages of each data source while mitigating the problems associated with their weaknesses. RP data are constrained by the limited range of conditions to which respondents can be exposed (e.g., historical conditions) and multicollinearity. SP data can be used to expose respondents to new conditions and break the multicollinearity. On the other hand, SP data may suffer from the unfamiliarity of the hypothetical choices whereas RP data are grounded in reality.

Another advantage of combining RP and SP data is the increased efficiency of estimation when combining RP and SP data. Increased econometric efficiency can be realized by employing the additional observations on each respondent in panel data models. Research budget efficiency can be achieved by employing smaller sample sizes while keeping econometric efficiency constant compared to independently estimated RP and SP models.

Jointly estimated RP-SP models will often be superior to independently estimated RP and SP models. There have been numerous applications of this approach in the environmental economics literature but studies that independently estimate RP and SP data still dominate the valuation literature. Future environmental valuation research should continue to explore the opportunities for joint estimation.

5.3 *Postscript*

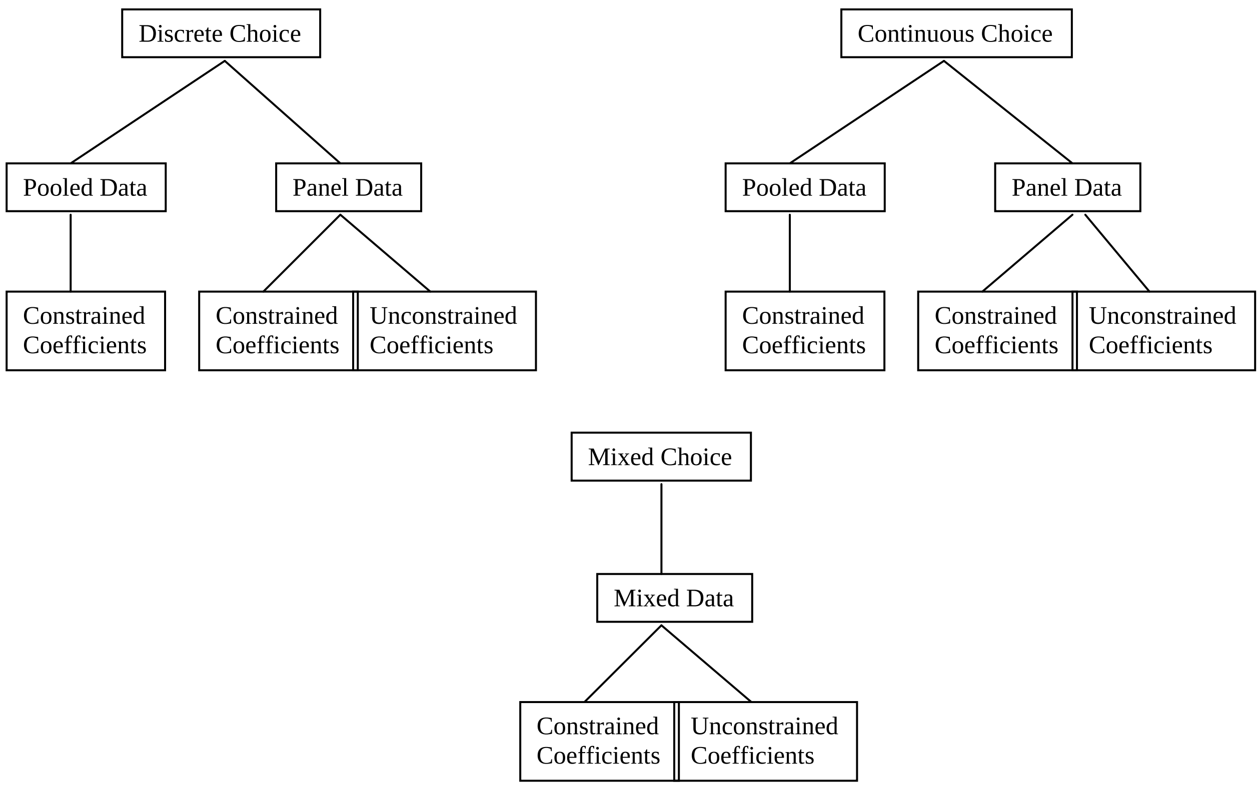
Data combination and joint estimation has evolved to such a point of maturity in the field of environmental economics that the 10th anniversary of the 1996 AERE Workshop was commemorated by an AERE sponsored session at the 2007 American Economic Association annual meeting (i.e., ASSA): “Joint Estimation and Environmental Valuation Ten Years Later.” The AERE session featured four new papers that covered a broad range of applications: (1) a significant advancement in mixed model estimation, (2) a model of attitudes and revealed preference recreation demand, (3) combining data for benefit transfer and (4) incorporating insights from the industrial organization literature into recreation demand analysis. Presentation of these papers and the subsequent discussant comments by the lead authors of the seminal papers in this field signal that data combination and joint estimation will remain an active research area in environmental valuation.

Table 1. Types of Combined Revealed Preference and Stated Preference Data Studies

	Independent and Identically Distributed Errors	Correlated Error Structure
Stacked Data	Pooled Data Studies	Panel Data Studies
Other Forms of Data Combination	Comparison Studies	Mixed Data Studies

Note: All but the comparison studies are jointly estimated.

Figure 1. Types of Econometric Models use to Jointly Estimate Revealed and Stated Preference Data



Acknowledgements

We thank Vic Adamowicz and an anonymous journal referee for many helpful comments. Financial support for this research was provided by the U.S. Environmental Protection Agency's Office of Water under contract 68-C-01-142. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the U.S. EPA.

References

Adamowicz, W., Louviere, J. and Williams, M. (1994) Combining revealed and stated preference methods for valuing environmental amenities, *Journal of Environmental Economics and Management*, 26: 271-292.

Adamowicz, W., Swait, J., Boxall, P., Louviere, J., and Williams, M. (1997) Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation, *Journal of Environmental Economics and Management*, 32: 65-84.

Allenby, G.M. and Ginter, J.L.. (1995) Using extremes to design products and segment markets, *Journal of Marketing Research*, 32(4): 392-403.

Azevedo, C. D., Herriges, J.A., and Kling, C.L. (2003) Combining revealed and stated preferences: consistency tests and their interpretations, *American Journal of Agricultural Economics*, 85: 525-537.

Bateman, I.J., Carson, R.T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Ozdemiroglu, E., Pearce, D.W., Sugden, R., Swanson, J. (2003) *Economic Valuation with Stated Preference Techniques: A Manual*, Edward Elgar.

Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H., and Rao, V. (1994) Combining revealed and stated preferences data" *Marketing Letters* 5(4): 335-350.

Ben-Akiva, M., and Morikawa, T. (1990) Estimation of switching models from revealed preferences and stated intentions, *Transportation Research A*, 24A(6): 485-495.

Bergstrom, J.C., Teasley, R.J., Cordell, H.K., Souter, R. and English, D.B.K (1996) Effects of reservoir aquatic plant management on recreational expenditures and regional economic activity, *Journal of Agricultural and Applied Economics*, 28: 409-422.

Berry, S., Levinsohn, J. and Pakes, A. (2004) Differentiated products demand systems from a combination of micro and macro data: the new car market, *Journal of Political Economy* 112: 68-105.

Bhat, C., (1994) Imputing a continuous income variable for grouped and missing income observations, *Economic Letters*, 46(4): 311-320.

Bhat, C. and Castelar, S. (2002) A unified mixed logit framework for modeling revealed and stated preferences: formulation and application to congestion pricing analysis in the San Francisco bay area, *Transportation Research B*, 36: 593-616.

Blackburn, M., Harrison, G.W. and Rutstrom, E.E. (1994) Statistical bias functions and informative hypothetical surveys, *American Journal of Agricultural Economics*, 76: 1084-1088.

Blomquist, G.C. (2004) Self protection and averting behavior, values of statistical lives, and benefit cost analysis of environmental policy, *Review of Economics of the Household*, 2: 89-110.

Bockstael, N.E., and Strand Jr., I.E. (1987) The effect of common sources of regression error on benefit estimates, *Land Economics*, 63: 11-20.

Bockstael, N. and McConnell, T. (1993) Public goods as characteristics of nonmarket commodities, *Economic Journal*, 103: 1244-1257.

Boxall, P.C.,Englin, J. and Adamowicz, W.L. (2003) Valuing aboriginal artifacts: a combined revealed-stated preference approach, *Journal of Environmental Economics and Management*, 45: 213-230.

Brownstone, D., Bunch, D.S. and Train, K. (2000) Joint mixed logit models of stated and revealed preferences for alternative fuel vehicles, *Transportation Research Part B*, 34: 315-338.

Cameron, T.A. (1992) Combining contingent valuation and travel cost data for the valuation of nonmarket goods, *Land Economics*, 68: 302-317.

Cameron, T.A., Shaw, W.D., Ragland, S.E., Callaway, J.M. and Keefe, S. (1996) Using actual and contingent behavior data with differing levels of time aggregation to model recreation demand, *Journal of Agricultural and Resource Economics*, 21: 130-149.

Carson, R.T., Flores, N.E., Martin, K.M. and Wright, J.L. (1996) Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public foods, *Land Economics*, 72: 80-99.

Carson, R.T., Mitchell, R.C., Hanemann, M., Kopp, R.J., Presser, S. and Ruud, P.A. (2003) Contingent valuation and lost passive use: damages from the Exxon Valdez oil spill, *Environmental and Resource Economics* 25(3): 257-286.

Carson, R.T. and Hanemann, W.M. (2005) Contingent valuation, Chapter 17 in *Handbook of Environmental Economics: Valuation of Environmental Changes*, Volume 2, edited by Karl-Göran Mäler and Jeffry R. Vincent, 821-936.

Champ, P.A., Bishop, R., Brown, T. and McCollum, D. (1997) Using donation mechanisms to value nonuse benefits from public goods, *Journal of Environmental Economics and Management*, 33: 151-162.

Chen, H.Z. and Cosslett, S.R. (1998) Environmental quality preference and benefit estimation in multinomial probit models: a simulation approach, *American Journal of Agricultural Economics*, 80: 512-520.

Clark, D.E. and Kahn, J.R. (1989) The two-stage hedonic wage approach: A methodology for the valuation of environmental amenities, *Journal of Environmental Economics and Management*, 16: 106-120.

Cooper, J.C. (1997) "Combining actual and contingent behavior data to model farmer adoption of water quality protection practices, *Journal of Agricultural and Resource Economics*, 22: 30-43.

Cummings, R.G., Brookshire, D.S. and Schulze, W.D. editors (1986) *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*, Rowman and Allanheld.

Cummings, R.G., and Taylor, L.O. (1999) Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method, *American Economic Review*, 89: 649-665.

Cunha-e-Sá, M.A. and Ducla-Soares, M.M. (1999) Specification tests for mixed demand systems with an emphasis on combining contingent valuation and revealed data, *Journal of Environmental Economics and Management*, 38: 215-233.

Cunha-e-Sá, M., Ducla-Soares, M.M., Nunes, L.C. and Polomé, P. (2004) Consistency in mixed demand systems: contingent valuation and travel cost data, *American Journal of Agricultural Economics*, 86: 444-454.

Dickie, M. (2003) Defensive behavior and damage cost methods, Chapter 11 in *A Primer on Nonmarket Valuation*, edited by Patricia A. Champ, Kevin J. Boyle and Thomas C. Brown, Kluwer.

Dosman, D. and Adamowicz, W. (2006) Combining Stated and Revealed Preference Data to Construct an Empirical Examination of Intrahousehold Bargaining," *Review of Economics of the Household* 4:15-34, 2006.

Earnhart, D. (2001) Combining revealed and stated preference methods to value environmental amenities at residential locations, *Land Economics*, 77: 12-29.

Earnhart, D. (2002) Combining revealed and stated data to examine housing decisions using discrete choice analysis, *Journal of Urban Economics*, 51: 143-169.

Earnhart, D. (2004) Time is money: improved valuation of time and transportation costs, *Environmental and Resource Economics*, 29: 159-190.

Egan, K. and Herriges, J. (2006) Multivariate count data regression models with individual panel data from an on-site sample, *Journal of Environmental Economics and Management*, 52: 567–581.

Eiswerth, M.E., Englin, J., Fadali,E. and Shaw, W.D. (2000) The value of water levels in water-based recreation: a pooled revealed preference/contingent behavior model, *Water Resources Research*, 36: 1079-1086.

Englin, J. and Cameron, T.A. (1996) Augmenting travel cost models with contingent behavior data, *Environmental and Resource Economics*, 7: 133-147.

Eom, Y.-S. and Larson, D.M. (2006) Improving environmental valuation estimates through consistent use of revealed and stated preference information, *Journal of Environmental Economics and Management*, 52: 501-516.

Fillig, D., Woodward, R., Ozuna Jr., T. and Griffin, W.L. (2003) Joint estimation of revealed and stated preference data: an application to recreational red snapper valuation, *Agricultural and Resource Economics Review*, 32: 209-221.

Fox, J.A., Shogren, J.F., Hayes, D.J. and Kliebenstein, J.B. (1998) CVM-X: Calibrating Contingent Values with Experimental Auction Markets, *American Journal of Agricultural Economics*, 80: 455-465.

Freeman III, A.M. (1993) *The Measurement of Environmental and Resource Values: Theory and Methods*, Washington, DC: Resources for the Future.

Fujii, S. and Garling, T. (2003) Application of attitude theory for improved predictive accuracy of stated preference methods in travel demand analysis, *Transportation Research Part A*, 37: 389-402.

Gelso, B. (2002) *Combining Revealed and Stated Preference Data to Estimate the Nonmarket Value of Green Technologies*. Doctoral Thesis. Kansas State University. Department of Agricultural Economics.

Gillig, D., Woodward, R.T., Ozuna Jr., T. and Griffin, W.L. (2003) Joint estimation of revealed and stated preference data: An application to recreational red snapper valuation, *Agricultural and Resource Economics Review*, 32(October): 209-221.

Golob, T., Bunch, D.S. and Brownstone, D. (1997) A vehicle use forecasting model based on revealed and stated vehicle type choice and utilisation data, *Journal of Transport Economics and Policy*, 31(1): 69-92.

Green, D., Jacowitz, K.E., Kahneman, D., McFadden, D. (1998) Referendum contingent valuation, anchoring, and willingness to pay for public goods, *Resource and Energy Economics*, 20: 85-116.

Grijalva, T.C., Berrens, R.P., Bohara, A.K. and Shaw, W.D. (2002) Testing the validity of contingent behavior trip responses, *American Journal of Agricultural Economics*, 84: 404-414.

Grijalva, T., Bohara, A.K., Berrens, R.P. (2003) A seemingly unrelated poisson model for revealed and stated preference data, *Applied Economics Letters*, 10: 443-446.

Haener, M.K., Boxall, P.C. and Adamowicz, W.L. (2001) Modeling recreation site choice: do hypothetical choices reflect actual behavior, *American Journal of Agricultural Economcis*, 83: 629-642.

Hanley, N., Bell, D. and Alvarez-Farizo, B. (2003) Valuing the benefits of coastal water quality improvements using contingent and real behaviour, *Environmental and Resource Economics*, 24: 273-285.

Harris, K.M. and Keane, M.P. (1999) A model of health plan choice: inferring preferences and perceptions from a combination of revealed preference and attitudinal data, *Journal of Econometrics* 89: 131-157.

Harrison, G.W. (2006) Experimental evidence on alternative environmental valuation methods, *Environmental and Resource Economics*, 34: 125-162.

Hensher, D.A. (1994) Stated preference analysis of travel choices: the state of practice, *Transportation*, 21(2): 107-133.

Hensher, D.A. (2001) The sensitivity of the valuation of travel time savings to the specification of unobserved effects, *Transportation Research Part E*, 37: 129-142.

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

Hensher, D., Louviere, J. and Swait, J. (1999) Combining sources of preference data, *Journal of Econometrics*, 89: 197-221.

Hensher, D. A. and Bradley, M. (1993) Using stated choice data to enrich revealed preference discrete choice models, *Marketing Letters* 4(2): 139–152.

Hensher, J., Rose, M., Greene, W.H. (2005) *Applied Choice Analysis: A Primer*, Cambridge.

Herriges, J.A. and Kling, C.L. editors. (1999) *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*, Edward Elgar.

Huang, J.-C., Haab, T.C. and Whitehead, J.C. (1997) Willingness to pay for quality Improvements: should revealed and stated preference data be combined? *Journal of Environmental Economics and Management*, 34: 240-255.

Huang, J.-C., Haab, T.C. and Whitehead, J.C. (2004) Risk valuation in the presence of risky substitutes: an application to demand for seafood, *Journal of Agricultural and Applied Economics*, 36: 213-228.

Hubbell, B.J., Marra, M.C. and Carlson, G.A. (2000) Estimating the demand for a new technology: Bt cotton and insecticide policies, *American Journal of Agricultural Economics*, 82: 118-132.

Kling, C.L. (1997) The gains from combining travel cost and contingent valuation data to value nonmarket goods, *Land Economics*, 73: 428-439.

Larson, D.M., Shaikh, S.L. and Layton, D.F. (2004) Revealing preferences for leisure time from stated preference data, *American Journal of Agricultural Economics* 86(2): 307-320.

Laughland, A.S., Musser, W.N., Shortle, J.S. and Musser, L.N (1996) Construct validity of averting cost measures of environmental benefits, *Land Economics*, 72: 100-112.

Layman, R.C., Boyce, J.R., Criddle, K.R. (1996) Economic valuation of chinook salmon sport fishery of the Gulkana River, Alaska, under current and alternate management plans, *Land Economics*, 72: 113-128.

Layton, D.E. (2000) Random coefficient models for stated preference surveys, *Journal of Environmental Economics and Management*, 40: 21-36.

Leggett, C.G. and Bockstael, N.E. (2000) Evidence of the effects of water quality on residential land prices, *Journal of Environmental Economics and Management*, 39; 121-144.

List, J.A. and Gallet, C.A. (2001) What experimental protocol influence disparities between actual and hypothetical values? *Environmental and Resource Economics*, 20: 241-254.

Loomis, J.B. (1997) Panel estimators to combine revealed and stated preference dichotomous choice data, *Journal of Agricultural and Resource Economics*, 22: 233-245.

Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000) *Stated Choice Methods: Analysis and Application*, Cambridge.

Marka, T.L. and Swait, J. (2004) Using stated preference and revealed preference modeling to evaluate prescribing decisions, *Health Economics*, 13: 563-573.

Morey, E., Thacher, J. and Breffle, W. (2006) Using angler characteristics and attitudinal data to identify environmental preference classes: a latent-class model, *Environmental and Resource Economics*, 34: 91–115.

Mitchell, R.C. and Carson, R.T. (1989) *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Washington, DC: Resources for the Future.

McConnell, K.E. (1990) Models for referendum data: the structure of discrete choice models for contingent valuation, *Journal of Environmental Economics and Management*, 18: 19-34.

Morikawa, T., Ben-Akiva, M. and Yamada, K. (1991) Forecasting intercity rail ridership using revealed preference and stated preference data, *Transportation Research Record*, 1328: 30-35.

Murdock, J., (2006) Handling unobserved site characteristics in random utility models of recreation demand, *Journal of Environmental Economics and Management*, 51: 1-25.

Nestor, D.V. (1998) Policy evaluation with combined actual and contingent response data, *American Journal of Agricultural Economics*, 80: 264-276.

Niklitschek, M. and León, J. (1996) Combining intended demand and yes/no responses in the estimation of contingent valuation models, *Journal of Environmental Economics and Management*, 31: 387-402.

Palmquist, R.B. (1991) Hedonic methods, Chapter 4 in *Measuring the Demand for Environmental Quality*, edited by J.B. Braden and C.D. Kolstad, North-Holland.

Parsons, George R. (2003) The travel cost model, Chapter 9 in *A Primer on Nonmarket Valuation*, edited by P.A. Champ, K.J. Boyle and T.C. Brown, Kluwer.

Parsons, G.R., Jakus, P.M. and Tomasi, T. (1999) A comparison of welfare estimates from four models for linking seasonal recreational trips to multinomial logit models of site choice, *Journal of Environmental Economics and Management*, 38: 143-157.

Pattanayak, S.K., (2001) How green are these valleys? combining revealed and stated preference methods to account for ecosystem costs of deforestation, Working Paper 01_02, RTI International.

Pattanayak, S.K., Smith, V.K. and Van Houtven, G. (2003) Valuing environmental health risks: from preference calibration to estimation, Working Paper 03_04, RTI International.

Pattanayak, S.K. and Van Houtven, G. (2003) Measuring the benefits of the safe drinking water act: a framework for combining CV and averting behavior bata," Research Triangle Institute.

Phaneuf, D.J. and Smith, V.K. (2005) Recreation demand Models, Chapter 15 in *Handbook of Environmental Economics: Valuation of Environmental Changes*, Volume 2, edited by K.-G. Mäler and J.R. Vincent, 671-762.

Portney, P.R. (1994) The contingent valuation debate: why economists should care, *The Journal of Economic Perspectives*, 8(4): 3-17.

Randall, A. (1994) A difficulty with the travel cost dethod, *Land Economics*, 70: 88-96.

Rosado, M.A., Cunha-e-Sá, M.A., Ducla-Soares, M.M. and Nunes, L.C. (2006) Combining averting behavior and contingent valuation data: an application to drinking water treatment in brazil, *Environment and Development Economics*, 11(6): 729-746.

Rosenberger, R.S. and Loomis, J.B. (1999) The value of ranch open space to tourists: combining observed and contingent behavior data, *Growth and Change*, 30: 366-383.

Schläpfer, F., Roschewitz, A. and Hanley, N. (2004) Validation of stated preferences for public goods: a comparison of contingent valuation survey response and voting behavior, *Ecological Economics*, 51:1-16.

Smith, V.K., Pattanayak, S.K. and Van Houtven. G. (2006) Structural benefits transfer: an example using VSL estimation, *Ecological Economics*, 60(2): 361-371.

Smith, V.K. (1991) Household production functions and environmental benefit estimation, Chapter 3 in *Measuring the Demand for Environmental Quality*, edited by J.B. Braden and C.D. Kolstad, North-Holland.

Swait, J., Adamowicz, W. and van Bueren, M. (2004) Choice and temporal welfare impacts: incorporating history into discrete choice models, *Journal of Environmental Economics and Management*, 47: 94-116.

Swait, J. and Louviere, J. (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models, *Journal of Marketing Research*, 30: 305-314.

Swait, J., Louviere, J.J. and Williams, M. (1994) A sequential approach to exploiting the combined strengths of SP and RP data: application to freight shipper choice, *Transportation*, 21(2): 135-152.

Taylor, L.O., (2003) The hedonic method, Chapter 10 in *A Primer on Nonmarket Valuation*, edited by P.A. Champ, K.J. Boyle and T.C. Brown, Kluwer.

Train, K.E. (1999) Mixed logit models for recreation demand, Chapter 4 in *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*, J.A. Herriges and C.L. Kling, editors, Edward Elgar.

von Haefen, R.F. and Phaneuf, D.J. (2007) Identifying demand parameters in the presence of unobservables: a combined revealed and stated preference approach, North Carolina State University, Department of Agricultural and Resource Economics Working Paper.

Vossler, C.A. and Kerkvliet, J. (2003) A criterion validity test of the contingent valuation method: comparing hypothetical and actual voting behavior for a public referendum, *Journal of Environmental Economics and Management*, 45: 631-649.

Wardman, M. (1988) A comparison of revealed preference and stated preference models of travel behavior, *Journal of Transport Economics and Policy*, 22(1): 71-91.

Whitehead, J.C. (1995) Willingness to pay for quality improvements: comparative statics and interpretation of contingent valuation results, *Land Economics*, 71: 207-215.

Whitehead, J.C. (2003) One million dollars per mile? the opportunity costs of hurricane evacuation, *Ocean and Coastal Management*, 46: 1069-1083.

Whitehead, J.C. (2005a) Combining willingness to pay and behavior data with limited information, *Resource and Energy Economics*, 27: 143-155.

Whitehead, J.C. (2005b) Environmental risk and averting behavior: predictive validity of revealed and stated preference data, *Environmental and Resource Economics*, 32(3): 301-316.

Whitehead, J.C. (2006) Improving willingness to pay estimates for quality iImprovements through joint estimation with quality perceptions, *Southern Economic Journal*, 73(1): 100-111.

Whitehead, J.C., Haab, T.C. and Huang, J.-C. (2000) Measuring recreation benefits of quality improvements with revealed and stated behavior dData, *Resource and Energy Economics*, 22: 339-354.