DISSERTATION

EXAMINING EFFICIENCY GAINS THROUGH COMBINING REVEALED AND STATED PREFERENCES, AND ISSUES RELATED TO SCOPE WITH CONTINGENT VALUATION

Submitted by

David J. Gebben

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2013

Doctoral Committee:

Advisor: John B. Loomis

Andrew Seidl
Gregory Graff
Joshua Goldstein

ABSTRACT

EXAMINING EFFICIENCY GAINS THROUGH COMBINING REVEALED AND STATED PREFERENCES, AND ISSUES RELATED TO SCOPE WITH CONTINGENT VALUATION

An increase in the statistical efficiency for non-market valuation techniques is often desired in order to narrow the confidence intervals and provide better policy recommendations for resource managers. This is important to assist the managers in conducting benefit-cost analysis for the scare resources at their disposal. This dissertation examines the gains that come from combining revealed and stated preference data, exploring how estimation techniques can reduce the variance of a WTP amount.

This first parts of this dissertation looks at why resource managers would be interested in methods of combining Revealed and Stated preference data and measurement of the gains. One chapter does this by combining DC CVM with an MNL travel cost study. The following chapter examines the role anchoring can play in DB CVM studies for an onsite user of a beach resource. The final part of this dissertation studies the issue of scope in CVM studies through a meta-analysis. This dissertation finds that, in organizing the collected survey data, there are low cost methods to increase the efficiency of estimators that provide a significant reduction in variance. This reduction is critical for the resource manager wanting to examine if the project or policy would pass a benefit cost test. It also finds that the key factors necessary to reflect scope require more research with CVM.

TABLE OF CONTENTS

CHAPTER ONE: Motivation and Background

The Manager's Problem

The manager's problem is to determine the optimal allocation of scarce resources to competing interests, in a way that maximizes ecosystem goods and services under his responsibility. The manager maximizes the total net present value of ecosystem goods and services for a resource. The manager's problem is clear when the entire social value of the resource is captured by the market or prevailing societal institutions. If the market fails or the institutions are weak, maximizing resource benefits presents a challenge. If the resource has public good attributes, e.g. non-rivalry and non-excludability, then efficient management becomes more difficult.

Economics is a decision science focused on tradeoffs. Resource managers may face tradeoffs among valuable ecosystem goods and services, the value of only some of which are fully reflected in markets. The manager needs a mechanism to value the various interests in a manner that is reasonable and tractable. One of the most useful methods of evaluating competing interests is benefit-cost analysis. It allows for a great number of interests and needs to be weighed against each other and can provide a common metric of comparison (Freeman 2003). Market failures and alternative uses for the resources are outside the focus of this dissertation. Examining public goods valuation and how to more efficiently capture the value of the natural resources will be the focus of this dissertation.

There is a need for common language regarding the role of value and valuation so ecologists and economists can provide useful information to a resource manager or policy maker. Both economists and ecologists understand the interconnectedness of the environment and ecosystem services, benefits that come from the environment and provide for human well-being, for maximizing the strengths of each field. Ecosystem Services directly and indirectly improve human well-being, yet they are rarely fully accounted in the economy. More accurate information for the values from nature enables the manager to make better decisions by having a more complete understanding the good.

Viewing natural resources through a lens of capital assets perhaps comes more easily to economists, but ecologists also term natural resources as “natural assets” or discuss the value of ecosystem services (Dominati, Patterson and Mackay 2010), (Daily, et al. 2009), (Ellison and Daily 2003). Determining the process for environmental valuation requires some understanding of the various interests and values under consideration. Part of the difficulty is the multiple meanings that the term value can take. In order to structure these potentially competing values, some common methods and definitions are required.

It is helpful to first think about from where ecosystem services come. First, the ecosystem has various functions that it performs. From these functions we then get ecosystem services that directly or indirectly benefit people. Ecosystem functions are connected to habitats, catchments, and ecosystems that do not follow political boundaries.

In order to analyze a problem in a meaningful way, all the costs and benefits of changes to the system need to be taken into account and properly evaluated. The services need to be valued and weighed against each other so that a trade-offs can be revealed and the implications of management decisions better understood. After the initial project has been implemented, rigorous evaluation needs to be conducted comparing with and without (Carpenter, et al. 2009).

By mapping the location and flow of the services, we can begin to visualize the services and then model those services. This helps create a good tool for assessing the trade-offs in land cover and land change on ecosystem services. Once that relationship is known, management decisions can be made (de Groot, et al. 2010).

From Ecosystem Function to Economic Valuation

Another way to think about how ecosystem services work is seen in Figure 1.1. In that figure, a diagram is presented that demonstrates the order from ecosystem production, to ecosystem service, and finally valuation method.

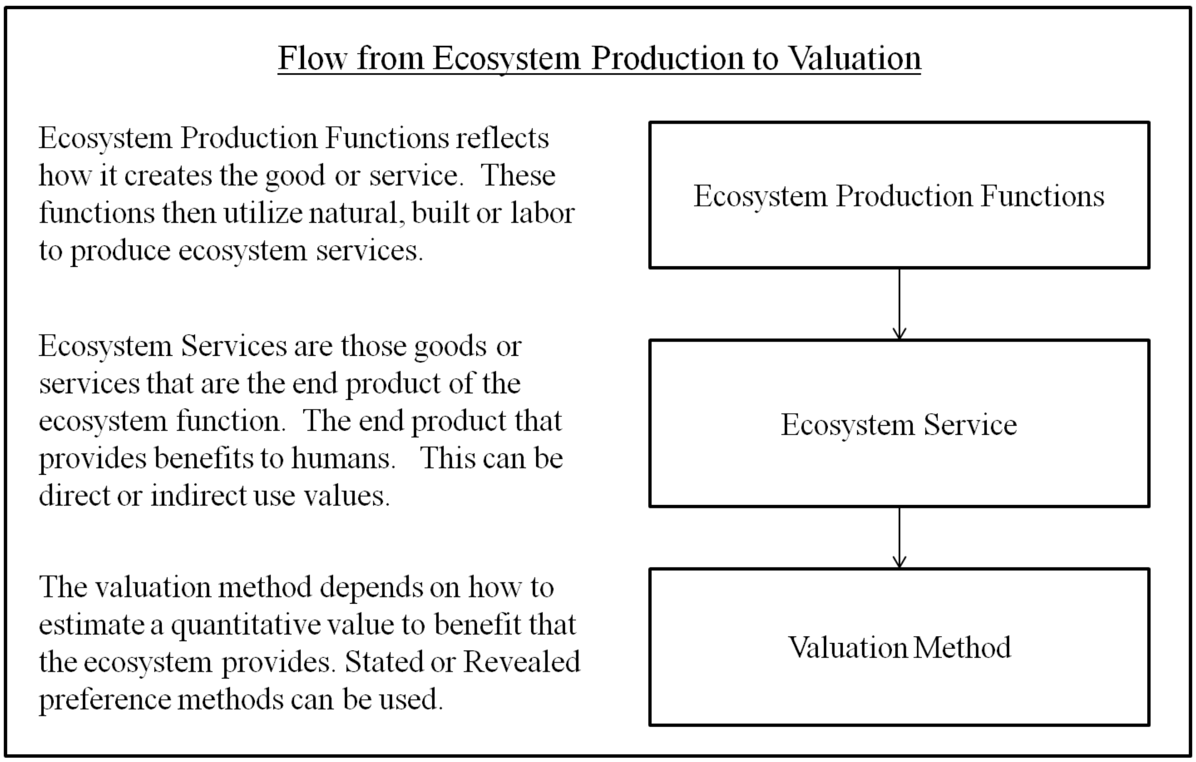

Figure 1.1: Flow of Ecosystem Services

From these production functions, the ecosystem services are produced and both ecologists and economists understand that production functions have spillover effects and often have less obvious connections. Framing the conversation in terms of ecosystem services and goods can be helpful when thinking about ecosystem valuation. Both can be defined as the "benefits people obtain from the ecosystem.[1]" Further investigation of that definition reveals that this often refers to the enjoyment or positive aspects from nature. Ecosystem services are often the end product of the natural ecosystem processes. These benefits include: purification of water and air, maintenance of soil fertility, pollination of crops and natural vegetation, erosion control of soil, and flood and drought mitigation. Ecosystem goods are those products that result from the ecosystem services. A few ecosystem goods are: wildlife and fish, plants (for food, fiber, or fuel), water, air, recreation, and landscape beauty.

To value the relevant goods and services, the researcher needs to determine a few things first. This is partly since as Koopmans pointed out in 1947, measurement without theory provides little in the way of useful predictive powers to policy makers (Koopmans 1947). In choosing which project a manager should finance, it is critical that the structure be well defined and tractable. It is advisable that prior to beginning a project, a few questions need to be answered to determine the state-of-the-art typology regarding the service:

- How can the landscape and the ecosystem characteristics be quantified and related to their associated ecosystem functions?
- What indicators and benchmark values will be used for measurement?
- How can the ecosystem characteristic be mapped or visualized? (de Groot, et al. 2010)

The answers to these questions can then help determine the best choice of options to be selected. Once we have a clear idea of what is possible, it then becomes feasible for trade-off analysis to begin to determine which options are most valued by residents. This also requires recognizing that there are ecological opportunity costs from emphasizing one service provision over another.

[1] Millennium Ecosystem Assessment 2005

Ecologists tend to think of natural resources in terms of the intrinsic value of a natural resource having value in and of itself. The intrinsic value of nature is that the resource has value just for existing. Accordingly, the value of the natural resource is inherent, requiring no anthropocentric reference. Conversely, the ecological value of a resource may be established by the choices made surrounding its use or management.

Economists tend to utilize value in the more instrumental sense of the word, that there is some equivalent monetary value of the resource that can be measured or that can provide a welfare measurement of utility that the resource provides (Freeman 2003). The instrumental value of nature requires that we have a specific goal in mind of what society wants from that resource. The end goal then allows for a discussion about how the various parts of the resource under consideration are utilized or how those other interests tradeoff. This view of value comes from the discussions of neoclassical economic welfare. The basic premise of economic analysis is the increase the welfare of society and Pareto optimality is a litmus test often used. Pareto optimality is the principle that no change should be made that would make another worse off. The 'potential' Pareto principle underlies BCA. This principle requires that the change is economically efficient if the gainers could compensate the losers (those bearing the costs) and the gainers would still come out ahead.

The distinction between intrinsic and instrumental values of natural resources is important, but it is also important to recognize where they have weaknesses and the two value types are much stronger when combined. If a resource unit only has intrinsic value then it is hard to quantify the tradeoffs that need to be determined when making management decisions. If a resource is only valued for the instrumental value it can provide, and then harm to the resource can occur.

The differences between the two values can create policy objectives in conflict with each other. This can be seen when considering the inter-temporal aspect of the valuation of ecosystem goods. If the policy objectives only maximize the instrumental facet of the good, it introduces the possibility that the present welfare maximization scenario might not have a long enough time frame and overvalue current users at the expense of future generations (Harris and Frasier 2002). However, only valuing the intrinsic facet could undervalue the immediate and direct value.

Capturing the Total Economic Value

Total Economic Value (TEV) is the combined value of use and non-use values that when combined provide an estimate of the contribution of the environment to human well being. Use values are often described in one of three ways. The first is direct values, the raw materials and physical products that come from production, consumption and sale of goods or services. The second use value is termed bequest values, the premium placed on preserving resources for future use that may have an economic value. A third type of use value is indirect values, the ecological functions that provide life supporting functions that maintain and protect natural and human systems. Non use values include existence and bequest values of a resource, regardless of current or future uses. The direct values and indirect values are more likely to be captured by market actions, while option and existence values are less likely to be captured by market actions. Figure 1.2 provides a graphical representation of TEV, and provides examples.

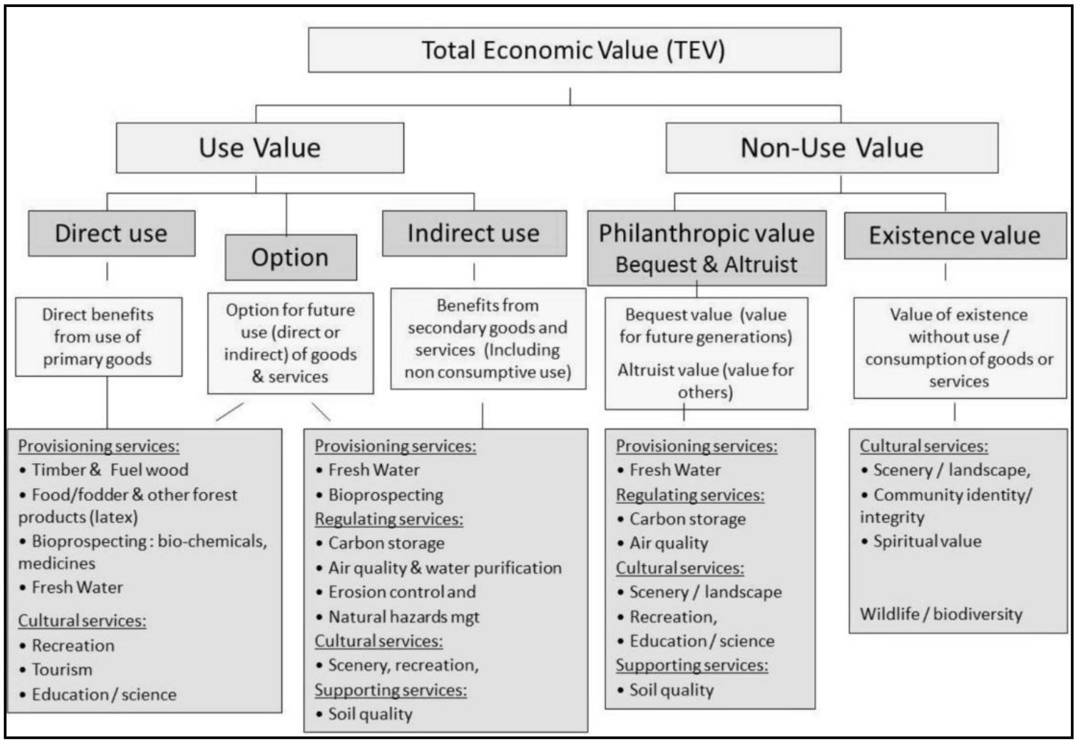

Figure 1.2. Total economic value framework. Source ten Brink et. al. 2011 *Estimating the Overall Economic Value of the Benefits provided by the Natura 2000 Network.*

To estimate the TEV, take use values and non-use values and add them together. Economic valuation falls into one of two categories: 1) direct market valuation, 2) nonmarket valuation. Each of these two values is useful on its own, though they are not exclusive of each other. Both of these values interact with one another to produce TEV. The advantage of TEV is it allows the manager to explicitly take into account values and benefits that would otherwise be undercounted. The basic management logic from TEV is that ecosystem managers often receive few benefits from land uses like forest conservation or watershed protection. The onsite benefits of such stewardship practices are often smaller than what could be provided by alternative land use practices. However, the stewardship practices and conservation measures often have greater benefits to those downstream, including, for example, avoided cost for water filtration or flood mitigation (Engel, Pagiola and Wunder 2008). The manager that understands both on and offsite effects of his/her decisions can help internalize what would otherwise be an externality.

Finding the balance between those two values can be a dilemma but this is where the components of TEV can aid the decision maker. Restoring a resource has a non-use or direct use values to current and future generations even if they do not use it. Thus, even though non-use value is an instrumental value, it may capture some facet of intrinsic values that ecologists emphasize.

Choosing the best unit of measurement is difficult. The focus on just monetary values from the established markets has the simplicity that monetary values are easily compared and added. If one wants to consider costs from production, things becomes less clear since the monetary value from the pollution is difficult to appraise.

Single sector or single purpose management leads to ecologically “unintended consequences” (Tallis and Polasky 2009). This is a principle that also has a long history in economic thought. Shifting towards a TEV for understanding resource values can assist the resource manager to construct a benefit cost analysis of the management problem. TEV could explicitly account for the interconnectedness within the system and recognize the importance between many key services and target functions (Tallis and Polasky 2009).

The first step towards understanding how a manager could conduct a total economic valuation analysis of the ecosystem service value of a given resource would be recognition of the need for marginal ecosystem service assessments. Demand curves are marginal benefit curves, and these curves are essential in formulating and quantifying the role that ecosystem services fulfill in generating human welfare or utility (Fisher, et al. 2008). It is the amount for an additional unit of the good we need to value for ecosystem services. As the resource becomes scarcer, the value ascribed to that source will increase (Fisher, et al. 2008).

The economic value of something is a measure of its contribution to human-wellbeing. Economic values reflect the preferences and actions of a given society (Brown, Bergstrom and Loomis 2007). It provides meaningful quantifiable analysis that can lend itself to a trade-off analysis subject to time and money constraints. An economic valuation is more likely to be relevant and accurate if the ecosystem change being evaluated is small relative to the total production or service of the geographic area of interest (Brown, Bergstrom and Loomis 2007).

Methods for estimating TEV

To analyze the economic value of protecting a given resource area, there are a few methods that could be implemented. Another option for measuring the value of protecting sensitive zones would be to conduct a hedonic study of zones that are in good condition. This would be conducted by examining the sale prices of relevant properties and categorizing the riparian zones according to some predetermined scale. Building the regression model and recovering the implicit price of a characteristic reveals information about the preferences for heterogeneous goods (Taylor 2003). This method, however, requires a large amount of data: a high amount of information about the various property information entering the regression analysis to ensure correct valuation of the value of the attributes that are included (Brown, Bergstrom and Loomis 2007).

Good data that allows for an examination of the marginal changes in ecosystem services are needed to get a better idea of the return on dollars invested. This is where high quality time-series data on land cover integrated into service valuation would be useful (Fisher, et al. 2008). This would also require the explicit acknowledgement of the importance of changes in service delivery across various disturbance states. Another useful source of data would be information relating to service valuations of alternative land use practices. However, these types of data are not as useful since such studies are not as amenable to marginal analysis.

If it is not possible to meet the above criteria, alternative valuation methods must be sought. These methods include stated preference studies, where the researcher creates an artificial market or scenario gauge a populations preferences, an example would a contingent valuation survey where the hypothetical change to the resource good is often framed in the form of a referendum. Other examples are revealed preference surveys where the researcher observes the consumer's activity. An example is the travel cost survey. Both revealed and stated preferences will be explored in greater detail below and in the following chapters.

One alternative method to determine the value of a resource is a stated preference study to be conducted to estimate the value of increasing such areas. This could be useful in helping to determine the TEV. By conducting a contingent valuation survey, respondents can help policy makers "zero in on a specific ecosystem good or service as long as a realistic payment scenario can be posited" (Brown, Bergstrom and Loomis 2007)

The weakness of such a study is that it can be costly to implement. Further, the danger of a low response rate is always present, thus making out-of-sample predictions difficult. Along with that, the risk of protest responses would need to be handled. A "casual" contingent valuation survey often results in a disastrous outcome, and the contingent valuation is still too subjective for its critics (Boyle 2003). Those barriers, however, are not high enough to exclude implementing such a study, just the recognition that it is important to do it well. Low cost methods of improving the efficiency and accuracy of such studies is important, hence a transfer of benefits from similar stated preference surveys may be possible. Once such valuation phase is conducted, it is then possible to conduct meaningful benefit cost analysis that can look at marginal changes.

Management decisions are not created in a vacuum by ecologists or economists. They develop in specific contexts and respond to specific problems being subject to the differing interests of the stakeholders (Engel, Pagiola and Wunder 2008). They are not even necessarily designed by official government agencies (Ostrom 1990). To monitor the effectiveness of a project, it must have a predetermined definition of successful provision of service. This definition can be determined by the relevant stakeholder group or government regulation.

Empirical challenges to valuation measures are numerous. One of these challenges is that of finding and utilizing the relevant data. Often managers are driven more by operational rather than conceptual issues of how to measure. The particular environmental or welfare issue of a country, the available data, funding options, and which components seem most tractable often play a major role in determining what goes into a national wealth account (Hetch 2007).

A final empirical difficulty in TEV that seeks to account for nonmarket values is that while it is possible to construct pseudo markets and conduct a nonmarket valuation study to provide a value, those are difficult and costly to implement. After a survey is done, transferring the benefits up to the national scale is difficult since land is very heterogeneous. The farther one goes from the specific valuation study area, the less valid it will be. At some point an original valuation study may be necessary.

Nonmarket Valuation

Valuation as a tool would best be applied in areas where there is a substantial and well established group of studies. Notably provisioning services for food, raw materials and recreation is a noteworthy example (Atkinson 2010). For the first two, they are close to market and often have and may have a commercial parallel. For recreation, valuation through revealed or stated preference has a long establish tradition (Atkinson 2010).

Nonmarket valuation techniques find values of goods not traded in markets. This can be done either through revealed preference methods or stated preference methods and can provide insights into how people value a resource. After the relevant person or group has been identified and the exact ecosystem service benefit has been described, both spatially and temporally, it is possible to construct a pseudo market valuation of that resource that can then contribute to benefit cost analysis of projects or programs. While there is some debate about the mixing of the market data with nonmarket data in green accounting, it has to be acknowledged that the development of nonmarket valuation tools was done with a great deal of critical reflection, which has resulted in ever greater sophistication in application, and ever greater scrutiny regarding the reliability and validity of the nonmarket values (Atkinson 2010). Both revealed and stated preference methods provide similar welfare measures.

Revealed preference approaches to the valuation of nonmarket goods have a long history, going back at least as far as Hotelling in 1947 when he, in response to a request from the Secretary of the Interior, started developing what became the travel cost model. The question was whether or not it was possible to value the National Parks. Hotelling replied that it is possible by approximating the value the people assign the parks based on how much they would be willing to spend to get there. That is, if a person is willing to spend the time and money to get to a recreation site, then it provides them at least that much value. The researcher observes the behavior and from that infers the value. The consumer is presumed to be making informed decisions under constrained choices, making tradeoffs and determining that the best source of increased welfare would be to visit that particular recreation site. So the researcher can then use travel costs as price of recreation visits and develop a demand curve for recreation visits.

The travel cost then provides a useful predictive model for visitors, types of usage and preferences of the users. The model requires that a price be constructed from responses on visitor travel costs, but some components of the full cost or price may not be observed. The number of trips must be collected through surveys. There are multiple examples of the travel cost model. The single site is the simplest to implement and is useful for valuing access. Though it is hard to value the changes in environmental quality with the single site model, typically there is no change in environmental quality. Multiple site methods are mostly based on various implementations of the random utility model, which will be discussed in greater detail in the following chapter.

To conduct a quick travel cost model, the researcher must:

- Define the site of interest
- Define the relevant substitute sites
- Decide if a system of demands is needed (small sets), or a random utility model (large sets)
- Develop a behavior survey tool to learn what the respondent did and how much they spent doing it
- Learn about the relevant environmental qualities or site attributes
- Estimate the model and interpret the welfare estimates and demand predictions.

For the resource manager seeking to inform management decisions, the travel cost model has many features that lend itself towards a benefit cost analysis and that can provide marginal effects from changes. If there is a large enough set of substitute sites, information about what the effect of the removal of a site would cause can be estimated.

Stated preferences ask consumers what they prefer or what they think is best. This method has the advantage of allowing for a change in an environmental quality or attribute that might not be observable in revealed preference data. Through the careful implementation of survey tools, the researcher can estimate monetary values, the economic effect of choices, or ratings of preferences allowing for meaningful measures of value (Brown 2003). The advantage of a survey is that it allows for the introduction of a new good, a hypothetical change or limiting the choice set. They provide opportunities beyond what revealed preferences alone would provide (Brown 2003).

An estimation method for stated preferences is the contingent valuation method. The first step in implementing a contingent valuation study is to identify the change in the good. This change can be either in quality or quantity. The goal of this first step is to measure the change from some baseline, often the *status quo*, to some new level. The second step is to then determine whose values are to be estimated. This step is important as it requires the researcher to decide if only those with direct use matter or if indirect users of the resource are to be sampled as well. Then the sample mode needs to be determined, followed closely by administering the survey. The final stage is then analyzing and interpreting the data for communication to the relevant stakeholders (Boyle 2003), (Brown 2003).

Both revealed and stated preferences approaches to estimate economic value using a random utility model have advantages and disadvantages. Speed/ease of estimation, accuracy and cost of valuation estimates are important considerations for the resource manager. Combining revealed and stated preference methods can allow for the advantages of both, while mitigating the limitations of both as well (Train 2009). This dissertation will explore further the effects of combining these two valuation approaches and how the efficiency gains can be captured.

In summary, to surmount the theoretical and empirical challenges of incorporating nonmarket valuation techniques into policy and resource decisions, we must use the tools correctly and within the proper framework for which they were intended. To that end the next chapter will provide a more detailed exploration of the random utility model.

WORKS CITED

Atkinson, Giles. "Environmenatal Valuation and Greening the National Accounts Challenges and Initial Practical Steps". Report, Washington D.C.: World Bank, 2010.

Boyd, James. ”Nonmarket benefits of nature: What should be counted in green GDP? Ecological Economics", 2007: 716-723.

Boyle, Kevin J. "Contingent Valuation in Practice." In *A Primer on Nonmarket Valuation*, by Patricia A. Champ, Kevin J. Boyle and Thomas C. Brown, 111-116. Norwell: Kluwer Academic Publishers, 2003.

Brown, Tomas C., John C. Bergstrom, and John B. Loomis. "Defining, Valuing, and Providing Ecosystem Goods and Services." *Natural Resources Journal* 47 (2007): 329-376.

Carpenter, S.R., et al. "Science for Managing Ecosystem Services: Beyond the Millenium Ecosystem Assessment." *Proceedings of the National Academy of Science*, 2009: 1305-1312.

Coase, Ronald H. "The Problem of Social Cost." *Journal of Law and Economics*, 1960: 1-44.

Daily, Gretchen C., et al. "Ecosystem Serivices in Decision Making: Time to Deliver." *Frontiers in Ecology*, 2009: 21-28.

de Groot, R. S., R. Alkemade, L. Braat, L Hein, and L. Willemen. "Challenges in integrating the concept of ecosystem services and values in landscape planning, mangement, and decision making." *Ecological Complexity* , 2010: 260-272.

de Groot, Rudolf S., Matthew A. Wilson, and Roelof M.J. Boumans. "A typology for the classification, description and valuation of ecosystem functions, goods and services." *Ecological Economics*, 2002: 393-408.

Dominati, Estelle, Murray Patterson, and Alec Mackay. "Modeling Ecosytem Services from Soils: A Case Study from New Zealand Dairy Farms". White Paper, Palmerston North: Ecological Economics Research Center New Zealand, 2010, 1-8.

Ellison, Katherine, and Gretchen C. Daily. "Making Conservation Profitable." *Conservation Magazine* 4, no. 2 (2003): 1-8.

Engel, Stefanie, Stefano Pagiola, and Sven Wunder. "Designing Payments for Environmental Services in Theory and Practice: An Overview of the Issues." *Ecological Economics*, 2008: 663-674.

Fisher, Brendan. Turner, Kerry. Zylstra, Matthew. Brouwer, Roy. de Groot, Rudolf. Farber, Stephen. Feraro, Paul. Green, Rhys. Hadley, David. Harlow, Julian. Jefferiss, Paul. Kirkby, Chris. Morling, Paul. Mowatt, Shaun. Naidoo, Robin. Paavola, Jouni. Strassburg, Bernardo. Yu, Doug. Balmford, Andrew. "Ecosystem Services and Economic Thoery: Integration For Policy-Relevant Research." *Ecological Application*, 2008: 2050-2067.

Freeman, A. Myrick III. "The Measurement of Environmental and Resource Values Second Edition". Washington DC: Resources for the Future, 2003.

Harris, Michael, and Ian Frasier. "Natural Resource Accounting in Theory and Practice: A Critical Assesment." *Australian Journal of Agricultural and Resource Economics*, 2002: 139-192.

Hetch, Joy E. "National Environmental Accounting: A Practical Introduction." *International Review of Environmental and Resource Economics*, 2007: 03-66.

Koopmans, Tjalling. "Measurment Without Theory." *The Review of Economic Statistics*, 1947: 161-172.

Ostrom, Elinor. "Governing the Commons." New York: Cambridge Univeristy Press, 1990.

Ruhl, J.B., Steven E. Kraft, and Christopher L. Lant. "The Law and Policy of Ecosystem Services". Washington: IslandPress, 2007.

Tallis, Heather, and Stephen Polasky. "Mapping and Valuing Ecosystem Services as an Approach for Conservation and Natural-Resource Management." *Annals of the New York Academy of Sciences*, 2009: 265-283.

Taylor, Laura O. "The Hedonic Method." In *A Primer on Nonmarket Valuation*, by Patricia A. Champ, Kevin J. Boyle and Thomas C. Brown, 331-394. Norwell: Kluwer Academic Publishers, 2003.

ten Brink P., Badura T., Bassi S., Daly, E., Dickie, I., Ding H., Gantioler S., Gerdes, H., Hart, K., Kettunen M., Lago, M., Lang, S., Markandya A., Mazza, L., Nunes P.A.L.D., Pieterse, M., Rayment M., Tinch R., (2011). *Estimating the Overall Economic Value of the Benefits provided by the Natura 2000 Network.* Final Synthesis Report to the European Commission, DG Environment on Contract ENV.B.2/SER/2008/0038. Institute for European Environmental Policy / GHK / Ecologic, Brussels 2011

Train, Kenneth E. "Discrete Choice Methods with Simulation 2nd Edition." Cambridge: Cambridge University Press, 2009.

CHAPTER TWO: Random Utility and Choice Modeling

Introduction

How humans make choices is a subject of endless speculation. It is a topic that is addressed in art, religion, literature, philosophy, natural and social sciences. Decision makers and those seeking advice also has an equally long history, the Oracle of Delphi was a place where those seeking information or wisdom would go to find answers. Today policy makers often turn to focus groups or surveys to gauge how a policy or program will be received. In the US Presidential election of 2012, the established pundits predicted a close electoral race, while the statistical analysis of New York Times blogger Nate Silver predicted a strong lead by Obama and correctly predicted 50 of 50 states. After the election the statistical analysis carried the day while the pundits were left wondering what had happened.

The resource manager also needs to predict what attributes of a resource or program is desired or preferred. Environmental economists have developed very robust techniques for making predictions on what people prefer through the use of what is termed random utility modeling. Originally, it was used by marketers to estimate demand from large aggregate data sets, while recent developments have examined individual preference heterogeneity. Since then statistical models have a long history and are useful tools for the resource manager seeking information on how incremental changes will influence changes in consumer welfare. In marketing, firms need to make predictions and determine what attributes a population desires. The literature on how to gauge population preference is extensive (Louviere, Hensher and Swait 2000), (Holmes and Adamowicz 2003). In modeling choices random utility theory has also found wide usage in transportation demand studies. Also, random utility is very helpful in ranking preferences or creating an ordinal ranking (Holmes and Adamowicz 2003). Random Utility Modeling is even flexible enough to be used to rank recycling preferences on a college campus (Gebben 2008). The framework from which the random utility modeling utilizes is the focus of this chapter.

Predicting Choices

Economists often theorize about how the individuals (consumers) make choices. These behavioral models should be internally consistent (Freeman 2003). A basic assumption about individual choice is that of utility maximization – that is, subject to constraints, consumers choose what they prefer. The choices made by the consumer are goods or programs consisting of various attributes that combine to contribute to consumers' utility (Louviere et al. 2000). Attribute based methods of analysis seek to further understand consumer choices in terms of the attributes (or characteristics) of the alternatives that are chosen. It further requires that the model used for predictive purposes have some underlying theory. As has been pointed out in the previous chapter, measurement without theory is at best unsuitable for explanation, and at worst meaningless (Louviere, Hensher and Swait 2000), (Koopmans 1947). The theoretical framework for this study is based on the Random Utility Theory (Adomowicz and Boxall 2001). I will first lay out the basic underlying principles of random utility and why it is used. Then, I will move into how it is defined functionally. Finally, I will close with a look at how revealed and stated preference data can both be utilized through random utility modeling.

Random Utility Theory and Modeling

The goal in the application of traditional random utility models (RUM) in nonmarket valuation is to model consumer choice, whether between, recreation sites (Parsons, Massey and Tomasi 2000) or whether they would pay for a given public good. Modern day travel cost models (TCM) and contingent valuation models (CVM) assume a basic RUM model. The early formulation of RUM as a behavioral model followed economists' theory of consumer behavior.

Therefore, preferences are heterogeneous across individuals and unknown to the analyst. The researcher can parameterize those preferences and the distribution of the random factors, thus arriving at a model that is reasonably able to describe a population's preferences (McFadden 2000).

Given that the goal is to make predicative statements about what the consumer will choose the random utility models is derived as follows. A consumer makes a choice from the given alternatives. RUT assumes the respondent makes the preferred choice (Holmes and Adamowicz 2003), and this choice mimics what goes on in actual markets. The underlying idea of random utility theory is pretty basic. That the consumer gets utility from some set of attributes that a private good or public site provides. These recreations sites or goods are mutually exclusive of each other. From these alternative choices, the consumer is chooses the best.

The resources are characterized by attributes and a price. As those attributes are added or subtracted, or if the quality of those attributes is altered, then welfare estimates are also changed. The consumer is then offered alternative recreation sites from which to choose (Hilger and Hanemann 2006).

At its most basic a random utility model, seeking to make a decision based on a program or policy attributes has a few basic steps to follow (Holmes and Adamowicz 2003):

- Characterize the problem
- Identify and describe the attributes
- Develop an experimental design
    - This can be either through a state or revealed preference
- Develop the survey tool

- Collect the data
- Estimate model
- Interpret the results for policy analysis or benefit cost analysis

The above steps are useful for nonmarket valuation techniques since they force the researcher at the beginning to identify the problem or question being asked. This then directs the following steps in determining what the relevant attributes. Once the attributes are determined, it then becomes possible to determine what design is most practical. That is, in those situations where the non-use or passive value is needed to be estimated, a stated preference design might be needed. However, in those situations where a direct use value is needed, a revealed preference design would be most applicable. This is where the theory is needed for accurate measurement. Without a theory leading the design, it is very difficult for the researcher to get results that are appropriate and meaningful. Following the determination of the survey design that is best, the researcher must collect the data, and ensure data quality. This is difficult and requires that the survey be field tested and during collection during implementation spot checked as it goes along. For a more full discussion of various survey collection techniques the interested reader is encouraged to read Salant and Dillman's How To Conduct Your Own Survey (Salant and Dillman 1994). Once the data is collected, estimation of the model is required, and ideally the estimation tool would have been considered prior to the survey's completion. That however does not preclude for interesting possibilities to arise after the survey is completed. The final step of interpreting the results in a format accessible to policy makers or resource managers is critical.

The platform of random utility theory enables the researcher to utilize the above steps, and where appropriate to merge stated and revealed preference data sets. The model has gained wide usage in environmental settings because of its flexibility and robustness. Random utility theory and models are often used in nonmarket valuation, and discrete choice models. As such they have been used in various forms of travel decisions models, for example, does the person take the car, bus or train? Each choice is a discrete option. Other common examples of where random utility models are used is fishing choices, both recreation and commercial fishing site choices. These studies examine what the particular site a characteristic of fishing ground has that helps the respondent choose that area over another (Train 2009).

The models have a wide presence in nonmarket valuation. Choice experiments are also common examples of random utility models. The travel cost model, described in greater detail in the previous chapter, can also make good use of the model when the number of alternative sites is large enough. And contingent valuation also employs random utility models.

Discrete choice models are generally derived under the assumption of utility maximizing behavior (Train 2009). However, it is worth noting that random utility models can also represent decision making made under criteria other than utility maximization. The model can also be seen as a describing the relationship of the explanatory variables to the choice probability, without making statement on how the choice was made (Train 2009).

Under the assumptions of Random Utility Theory (RUT), the general model of choice starts with the indirect utility, U, a consumer is able to obtain which is a function of income I, a vector of attribute levels, $\mathbf{x}$, and individual tastes, T.

$$U = u(\mathrm{I}, \mathbf{x}, T) \qquad (1)$$

Respondents are asked to choose between two bundles of attributes, each with a price. Therefore, the indirect utility for an amount p paid for an attribute would be:

$$U_1 = u(I - p_1, x_1, T_1). \qquad (2)$$

where, utility U_1. is a function of x_1 a vector of attributes provided to the consumer at prices p_1. The respondents are then presented with two program choices, and were asked to choose the program that they preferred. Therefore, the indirect utility associated with the second option can be expressed as:

$$U_2 = u(I - p_2, x_2, T_2). \qquad (3)$$

Therefore, we can assume that an individual would select (be willing to pay p_1) the bundle of goods x_1 if the utility from x_1 was greater than the utility from x_2 at p_2

(i.e., $U_1 > U_2$). So the probability that a respondent would say *Yes* when asked if x_1 is preferred to x_2, may be expressed as:

$$\begin{aligned} \Pr(yes) &= \mathbf{P}r[u_1(x_1, I - p_1, T_1) \geq u_2(x_2, I, -p_2, T_2)] \\ \Pr(yes) &= \Pr[\Delta u > 0]. \end{aligned} \qquad (4)$$

Using the basic approach embodied in (4), we can parameterize the utility functions to quantify preferences. Let utility be given by the sum of observable and unobservable components:

$$u_{ij} = v_{ij} + e_{ij}. \qquad (5)$$

Where v_{ij} represents the observable portion of utility and e_{ij} represents the unobservable portion of utility. With that, we can then modify (4) so that:

$$\Pr(yes) = \Pr[v_{ij} + e_{ij} > v_{ik} + e_{ik}] \qquad (6)$$

The observable portion of utility is made up of the following characteristics:

$$v_{ij} = \alpha x_j + \beta(I_i - p_j) \;\;, \qquad (7)$$

where x_j is a vector attributes associated with alternative *j*, α is a vector of estimable parameters, β is an estimable parameter, I_i is income of respondent *i*, and p_j is the price of alternative *j*.

Substituting (7) into (6) we can write the probability of a *yes* answer to a preference questions regarding *j* versus *k* (a respondent prefers program *j* to program *k*) as follows:

$$\begin{aligned}
\Pr(yes) &= \Pr[\alpha x_j + \beta(I_i - p_j) + e_{ij} > \alpha x_k + \beta(I_i - p_k) + e_{ik}] \\
\Pr(yes) &= \Pr[(\alpha x_j - \alpha x_k) - (\beta p_j - \beta p_k) > e_{ik} - e_{ij}] \qquad (8) \\
\Pr(yes) &= \Pr[\alpha(\Delta x\ ) - \beta(\Delta p\ ) > e_{ik} - e_{ij}].
\end{aligned}$$

It is typically assumed that the marginal utility of income does not change from one choice to another. Since the utility of income does not change from one state to another, we can then drop the ΔI_i terms as in the second line of (8). We can next make the assumption that the difference in the error terms is normally distributed allowing for a model that gives the probability of an individual choosing alternative *j* over alternative *k* :

$$\Pr(yes) = \Pr[\alpha(\Delta x) - \beta(\Delta p) > e_i] = \Phi[\alpha(\Delta x\ ) - \beta(\Delta p)], \qquad (9)$$

where Φ is the cumulative distribution function (cdf) of the standard normal distribution, assuming that $\sigma = 1$. Equation (9) can be estimated using the maximum likelihood estimation procedure for a probit model. Due to the discreet nature of the choices, σ cannot be identified in a probit, so one can take it as if $\sigma = 1$ or simply recognize that all parameters are identified up to the unknown variance term. As such, model comparisons that involve parameter ratios are fully identified since the unknown variance term will cancel.

Random utility can also easily be modeled as a logit if we assume the error term is logistically distributed. If we return to equation (5), the logit model is obtained if we assume that e_{ij} is independently, identically distributed (iid) extreme value. This is sometimes called the

Type I extreme value or Gumbel distribution. The density for each unobserved component of utility is then:

$$f(\varepsilon_{nj}) = e^{-e_{nj}} e^{-e^{-\varepsilon_{nj}}} \quad (10).$$

Where the cumulative distribution is

$$F(\varepsilon_{nj}) = e^{-e^{-\varepsilon_{nj}}} \quad \text{(11) (Train 2009).}$$

The variance in this situation is defined as $\pi^2/6$. By assuming this variance it implicitly normalizes to a scale utility, for a full discussion of why this happens the interested reader is directed to Train's (2009) discussion of the matter in Discrete Choice Methods with Simulation . If we assume that the error terms between alternatives e_{ij} and e_{ik} are iid, the $e_{ijk}^{*} = e_{ij} - e_{ik}$. Which follows the logistic distribution:

$$F(\varepsilon_{ijk}^{*}) = \frac{e^{\varepsilon_{ijk}^{*}}}{1 + e^{\varepsilon_{ijk}^{*}}} \quad (12).$$

Equation (12) is sometimes used to describe binary logit models, this model can then be extended to multinomial logit models.

Taking this reduction of the realm of possible site choices down to a manageable set, we can estimate the RUM as a multinomial logit. If we assume that the user can choose from a universe of *C* possible site locations, we can then say that the consumer will choose a site based on the following utility:

$$U_i = X_i\beta + \varepsilon_i \quad (13)$$

where X_i is a vector of site characteristics at site i $(i = 1, \ldots, C)$, including travel time and cost, site amenities and so forth. β is an unknown parameter vector, and ε_i are independent and identically distributed type 1 random variables. From this we can derive the basic logit probability for visiting a given site k in the choice set

$$\Pr(visit)_k = \frac{\exp(X_k,\beta)}{\sum_{i=1}^{C}\exp(X_k,\beta)} \quad (14) \text{ (Parsons, Massey and Tomasi 2000).}$$

The numerator of (14) provides the exponential of the visited site utility, while the denominator is the sum of the exponential over all site utilities in the universe of sites. With the observational data on people having visited a given site in the set of C possibilities and characteristics X_i, the logit probabilities are used in a standard likelihood function to estimate the parameters β.

Any given individual will then have the expected utility of visiting a site on a given choice occasion in the logit model as:

$$ln \sum_{i=1}^{C} \exp(X_i, \beta) \quad (15) \text{ (Parsons, Massey and Tomasi 2000).}$$

Equation (15) is the natural log of the denominator for the logit probability, summed over the *C* sites in the choice set.

With these pieces one can then begin to model how an individual chooses a discreet option from a finite number of possible options. One assumes that the individual will choose the option that has the most utility relative to the other options available. While it is possible that a respondent could be asked to rank the alternatives, this study chose to look at only the first preference. The general form of this discreet choice would be:

$$U_{qi} = \beta_q X_{qi} + e_{qi} \quad (16) \text{ (Hensher and Greene 2003).}^{2}$$

Where the subscript q is the individual, *i* is each alternative choice. The *X* is then a vector or non-random explanatory variables while β and *e* are random non-observed variables.

[2] Hensher and Greene include a third subscript *t* representing the choice situation, this would be applicable if data used a choice experiment matrix, this study is using a DC CVM follow-up question instead.

The RUM model making use of the multinomial logit allows us to let person *n* gain utility from alternative *j* in situation *t* and building from that the functional form represented as $U_{njt} = \beta' x_{njt} + e_{ntj}$ where x_{ntj} does not contain any alternative specific constraints, and e_{ntj} represents constraints not observed by the researcher. The factors have a mean and a distribution around each mean. The alternative specific constant can be labeled c_j, and for a standard logit model, this distribution around the mean is extreme value with variance $\lambda^2\pi^2/6$ (Train 2009).

With the RUM model described, we can now turn to utilizing it in practice. The following chapters will discuss applications of the model.

WORKS CITED

Freeman, A. Myrick III. "The Measurement of Environmental and Resource Values*" Second Edition.* Washington DC: Resources for the Future, 2003.

Gebben, David J. "An Attribute Based of Recycling Preferences At Michigan State University." *Masters Thesis.* East Lansing : Michigan State University, 2008.

Hensher, David A., and William H. Greene. "The Mixed Logit Model: The State of the Practice and Warnings for the Unwary." *Transportation*, 2003: 133-176.

Holmes, Thomas P., and Wiktor L. Adamowicz. "Attribute-Based Methods." In *A Primer on Nonmarket Valuation*, by Patricia A. Champ, Kevin J. Boyle and Thomas C. Brown, 171-220. Norwell: Kluwer Academic Publishers, 2003.

Koopmans, Tjalling. "Measurment Without Theory." *The Review of Economic Statistics*, 1947: 161-172.

Louviere, Jordan J., David A. Hensher, and Joffre D. Swait. *State Choice Methods Analysis and Application.* Cambridge: Cambridge University Press, 2000.

Parsons, George R., Matthew D. Massey, and Ted Tomasi. "Familiar and Favorite Sites in a Random Utility Model of Beach Recreation." *Marine Resource Economics*, 2000: 299-315.

Salant, Priscilla, and Don A. Dillman. "How to Conduct Your Own Survey*."* New York: Wiley, 1994.

Train, Kenneth E. "Discrete Choice Methods with Simulation 2nd Edition." Cambridge: Cambridge University Press, 2009.

CHAPTER THREE: Augmenting Revealed Preference MNL Models with Stated Preference Contingent Valuation Dichotomous Choice Data

Introduction

The goal of this paper is to explore whether or not multinominal logit (MNL) Revealed Preference (RP) data combined with a simple Dichotomous Choice (DC) Contingent Valuation Model (CVM) Stated Preference (SP) can provide significant gains in efficiency at a low marginal cost. Taking an RP data set, and utilizing the MNL regression, it is possible to estimate how a respondent would alter their choice set based on how they decided where to go.

The proposed approach has the advantage that this method is simpler for the respondent than conducting a full choice experiment. This method also has the advantage of being less hypothetical than varying all the attributes simultaneously. By focusing on just a few potential site choices, the question becomes less hypothetical and also is consistent with the NOAA Blue Ribbon (1994) recommendations.

Background

As discussed in previous chapters, the need to effectively measure and calibrate the preferences of a population requires the resource manager to weigh various interests against each other. At times those interests are in competition with each other at best or at worst are in outright conflict. The use of benefit cost analysis provides a lens to compare what otherwise might be apples to oranges. To value the competing interests is important, and in the process of creating a useful benefit cost analysis, confidence intervals around a willingness-to-pay estimate is often employed. Now, for the resource manager looking at the upper and lower bounds of the confidence interval, a given project would ideally pass at either bound. This is true in a number of cases, including situations of uncertainty, and there are times where the parameters of a given project cannot be known due to technical, physical or economic constraints. The state of the art may not be sufficient to provide that certainty. In such situations providing the upper and lower bounds for a benefit cost analysis would is very useful (Freeman 2003). So, for the researcher the goal is to provide a measurement that has as narrow an interval as possible.

There are various methods to accomplish that goal - one obvious method in statistical analysis is increasing the number in the sample. However, given that even in survey research the interested parties operate under constraints, increasing the total number of respondents is costly. Further, increasing the total number might reduce the quality of the data collected, either by gaining less information per person or introducing the potential of respondent self-selection. Gaining a sample that is representative of the general population takes money and time, and if there are methods to gain the same precision in estimates at a lower cost those options need to be considered.

An alternative method is to increase the information collected from each respondent. By increasing the amount of information provided, it is possible to increase the efficiency of the estimators thereby reducing the length of the confidence intervals. Developing valuation techniques which can accomplish that provides the resource manager another tool for analysis.

More narrowly environmental economics has developed valuation techniques for those resources that do not get valued by normal market mechanisms. These non-market valuation techniques have become quite common, and the U.S. Government and other policy bodies often require valuation information as part of the planning process for new programs or projects.

As discussed in previous chapters, the valuation of non-market resources often falls into one of two categories: Stated Preference or Revealed Preference. Both categories provide information about the underlying preferences for some type of good or resource. Both revealed and stated preferences have qualities that allow them to function well on their own, and many researchers utilize them as distinct models. Often the two types of information are utilized separately from each other, but each has the potential to improve on the weakness of the other. The possibility of combining the two methods into a single model has appeal, especially when the data for both was constructed through a survey response format allowing for a panel data set. The combining of both allows for the stated preference data to add variation to the revealed preference data, while the revealed data can ground the stated data in reality (Train 2009). As the amount of information increases in the model, the efficiency increases as well. The combining of stated and revealed preferences has the potential to lower the cost of gaining the needed information that the researcher needs to provide the resource manager. The combining of these two sources of information can then help narrow the gap on the confidence interval thereby helping the resource manager to have a better idea of at what point the project is or is not feasible.

Various others have explored the combining of stated and revealed preferences (Hensher, Louviere and Swait 1999), (Louviere, Hensher and Swait 2000), (Cameron 1992). Often, combining the two is taking a specific type of data, either stated or revealed, and building a new model with only a slight change from one of the primary source. The point is not to create new models, rather the goal is to develop a “model and estimation procedure that seem appropriate for that particular situation, drawing from, yet adapting, the standard set of models and tools” (Train 2009). Building upon the foundations of stated and revealed preferences, and thereby increasing the effectiveness by addressing a specific issue.

The specific question here is how might a simple dichotomous choice contingent valuation question inform the valuation of a river site? As was discussed in the previous chapter, a travel cost model is a revealed preference model that can be estimated through a RUM model (Parsons, The Travel Cost Model 2003), (Train 2009). Often the resource manager utilizing the travel cost model is seeking information about the economic value a recreation area provides. Suppose that the resource manager has more than one location he or she needs to manage, but the cost of getting to any one site is not significantly different from another site. For example, the cost of getting into a park is fixed, and once in the park all areas regardless of location cost the same. The resource manager might then want to find out if there are specific parts of the park that visitors would pay more for or would like a change in some quality. In this case, the travel cost model may only provide a lower-bound on the willingness-to-pay estimate. If the resource manager wants to know if there is a higher economic value, it may be difficult to estimate with only a revealed preference study. Since the travel cost model can only provide information about the current state of the resource and consumer surplus as estimated as a function of the current amount paid by people to get there, the model is limited in its flexibility.

In those situations where the resource manager needs more specific information about the upper bounds people would pay or what resource alternatives people would prefer, the travel cost model is not be able to provide that needed information for a meaningful tradeoff analysis or benefit cost analysis. The needed variation is not present in the data, but a stated preference survey would be able to get at those details. By including a stated preference section in a travel cost survey, the researcher can explore if there are specific parts that visitors might be willing to pay more for or to consider an alternative list of qualities. Depending on the format the stated preference information might also narrow confidence interval length, or provide a more precise measure of the willingness-to-pay.

Combining SP and RP

The combining of Stated Preference and Revealed Preference data has a history dating back to when Cameron first (1992) suggested it. By increasing the information on contingent valuation data, the RP data provides data points that are less hypothetical than the SP data. The advantage of combining the SP and RP data is that it allows the RP to “discipline the SP data” as Cameron (1992) phrased it. The combination of the two data sources allows the researcher to impose consistency between the two response formats for estimating willingness-to-pay (WTP) (Loomis 1997). This often requires multiple responses from the same individual, necessitating some form of panel estimation technique (Loomis 1997). Further, depending on the survey format used, there is a low marginal cost to the additional information. However, caution must be used since the researcher must consider the total cost as well as the marginal cost.

The travel cost model is useful for estimating the value of recreational uses of a site. It mimics the downward sloping demand of a market demand curve and can provide useful insights into how a recreation activity would be affected by an attribute change or a policy change (Parsons 2003). To ask a single question regarding a respondent in a referendum style of their preferences is often asked in the typical contingent valuation study. Occasionally, there is need for a follow-up question which is dependent on the respondents’ previous answer (McFadden 1994). Asking a dichotomous choice CVM question to follow up survey information on a TCM survey is not only a reasonable method to gain information but also offers a method to combine the two pieces of information. Since in both the CVM and TCM the price variable would be affected, it is a reasonable merging of information.

Merging RP and SP through RUM

Utilizing the RUM model to combine revealed and stated preference data has the intuitive appeal of utilizing actual behavior from the former and merging that with explicit policy-relevant scenarios not otherwise available except through the latter. It also has the desired feature of restraining unobservable factors of both models to the error terms, suggesting that the error terms could be related (Gonzalez, Loomis and Gonzalez-Caban 2008).

It is possible to normalize the utility to some standard scale Train (2009) goes through this in great detail. Suffice it to say here that the unobserved portions of the error terms in the random utility model are set to $\lambda^s$and to $\lambda^r$for stated and revealed respectively. To set the overall scale utility we set either λ^sor λ^r1 and then make the other equal to the ratio of the original scale parameters (Train 2009). The model can then be estimated from both the revealed and stated data, with both sets of data able to be estimated through "stacking" the information in a logit type of estimation.

The specific estimation utilized here is the multinomial logit (MNL) model. By merging the revealed preference data in a MNL and a stated preference dichotomous choice CVM model creating a hybrid model. The MNL - Site choice model of understanding how people make their choice based on the TCM model is a very commonly used RP method. The single dichotomous choice question is asked in a standard CVM referendum type question as follow-up. This allowed us to change only the Travel Cost (TC), though in other settings it could be framed to lower a single site attribute.

By merging those two models, it enables the researcher to utilize two very strong models and determine if the sum are then greater than the parts. The resource manager would need to know this in cases where an increase in the precision of the willingness-to-pay is needed, or where the cost of getting the same precision from a larger survey would be more than just getting respondents to answer one more question. Getting a respondent to answer a survey is the most expensive part of a survey response; the marginal cost of getting the respondent to answer one more question is very low, especially if it is a simple dichotomous choice CVM question.

This study is unique in that it combines the MNL with the single bounded dichotomous choice CVM response information. The revealed preference MNL allowed for a choice of ten sites at a given travel cost (TC). It is then possible to augment that with a dichotomous choice CVM question given this new piece of information. Other studies have combined TCM and CVM models (Loomis 1997). Adamowicz et. al. (1994) suggested using the MNL model with multi-attribute Choice Experiments. Utilizing the MNL with a dichotomous CVM is a unique combination of the two, and the utilization of in person data is also different from most studies that have looked at these types of hybrid models. While choice experiment does has the advantages of a large amount of information can be fed into the model, the cost to the respondent is higher since it can also be more cognitively taxing. If the resource manger is only seeking information on one quality change or price change, then a full choice experiment would be unlikely to be the best stated preference format.

Data

The data used in this paper was collected from the El Yunque National Forest in northeastern Puerto Rico. This information was collected during the summers of 2004 and 2005 as part of a comprehensive study examining the impact of site characteristics in and around forest streams. There were over 700 surveys collected in person with a total of 450 that were usable. A large number of those excluded were those persons who responded that the visit to the river was not the sole purpose of the trip. In order to get an accurate valuation, there cannot be multiple reasons for the trip. It is difficult to separate the portion of a trip spent at the site from the rest of the trip. Inclusion of visits where the river was not the sole reason for the trip would violate the assumption that all out of pocket expenses are acting as a proxy price for only the trip and nothing else (Loomis, Yorizane and Larson 2000), (Parsons 2003). The wording of the question is in Figure 3.1.

As you know the price of gasoline often goes up. Taking into consideration that there are other rivers as well as beaches nearby where you could go visit, if the cost of this visit to this river was $____ more than what you have already spent, would you still have come today?

____ Yes ____ No

Figure 3.1. Wording of Follow-Up Question

This allowed us to then vary the cost across the various sites from $5 to $200. The data was collected on site in the El Yunque National Forrest. The site characteristics are summarized in Table 3.1.

Table 3.1. Site Characteristics of Study Area

River	Espiritu Santo	Fajardo	Mameyes
Number of Observations	**164**	**235**	**593**
Natural and Built Site Variables			
Presence of Waterfalls	**3 of 5 sites**	**2 of 5 sites**	**2 of 7 sites**
Presence of Scenic Views	**3 of 5 sites**	**3 of 5 sites**	**5 of 7 sites**
Presence of Formal Trails	**1 of 5 sites**	**0 of 5 sites**	**2 of 7 sites**

The approach for combining the revealed preference MNL and stated preference CVM data was to allow the choice set from among 10 sites at travel cost. Since Puerto Rico is an island, it is reasonable to assume that all ten sites are within the respondents' choice set. We then augment that with SP DC CVM in the following ways. A **Yes** response to the CVM means the person would continue to visit or select their current site at the travel cost + bid amount. We create another set of site choice observations for these responses, in which the travel cost + bid is added to the revealed preference Travel Cost model for the current selected site, thus building a new travel cost variable with greater willingness-to-pay variation for each of our ten sites. Since the dichotomous choice CVM questions were asked of visitors to all the sites, we gain additional variation in TC for all the sites in the choice set.

A **No** response would not pay the bid amount and does not provide information regarding what other site (if any) would be chosen, because these responses do not create another set of site choice options. Asking if they would visit an alternative site with a lower travel cost would be an area for future improvement in the survey collection.

Results

Our primary concern was to look at the improvement in the cost variable. When comparing the TC coefficients, we did see an improvement in the form of a reduction of variance and a higher level of significance. When looking at the gains in efficiency based just on variances, there was a substantial amount from just a simple dichotomous choice CVM question.

The variance was reduced from 0.0127 for just the RP cost coefficient, to 0.0023 for both the RP and SP cost coefficient. The length of the confidence interval of 95% for just the RP data is from -0.0224 to -0.248, a length of 0.0023. While the 95% confidence interval for the combined data is from -.0096 to -0.0100, a length of just 0.0004. These are summarized in Table 3.2.

Table 3.2 Results of Models

Model	Travel Cost Coefficient	Std. Error	T-Statistic	Weighted Average of Consumer Surplus
RP only	-0.0236	**0.0127**	**-1.86 ***	**-3.33**
RP & SP (DC-CVM)	-0.01	**0.0023**	**-3.41 *****	**-10.98**

Conclusions and Implications

We see the gains provided by a simple and low cost follow-up question to travel cost values. This additional piece of information helps to better calculate the surplus values of a recreation area and provides a narrower bounds to aid policy makers or resource managers. Given that the largest cost in a survey data is often getting the respondent to participate in the survey, gaining one more piece of information though a dichotomous choice question would be a worthwhile use of the researcher's time and energy. While a full choice experiment has the appeal of just that much more variation of the attributes and responses from a respondent, the marginal cost might be prohibitively high for the respondent. Thus, avoiding the perils of diminishing marginal returns in the survey design needs to be observed.

The other advantage of this method is that it provides a less hypothetical scenario since only a single attribute is being altered which in this case is a higher cost. With just a single alteration in the scenario, the statistical efficiency went from 10% to 1% significance level. The gains in statistical efficiency likely outweigh the costs of an additional survey question. These gains then provide a more full insight into tradeoff analysis, allowing the resource manager to better allocate the scare resources like time and money where they can be best allocated.

In addition, with the knowledge of what the upper bound is on a resource, it can help expand the discussion of what the economic value is of a resource since it no longer is limited to just the lower bound that would have been provided from just a revealed preference travel cost study. Expanding the discussion of value to look at the losses of that might be incurred by the population by an entrance fee increase can demonstrate who in policy changes who would be most affected.

Policy makers might find research results that take the best attributes of RP data combined with SP data to be of more use than their separate parts. The sum of the two working together provides a more efficient assist for estimation of WTP values. This could be of particular value when conducting a benefit transfer study. If there are other areas or studies that have relevant data, an augmented estimate would likely provide greater accuracy.

Obtaining more information from the “No” responses would also be an area for improvement. This could be as simple as asking either what alternative site might they visit or what other activity would they choose, or would they simply stay home and opt out of the recreation.

The times when such a revealed and stated preference study would not be encouraged are those situations where the resource manger is not interested in the change of a quality or price change. If the manager is seeking information only about the lower bound valuation estimate from direct use, then a stated preference question from people not enjoying the site would be of little value. Further, a purely revealed preference study into the existence or indirect value of a resource is infeasible. A study exploring the change in a resource attribute quality is much more suited to a stated preference format. The merging of the two types of data in either of the above would play off the weaknesses of both formats, rather than trying to build on the strengths of both.

In this instance the combining of both types of data provided a more precise estimate of the value that would be lost in case of an increase in price to visit a recreation site along a river increased. The resource manager looking to increase the efficiency of the estimate can with a simple dichotomous choice question gain greater insights into the estimates of consumer surplus loss from an increase in price. At much lower cost than would have been required for a similar increase in precision by increasing the sample size.

WORKS CITED

Cameron, Trudy A. "Combining Contingent Valuation and Travel Cost Data for the Valuation of Nonmarket Goods." *Land Economics,* 1992: 302-317.

Cameron, Trudy A., and John Quiggin. "Estimation using contingent valuation data from a "Dichotomous Choice with follow-up" questionaire." *Journal of Environmental Economics and Management*, 1994: 218-234.

Freeman, A. Myrick III. "The Measurement of Environmental and Resource Values Second Edition." Washington DC: Resources for the Future, 2003.

Hanemann, W. Michael. "Valuing the Environment Through Contingent Valuation." *The Journal of Economic Perspectives,* 1994: 19-43.

Hensher, David, Jordan Louviere, and Joffre Swait. "Combining Sources of Preference Data." *Journal of Ecnometrics,* 1999: 197-221.

Loomis, John B. "Panel Estimators to Combine Revealed and Stated Preference Dichtomous Choice Data." *Journal of Agricultural and Resource Economics*, 1997: 233-245.

Loomis, John, Shizuka Yorizane, and Douglas Larson. "Testing Significance of Multi-Destination and Multi-Purpose Trip Effects in a Travel Cost Method Demand Model for Whale Watching Trips ." *Agricultural and Resource Economics Review,* 2000: 183-191.

Louviere, Jordan J., David A. Hensher, and Joffre D. Swait. *State Choice Methods Analysis and Application.* Cambridge: Cambridge University Press, 2000.

McFadden, Daniel. "Contingent Valuation and Social Choice." *American Journal Agricultural Econoimcs,* 1994: 689-708.

McFadden, Daniel. "Disaggregate Behavioral Travel Demand's RUM Side." Resource Paper, Berkeley: Department of Economics, University of California, Berkeley, 2000.

Parsons, George R. "The Travel Cost Model." In *A Primer on Nonmarket Valuation,* by Patricia A. Champ, Kevin J. Boyle and Thomas C. Brown, 269-330. Norwell: Kluwer Academic Press, 2003.

Train, Kenneth E. "Discrete Choice Methods with Simulation 2nd Edition". Cambridge: Cambridge University Press, 2009.

CHAPTER FOUR: Is the Bias-Efficiency Estimation Tradeoffs in Double Bounded Choice Less Pronounced Using Visitor Survey Data?

Introduction

Part of the manager's predicament when investigating different management options is to consider alternative states is often quite impractical through revealed preferences. The estimation of preferences for consideration of changes in quality that is needed for decision making is often simple not present in a revealed preference format. One method for eliciting preferences for different quality levels is through a stated preference (SP) survey. The advantage is that it allows the researcher to pose a question framed in a manner to evaluate the world as it might be. SP methods involve eliciting responses from people, often in the form of ordinal rankings or choice sets for a given list of options. An SP model might ask respondents how much they would pay to restore a watershed area that would result in a clearer lake.

A stated preference approach allows a researcher to value attributes as well as situational changes. As such it allows the researcher to examine a range of ecological impacts (Louviere, Hensher and Swait 2000). One specific type of stated preference method is the continent valuation method (CVM). Modern CVM studies utilize a referendum style of questioning to minimize strategic behavior since the format appears to induce individuals to reveal their true preferences or values (Louviere, Hensher and Swait 2000). This property of the CVM allows for a tradeoff analysis so that the manager can weigh alternative plans and options that might not be feasible using only a revealed preference format.

Economists began using contingent valuation when Robert Davis collected surveys to gauge the value of big game hunting in Maine (Boyle 2003). According to Boyle, this marked the beginning of surveys used for nonmarket valuation (Boyle 2003). The recommended method of CVM is a single-bounded (SB) discrete choice (DC) offered in the form of a referendum response mechanism (Boxall, et al. 1996). The credibility of that format was enhanced by the NOAA Blue Ribbon Panel (NOAA 1993) recomendataions.

The referendum style of CVM has much in its favor. Not least is that it reflects what a consumer would often face in the real world. An example could be a referendum on a ballot initiative or other take-it-or-leave-it market situations. The method requires a binary choice, and has the advantage that since it often only requires a respondent to consider only one good or one change in the attributes (Brown 2003), it is not very cognitively taxing. The willingness-to-pay (WTP) value found in a CVM study is often quite similar to a travel cost study, and can provide information not otherwise possible.

The single bounded model with its incentive compatible framework has at least one drawback; it is somewhat inefficient in gathering information. That is, since the respondent has only one price that he or she faces in the format, the researcher needs to ask a larger sample size to get a WTP estimate with a narrow confidence interval. The older method of asking an open ended “How much would you pay question?” does not have the incentive compatible nature of the binary take-it-or-leave-leave-it referendum format. To address this issue, one option is to extend the potential of the CVM format through what is known as the follow-up question or double bounded CVM study.

Research has shown that the double-bounded and interval models are statistically more efficient than single payment dichotomous choice questions (Albernini 1995). The role of the dichotomous choice double-bounded contingent valuation method in recreation valuation has been an important estimation tool since the publication of “Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation” (Hanemann, Loomis and Kanninen 1991).

A Google Scholar search on the paper resulted in 893[3] citations of the paper. Considering that the median citation of articles from the AJAE in 1991 was three (Hilmer and Lusk 2009), it has clearly had a large impact in the field of resource valuations. However, past studies have also shown the potential of starting point bias between the first bid amount and second bid amount in the follow-up WTP questions can lead to respondent uncertainty about the actual bid amounts they are paying (Herriges and Shogren 1996). The respondent might question whether or not the quality of the good is still the same at the new bid amount.

Exploring the advantages of the double bounded format is the interest of this chapter. How can the researcher provide an increased efficiency to the CVM format, without increasing the cost of the survey exorbitantly? If the same level of precision can be provided through a double bounded method, at a lower cost than it would be a useful means to provide a resource manager information about what a potential quality change might provide.

Background

The simple take-it-or-leave-it format of the single-bounded dichotomous choice can only provide a single bound on the respondents' willingness-to-pay. This therefore requires a larger amount of respondents to get a precise estimate (Herriges and Shogren 1996). The extension of asking a follow up question to the initial bid amount was first proposed by Hanemann (1985). These double-bounded questions provide more statistical efficiency to the welfare estimates derived from CVM studies (Roach, Boyle and Welsh 2002). However, these gains in efficiency come at a cost. The double-bounded response is more taxing on the respondent than the single-bounded dichotomous choice question. To achieve the same level of precision that the double

[3] May 5, 2012

bounded (DB) estimation offers usually requires the use of a larger sample size from a single-bounded (SB) question (Hanemann, Loomis and Kanninen 1991).

The trade-offs between the statistical efficiency gains of the double-bounded modeling method against the biasing effects of the second question must be considered. As Kanninen notes, this boils down to a bias versus efficiency debate when choosing between uses of the double-bounded and single-bounded dichotomous WTP question format. From a statistical perspective, the concern relates more towards the magnitude of variance for the estimated WTP than a point estimate, since the length of the confidence interval indicates the degree of uncertainty for the estimated welfare benefit (Kanninen 1995).

When trying to increase efficiency of a model, one area of concern with the DB model is starting point bias. Starting point bias refers to the iterative nature of the double-bounded dichotomous choice format (Herriges and Shogren 1996). Ideally the first bid amount acts only as a useful tool to begin the estimation process. The second bid continues the narrowing down process, determining the respondents' WTP. In other words, the start point does not influence the final WTP of the respondent or act as an "argument in the respondents' utility functions" (Boyle, Bishop and Welsh 1985). The final bid should signal to the researcher the Hicksian compensating or equivalent surplus of the good or service being valued (Boyle, Bishop and Welsh 1985). However, the starting bid can act as a signal as to what "should" be paid (Herriges and Shogren 1996). That trade-off between bias and efficiency requires consideration.

Previous investigations have looked mainly at either household data or simulated data sets. Cooper and Loomis (1992) utilized household surveys from 10 WTP estimates for example, while Kanninen (1995) utilized simulated data to examine the effect that starting point bias could have. This paper differs by examining the potential starting point bias of surveys collected on-site. Most past studies utilized passive users of the resource while our respondents were active users. The advantage of the on-site survey implementation is that it lowers the hypothetical nature of the good. Presumably if the respondent is on the beach that is being valued, they have a clear idea of the good under consideration. There is likely to be much greater household preference uncertainty toward passive use values of distant resources than visitors would have toward a recreational resource they have visited repeatedly.

Statistical Modeling

Since CVM studies employ binary variables, it is required to use a statistical model appropriate for a discrete dependent variable. These models have been utilized in various other fields of economics such as labor. Such models were first developed in the field of biometrics where a stimulus was applied and the result observed; logit and probit models are the two primary examples (Hanemann and Kanninen 1999). CVM environmental studies typically involve describing precise changes in environmental goods or services. This makes the reliability of CVM studies dependent both on the accuracy of information presented and on the understanding of the information (Boxall, et al. 1996). This is another reason why the use of onsite visitor data might make a difference as a direct user ought to be more familiar with the recreation site.

Inherent to all econometric estimators is the need to balance the gains of efficiency with the biasing factors while still maintaining useful precision in estimates. The choice of which estimator to use is often determined by a matter of statistical properties of the potential variables informed by economic theory. The relevant statistical properties refer to the level of unbiasedness, efficiency, and precision (Greene 2008) a given estimator has. The researcher has to balance the gains in efficiency against possible increases in bias by choosing a particular estimation method over another. Further, if the decrease in precision is to such a point as to not be reliable or useful to policy makers, it would indicate that an alternative estimator ought to be utilized. From a statistical standpoint, we should be interested more in the magnitude of variance of the estimated for a willingness-to-pay study than the magnitude of bias from the estimated willingness-to-pay (Kanninen 1995). It is the confidence interval around the point estimate that determines how much variance a point estimate has.

Thus, the DB model can overcome a poor bid design by giving the researcher a second chance to provide information about their WTP whereas the SB model only provides the researcher one chance. This feature of the DB model increases the robustness of what might otherwise be poor bid designs compared to an SB model (Kanninen 1995). By increasing the amount of information provided by a follow-up question, the confidence interval is narrowed reducing the variance of a WTP estimate because information is added to the overall model.

Theory and Modeling

The first objective is to investigate whether or not the DB-DC CVM applied onsite for use value of a resource also suffers from an anchoring or starting point bias. In most CVM studies that employ some form of bidding game, the respondent is first asked, either in person or on the phone, if they would be willing to pay some initial bid amount. If the respondent is willing-to-pay, then the enumerator updates the information by asking if the respondent would be willing to pay an increased bid amount. If the response to the initial bid is in the negative, the enumerator asks if the respondent would pay a lower amount. The researcher can then use that information to estimate a Hicksian Compensating or Equivalent Surplus for the item under consideration (Boyle, Bishop and Welsh 1985).

The hypothesis is that the single bound interval WTP estimate lies between the two bivariate probit estimates of WTP values. Furthermore, the study hypothesizes that there will be less “starting point” bias or influence of the first bid amount on the response to the second bid amount when visitors are asked their WTP. In this study, the on-site nature of resource valuation might also lower the starting bid since the experience of the direct benefit of the resource could better assist the respondent in judging WTP for that resource. This is due to the fact that the visitors have more well established preferences for the recreation site which they have frequently visited. The standard double-bounded model of Hanemann, Loomis and Kanninen is also known by the statistical technique used, which is the interval method. It offers the greatest amount of increased efficiency with the least amount of ambiguity regarding preferences (Haab and McConnell 2002).

The interval data model operates with the assumptions that the respondent works under a single WTP value. The actual value is found through information that the respondent provides; the second question is thus updated with information from the first question. Therefore, the correlation between bid amounts one and two are equal to one. This is in contrast to the bivariate model of WTP measures. The bivariate model allows that the respondent might have two WTP measures. In other words, if the respondent’s exact WTP values could be observed, then the first response could predict a linear and unbiased estimate of the second value, subject to random response variables (Alberini 1995). The seemingly unrelated bivariate probit model of estimation is often used for estimation of double-bounded estimates.

Statistically, CVM models can be modeled along nominal or ordinal scales. In general, they can take on a finite number of values which can then be indexed. For *i* observation it will take on a particular value that can be represented as a function such that:

$$\Pr\{\,response_i = j\} = H_j(A_i; Z_i; \gamma) \qquad \text{(1) (Hanemann and Kanninen 1999)}$$

where A_i represents the bid on that occasion and Z_i represents the covariates that describe the subject, and γ is a vector of covariates estimated from the data (Hanemann and Kanninen 1999).

After the statistical estimation of the CV is accomplished, useful modeling then requires that it satisfy the economic role. This necessitates the modeling, and questions provide meaningful economic information. This is often accomplished through designing a model that is consistent with models of economic utility (Hanemann and Kanninen 1999).

If we assume that a respondent has some underlying utility function that can be defined, then we can build up a representation of that function. Under the assumptions of the Random Utility Theory (RUT), the general model of choice starts with an indirect utility, U, a consumer with income, I, a vector of attributes, $\boldsymbol{x}$, and individual preferences, e. We can then form the function:

$$U = v(I, \boldsymbol{x}, e) \qquad (2)$$

In the single-bounded CV study it involves asking the respondents if they would pay a given amount of money, B, for an amenity or improvement to quality for some good. The probability of a "yes" or "no" response can then be represented as:

$$\pi^n(B) = G(B; \theta) \qquad (3)$$

$$\pi^y(B) = 1 - G(B; \theta) \quad (4)$$

where $G(\cdot; \theta)$ is some statistical distribution function with parameter vector θ (Hanemann, Loomis and Kanninen 1991). This statistical model can then be interpreted as a utility maximization response within a random utility context where G is the cumulative density function (cdf) of the respondents' true maximum WTP because utility maximization would imply that:

$$\Pr\{NO\ to\ B\ \Leftrightarrow \Pr\{B > maximumWTP\}$$

$$\Pr\{YES\ to\ B\} \Leftrightarrow \Pr\{B \leq maximumWTP\}$$

(Hanemann, Loomis and Kanninen 1991).

The double-bounded statistical distribution, however, is more complex. A new consideration is that there are two responses, and the second response is contingent upon the first. Thus there are four possible responses, (a) yes to both, (b) yes to the first, no to the second, (c) no to the first, yes to the second, and (d) no to both. The likelihood of each response can be represented as π^{yy}, π^{yn}, π^{ny}, π^{nn}, respectively (Hanemann, Loomis and Kanninen, 1991). Continuing to follow the model set out by Hanemann, Loomis and Kannimen, then the term B_i^u represents the second bid amount for a "yes" first response, and the term B_i^d represents a "no" response to the first bid amount.

Utility maximization theory still allows building up a probability distribution for these cases. In the first one, we can model the probability as:

$$\pi^{yy}(B_i, B_i^u) = \Pr\{B_i \leq maxWTP\ and\ B_i^u \leq maxWTP\}$$

$$= \Pr\{B_i \leq maxWTP | B_i^u \leq maxWTP\} \Pr\{B_i^u \leq maxWTP\}$$

$$= \Pr\{B_i^u \leq maxWTP\} = 1 - G(B_i^u; \theta).$$

(Hanemann, Loomis and Kanninen 1991).

The statistical efficiency gains from the double-bounded CVM come from the additional information that is provided by the follow-up question, narrowing the range of the maximum WTP. We can build the probability function as follows:

$$lnL^s = \sum_{i=1}^{N}\{d_i^{yy} ln\pi^{yy}(B_i, B_i^u) + d_i^{nn} ln\pi^{nn}\left(B_i, B_i^d\right) + d_i^{yn} ln\pi^{ny}(B_i, B_i^u) +$$

$$d_i^{ny} ln\pi^{ny}\left(B_i, B_i^d\right)\}$$ (Hanemann, Loomis and Kanninen 1991). (5)

The statistical efficiency gains from the additional information are important. However, in this work it must be kept in mind that by increasing efficiency of the model, we may be doing so by accepting a higher amount of bias in the respondents. This biasing effect comes from the possible confusion or frustration on the respondents' part after being given a set price and then being asked if they would pay more or less for the same good. While that is the mechanism of auctions (raising or lowering the price until a purchase is made), that is not the intended design function in a CVM study.

To think of it more in terms of just WTP, the view the true vales as follows, where b_1 and b_2 are the first and second bid amounts respectively:

1. $b_1 \leq$ WTP $\leq b_2$ for yes – no responses
2. $b_1 >$ WTP $\geq b_2$ for no – yes responses
3. WTP $\geq b_2$ for yes – yes responses
4. WTP $\leq b_2$ for no – no response

The more general model then helps to understand the action that the double bounding takes. It helps the researcher to estimate the actual areas where the values for WTP lay.

Hypothesis

I tested the hypothesis of consistency between responses to first and second bid amounts assumed by the interval model. This was done by testing whether or not the correlation between the response to the first and second bid amounts as estimated from the bivariate probit model is equal to one. That is Ho: r=1 vs Ha r<1.

Data

The survey data is from a 2009 in-person and on-site survey conducted in Puerto Rico. It is presumed that the respondents are familiar with the resource being valued as they can not only see the resource but can also touch it. In the survey, each respondent had two questions in the survey regarding the value of the beach: an initial random bid amount and a follow up bid amount, along with standard socio-demographic questions.

If the response to initial random bid was “Yes,” the follow-up bid = (2*initial bid amount). Conversely, a “No” response leads to a follow-up bid = (initial bid amount/2). The initial bid amounts were from $5 - $150. The bid amount was in addition to the amount that was paid to travel to the beach. Of the 660 survey responses, 657 were usable for this analysis. Three surveys had to be purged due to enumerator error in recording a “zero bid”. This error was noticed on the first day of survey collection and fixed on subsequent days.

The starting point bias is also referred to as an anchoring effect: if one person offered someone else a price of $X, why should he now have to pay $X + 2X?

In most previous literature the follow-up question was asked at the household level. An issue that has arisen in the literature is the effect that anchoring might have on the follow up bid. Herriges and Shogren (1996) point out respondents combine the prior WTP value provided by the first bid, thus anchoring the WTP. They model as:

$$W_{1i} = w_{0i} \text{ and } w_{2i} = (1-\gamma)W_{1i} + \gamma B_{1i}$$

where $0 \leq \gamma \leq 1$ (Herriges and Shogren 1996).

Results

Table 4.1. Results from testing gains of alternative models.

Model	Constant	Prob>Z	Price Coefficient	Prob>Z	Log Likelihood
Single Bounded Logit	1.159902 (0.141577)	0	-0.019501 (0.001872)	0	-35.101
Double Bounded Logit	1.256759 (0.097217)	0	-0.02647 (0.001369)	0	-891.881
Bivariate 1^{st} Equation	0.672053 (0.084653)	0	-0.01123 (0.0010513)	0	-804.593
Bivariate 2^{nd} Equation	0.194148 (0.079699)	0.015	-0.00693 (0.00099)	0	

Standard Errors in Parenthesis

In this analysis, we reject that the null hypothesis of the correlations between the BVP equations 1 and 2 are equal to 1. The correlation between the 1^{st} and 2^{nd} BVP equations is approximately 0.32, with a standard error of 0.08. This is very different from a correlation of 1 between the two equations that the DB model assumes is present when it estimates a single constant and bid coefficient. This implies that even for respondents who visit a given site there is a marked difference between the WTP for first bid and second bid. This is surprising since one might expect that experienced users have a clear idea of the good being valued relative to household surveys. This could indicate that the shift is intrinsic to the DB model. As Carson points out, if the good was possible to be purchased at $25, then paying $50 would be ridiculous for the same good. Additionally, if it was $50, then how can the good have equal quality at $25?

We used a double bounded logit model. As expected, the confidence intervals show the DB logit with the shortest length. In Table 2, the WTP's are reported. We see that that there is a reduction in the variance from including this second piece of information as a means to increase efficiency. However, an improvement in the point estimate or improvement in the stability of the mean WTP does not necessarily provide evidence that anchoring is at work (Herriges and Shogren 1996).

We accept Null Hypothesis of WTP SB=BVP(1st Eq)=DB, while we reject Null Hypothesis of WTP BVP(2nd Eq)=DB. There is a difference between response to the first and second bid, even for those who visit the sites. This is interesting because the interval median WTP does not fall halfway between the two bivariate probit estimates, but it is closer to the higher WTP than the lower. The median WTP for the double bounded logit does not fall halfway between the bivariate probit. There is a marked difference in the upper end of the WTP curves of the Bivariate Probit models & the DB, suggesting that differences in Mean WTP could be substantial.

In figure two, we see the shape of the WTP curves plotted against the probability of a yes response. The steepness of the Bivariate Probit 2nd Equation is worth noting. It likely indicates that there is something intrinsic about the second bid that causes a “NO” response. This would not be expected if dollar amounts are held constant. For example, 20% of respondents would say “NO” to a first time bid amount of $50, and 50% would say “NO” if that same amount was the second bid. This likely indicates some level of anchoring at work in the model. It could be that respondents just do not like that second bid amount if they have already been provided a prior bid, either higher or lower. The marked difference between upper end of WTP curves for the BVP and DB models could indicate that there are substantial differences between the two models.

Table 4.2. Confidence Intervals From Different Models

MODEL RESULTS

LOGIT		BIVARIATE PROBIT	
BINARY LOGIT	90% CI's	1st EQ	90% CI
Upper Bound	$ 67	Upper Bound	$ 68
Median WTP	$ 59	Median WTP	$ 60
Lower Bound	$ 52	Lower Bound	$ 52
DOUBLE BOUNDED LOGIT		2ND EQ	90% CI
DB LOGIT	90% CI	Upper Bound	$ 41
Upper Bound	$ 52	Median	$ 28
Median	$ 47	Lower Bound	$ 11
Lower Bound	$ 43	AVERAGE	$43

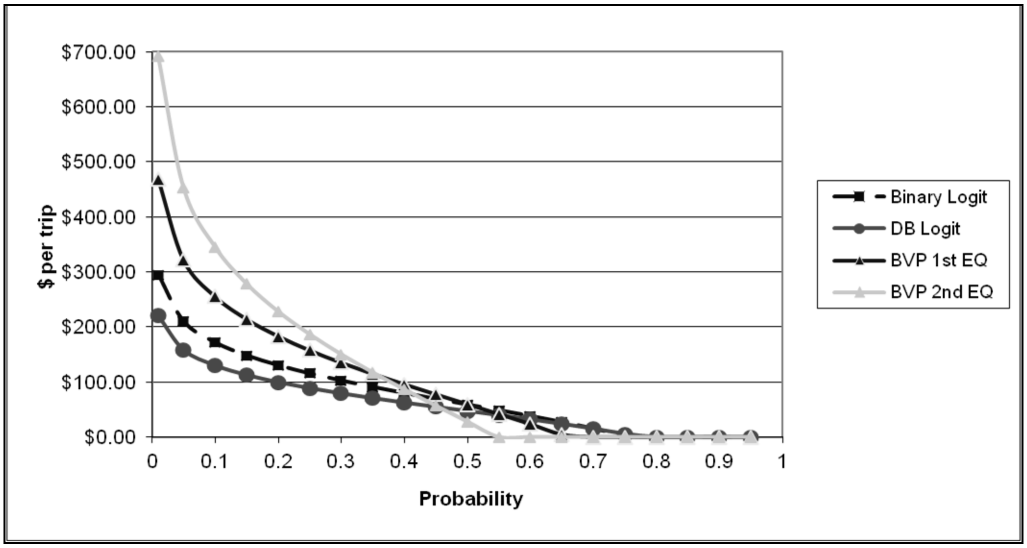

Figure 4.2. Comparison of WTP Curves under different models

Conclusions

Finding that there is a difference in the WTP values based on if it is the second or first bid, determining that there is a shift in WTP from the first to second bid, all indicates that the shift could be intrinsic to the DB model. This could be due partly to the fact that the second bid amount has endogeneity (Cameron and Quiggin 1994). The second bid after all is often one half or two times the original bid amount. Another possibility is that the respondent is updating his or her information based on the previous bid amount offered; somehow inferring that that amount is the “correct” amount.

There is also a marked difference between the first and second response to the bids by respondents who visit the recreation site. This is surprising as one might expect that experienced users would have a firmer idea of what their WTP would be relative to household surveys. Although the follow up question provided an increase in information, and the resulting reduction in variance is useful, questions about the model still remain.

It would be worthwhile to explore how this shift might be mitigated in future surveys, particularly if the efficiency gain from the DB model is valuable to the researcher. One possible avenue of exploration might be to test whether or not smaller follow-up bid amounts also exhibit a substantial shift in response. This could be carried out perhaps by asking a portion of the respondents one follow-up bid amount, while asking another portion a smaller change in the follow-up bid amount.

The efficiency gains from the double-bounded dichotomous choice CVM are very substantial and offer researchers the potential of communicating better information to policy makers. As we see in Table 2, the double bounded interval model provides a 90% WTP confidence length of $9.00, compare that to the single bounded logit or the bivariate probit 1st equation lengths of $15 and $16 respectively, and we have greatly reduced the variance. For the resource manager desiring to keep make sure the benefit cost analysis would pass at either end, the double bounded model provides the best result.

In addition to that efficiency gain, since the marginal costs of asking one extra question is low after the respondent has agreed to answer a survey, abandoning the DB format would seem foolish. That said, the effects of anchoring do need more study and consideration for useful estimation to ensure the model's weaknesses are understood and overcome.

Environmental economists need to utilize all the tools that are available in order to be as effective as possible. As the requirements for more resource valuation continue to increase, refining the accuracy is also needed.

WORKS CITED

Albernini, Anna. "Efficiency vs Bias of Willingness-to-Pay Estimates: Bivariate and Intravel-Data Models." *Journal of Environmental Economics and Management*, 1995: 169-180.

Boxall, Peter C., Wiktor L. Adamowicz, Joffre Swait, Michael Williams, and Jordan Louviere. "A Comparison of stated preference methods for environmental valuation." *Ecological Economics,* 1996: 243-253.

Boyle, Kevin J., Richard C. Bishop, and Michael P. Welsh. "Starting Point Bias in Contingent Valuation Bidding Games." *Land Economics* 61, no. 2 (1985): 188-194.

Brown, Thomas C. "Introduction to Stated Prefernce Methods." In *A Primer on Nonmarket Valuation*, by Patricia A. Champ, Kevin J. Boyle and Thomas C. Brown, 99-110. Norwell: Kluwer Academic Press, 2003.

Cameron, Trudy A., and John Quiggin. "Estimation using contingent valuation data from a "Dichotomous Choice with follow-up" questionaire." *Journal of Environmental Economics and Management*, 1994: 218-234.

Greene, William H. “Econometric Analysis.” 6th. Upper Saddle River, New Jersey: Pearson Prentice Hall, 2008.

Haab, Timothy C, and Kenneth E McConnell. “Valuing Environmental and Natural Resources.” Northampton: Edward Elgar Publishing, Inc., 2002.

Hanemann, Michael, John B. Loomis, and Barbara Kanninen. "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation." *American Journal of Agricultural Economics* 73, no. 4 (1991): 1255-1263.

Hanemann, Michael, John Lommis, and Barbara Kanninen. "Statistical Efficiency of Double-Bounded Dichtomous Choice Contingent Valuation." *American Journal of Agricultural Econoimcs*, 1991: 1255-1263.

Hanemann, W. Michael, and Barbara Kanninen. "The Statistical Analysis of Discrete-Response CV Data." In *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EC, and Developing Countries*, by Ian J. Bateman and Ken G. Willis, 302-402. New York City: Oxford University Press, USA, 1999.

Herriges, Joseph A., and Jason F. Shogren. "Starting Point Bias in Dichotomous Choice Valutation with Follow-Up Questioning." *Journal of Environmental Econoimcs and Management*, 1996: 112-131.

Hilmer, Christina E, and Jayson L. Lusk. "Determinants of Citations to the Agricultural and Applied Economics Association Journals." *Applied Economic Perspectives and Policy* 31, no. 4 (2009): 677-694.

Kanninen. "Bias in Discrete Repsonse Contingent Valuation." *Journal of Environmental Economics Management*, 1995: 114-125.

National Oceanic and Atmospheric Administration (NOAA). “Natural Resource Damange Assesment Under the Oil Pollution Act of 1990.” Federal Government, Federal Register, 1993.

Roach, Brian, Kevin J. Boyle, and Michael Welsh. "Testing Bid Design Effects in Multiple-Bounded Contingent-Valuation Questions." *Land Economics*, 2002: 121-131.

CHAPTER FIVE: Scope Sensitivity in Contingent Valuation: A Meta-Analysis

Stated Preferences, Contingent Valuation Tests and Background

Stated preference studies have to pass a number of tests. One test is what is termed the validity test, in a nonmarket valuation context it is checking if the estimated value measures theoretical construct under consideration. Does the measurement provided accurately reflect the value under consideration? The surest way of verifying that a CVM study passes the validity test to compare the values with actual purchases, but in the nonmarket context that is nearly impossible, particularly since the estimation is taking place because of the lack of a direct market example. Given that difficulty it is often needed to examine three parts of what go into provide validity in a study. Criterion, construct and content validity are all pieces that go into the validity test of a CVM study. Criterion validity refers to the comparison of stated preference values with other values. For example, a referendum vote compared to a controlled experiment where the participants actually pay for the good under consideration (Brown 2003).

The construct validity examines how the measure under consideration relates to other measures under consideration. For instance, if a travel cost study for a resource provides a value, does a contingent valuation study provide a similar value for that resource? Construct validity also examines how theory predicts certain variables relate to the expected relationship measure of the good. For example, how income relates to variables thought to influence willingness-to-pay.

The third part is content validity. This takes a different approach from criterion or construct, content asks if the survey tool itself is asking the needed question and getting the needed responses. Content validity asks if the questions are clear and unambiguous, are the statistical functions correct to the problem at hand and correctly implemented. The payment mechanism meets the expectations of being realistic and credible. Those pieces need to fit together to create a CVM study that is valid and provides a reasonable valuation.

To conduct a good contingent valuation study certain steps need to be taken. The data collection methodology and order was discussed in previous chapters. One critical step in the data collection process is selecting the data collection mode. The most common method is the mail survey, though the in-person survey is advocated by the NOAA Blue Ribbon panel (1993) as the optimal method. Phone surveys are often used as well, and mixed mode surveys are also used (Boyle 2003). Each mode has their relative strengths and weaknesses, not the least of which is the relative costs of implementing each (Boyle 2003).

Once the relevant population and mode is determined, the sample size is needed with the corresponding response rate to achieve the needed level of precision. Once that is done, the researcher needs to estimate the content of the information of the survey; this includes describing the good to be valued and the payment mechanism. Before going to the field, pretesting the survey is critical to ensure that respondents and researcher are valuing the same goods. Once all of that has been accomplished, then the survey can be implemented. Final step is analysis of the data and communicating the results to the relevant persons. The reader interested in a more detailed description of each step is encouraged to read Boyle's 2003 "Contingent Valuation in Practice". At the end of the research, the preferences of the people should have been clearly communicated in a way that is consistent and provides meaningful information.

After the researcher has presumably accomplished correctly the above steps, the survey meets the three validity tests, is the value found in a CVM study meaningful? Some critics of the method would say the next test required is the scope test.

Scope

Within the realm of contingent valuation (CV), a survey method to model respondents' preferences, most often for environmental goods, a discussion has been going back and forth regarding the scope of the results found.

Scope refers to the idea that a CVM study should reflect the theoretical predictions. So if more (or less) of a good is provided, there should be an increase (or decrease) in willingness-to-pay for the good. The idea is that a "good survey will show respondents are sensitive to significant and show substantive differences in the public good" (Haab and McConnell 2002). One result from the Exxon-Valdez oil spill and the resulting controversy about the validity of CVM studies is that scope criterion. A more full discussion of that is talked about in "Valuing Environmental and Natural Resources" by Haab and McConnell (2002).

Economic theory dictates that more of a desired good leads to greater consumer utility and thus providing greater economic value (Heberlein, et al. 2005). The NOAA panel (Arrow et al., 1993) recommends that any CV study ought to pass a scope test, following rational expectations and real marketplace activates. The concern that CVM studies are insensitive to scope arose in 1986 when Kahneman presented a graph pointing out that the willingness-to-pay (WTP) for cleaning up all the polluted lakes in Ontario was the same as cleaning up all the lakes in Muskoka, a small part of Ontario (Carson 1997). Since Kahneman is respected in the field, his critiques were taken seriously. Carson (1997) in a rebuttal points out that respondents were expressing an ideological value rather than an economic value. A study will be said to have passed the scope test if two levels of the good have differing WTP values that are statistically different and concur with the expected direction of change.

“Sensitive to significant and substantive differences” (Haab and McConnell 2002), part of the controversy regarding CVM is that people have conducted studies where respondents are willing to pay the same for 10,000 or 100,000 additional birds in a fly way. Economic theory would predict that if WTP is linear, there should be about 10 times more WTP for the additional birds, however no significant difference was found. But given diminishing marginal utility, WTP is likely to be nonlinear and increase but at a decreasing rate. How fast WTP should level off is an empirical question and one that may vary with the initial baseline quantity of the good and the amount of the increment or decrement of the good offered in the survey. Is the failure to find scope an indictment of the whole field of study, or is there something else at work? This is an important aspect of CVM since it lends itself to content validity. While that is a purely theoretical concern, it does have importance on a number of levels.

For the CV study to hold influence it must meet above three internal consistency checks some are common to the stated preference tests listed above and the third is unique to CVM. Smith (1996) lists these as: 1) the choices should be responsive to the scope or amount of the good being offered by the respondents, 2) the CV study should pass the construct validity test, that is, the survey needs to be put together in such a way that the relevant economic factors are accounted for, and 3) CV studies should find significant differences between goods that are generally agreed to be different (Smith 1996). This paper will seek to explore the variation in tests of scope in Contingent Valuation (CV) studies.

The wide variety of studies done has shown that CV studies can be surprisingly sensitive or insensitive to scope or scale depending on the information presented (Brown and Duffield 1995). The effect of information provided on the sensitivity has been termed the *part-whole* valuation effect by Mitchell and Carson (1989), or the *embedding effect* by Kahneman and Knetsch (1992).

This also relates to the "part-whole" or "embedding effect" that can arise in scope tests. That is, if a respondent would be willing to preserve one part of a ecosystem, economic theory would indicate that they ought to be willing to pay more to preserve the whole ecosystem. Yet at times respondents would pay more to preserve a single species than a whole ecosystem. This distressing result has caused the researchers to question the usefulness of the whole CV method. Embedding effects can arise from either a flawed survey design, or from the inability of the respondent to distinguish the parts (Boyle, Desvouges, et al. 1994), (Schulze, et al. 1998).

Desvousges et al (1993) provided the example of a) 2000, b) 20,000 or c) 200,000 birds being prevented from being killed, where the WTP for those widely different values was only a) $80, b)$79, and c) $88. This did not show significant differences in WTP, nor the satisfactory sample validity and reliability requirements one would hope for (Desvousges, et al. 1993). However, it should be noted that Boyle et al (1994) pointed out that while the numbers can be seen as large; there is also the issue of percentage changes. The largest number is less than 2% of the total bird population. This perhaps indicates that at the margin the respondent is telling the researcher how much they would pay which just happens to differ insignificantly across the board. These features do not need to correspond with irrational economic behavior (Ojea and Loureiro 2009). Determining what is going on is a problem for the researcher not the respondent.

The issue of insensitivity to scope can be described as a "weak test of economic theory" (Veisten, et al. 2004). Understanding why respondents provide the same WTP for a given set of goods of significantly different value size, leads to the question of the accuracy the value provided. Lack of such sensitivity is termed embedding at times. Economic theory would suggest that if someone was willing to pay something to obtain a certain environmental good, that person would also be willing to pay more to obtain more of that good (Pouta 2005). However, that assumes that the utility gained from the resource is linear. If the good exhibits diminishing marginal returns, meaning it is a nonlinear function, then the significant gains in utility would come from the first few units of the good provided. Additional units would not provide as much utility as the initial increase, so testing for sensitivity to scope through a linear function would provide misleading results.

Another area where scope issues arise is in what can be termed the part-whole issue. That is, the researcher sees a good as a single item while the respondent views the good as two distinct items. This can occur in environmental economics when joint-production from an ecosystem service occurs (Carson 1997), (Pouta 2005). Since ecosystem services often contain joint production functions, survey design that is weak in the area of construct validity would likely be subject to such concerns.

Insensitivity to scope violates the rational choice behavior that is axiomatic for neoclassical economic theory. Such assumptions lead to the assumption of non-satiated utility theory that WTP should be higher for a ‘higher number’ than a ‘lower number’. It also follows neoclassical theory, which would predict a higher price would be paid for the whole set of goods than the subcomponents that go into the good. Various explanations have been offered as to why scope insensitivity arises in research. It could be as Carson and Mitchell (1993) argue that design flaws are the primary reason for scope insensitivity. The insensitivity is a result of poor survey design or problems while implementing the survey. A poorly designed or explained good might be understood by the respondents in a different unit than the researcher intends (Pouta 2005).

Economic theory also provides insights into why there could be scope insensitivity. One such example would be the argument of diminishing marginal values. That is, restoring or preserving the first unit of the resource would have a strong value, while additional unit(s) of the resource would have positive but lower per unit values. When that is the case, it would be required to have a baseline scarcity of the resource to build from in order to test the level of scope sensitivity (Veisten, et al. 2004). There is also the mental model of a joint product in that the researcher might see the goods as additive or increasing; the respondent sees it as only a whole unit regardless of how the researcher wishes the respondent to understand the good (Schulze, et al. 1998).

A similar example would be the effect that income effects have on a respondents' scope sensitivity. The respondents' have a limited budget, and as such their spending is constrained. So when they are asked to pay a hypothetical amount in a bid design, respondents are still expected to pay a limited WTP and that value would hold when asked to value another good as well (Veisten, et al. 2004).

Embedding effects can result from a variety of sources. The argument is made that there is a moral satisfaction to giving towards a cause or good. That satisfaction then declines quickly once the initial amount is offered (Schulze, et al. 1998). Thus, if an individual chooses to preserve a single species in an ecosystem and does so, giving more towards that ecosystem does not have as much value as making that first bid.

Another reason for why critics of the CV method argue that insensitivity exists because of the hypothetical nature of the format. Respondents will always express a similar WTP across similar goods. Critics argue that what is being measured is a 'warm glow' or altruism effect of stating that they would contribute to the good, regardless of the scope (Pouta 2005). Regardless of the environmental good, respondents have a certain amount that they would contribute, testing the WTP of people wanting to give to a good cause (Smith 1996).

Testing for scope sensitivity has been done internally within subjects, as well as externally between subjects (Pouta 2005). Internal tests examine the WTP of the same subject to pay for differing amounts of a good, while external tests between subjects use split sample survey designs as recommended by Arrow et al (1993). A scope test requires people to be willing to pay more for a progressively larger amount and less for a smaller good (Smith 1996).

In order to investigate whether most CVM studies exhibit scope or not, and what survey design features lead to scope I undertake a meta-analysis of past CVM studies that have tested for scope.

Methods

Why use meta-analysis? It is a concise method of quantifying a diverse set of results into a single study. As Stanley (2001) pointed out, meta-analysis reveals unexpected results and ordering to areas "infused with controversy" (Stanley 2001). Meta-analysis was first proposed in 1976 by Glass as a method to study evidence across empirical studies. The use of meta-analysis dates back to its first usage in medical literature as a means to examine a large number of studies that have potentially inconclusive results to determine if a conclusive result could be found. The use of meta-analysis as a tool was resisted at first by the medical profession who preferred to rely instead upon more traditional clinical trials. Economists also have been reluctant at times to adopt the results of meta-analysis since they could be mixing the results of "good" studies with those of the perceived "bad" studies. However, no less an august body than the American Statistical Association supports the use of meta-analyses, even when based only on small samples (Hunt 1997). The use of meta-analysis in an economics realm began in 1989-

1990 when Stanley and Jarrell (1989), Jarrell and Stanley (1990), and others began utilizing the technique. Nelson and Kennedy (2008) point out that one-half of 140 studies in the environmental area have been done since 2004, so meta-analysis is a growing area of research. Stanley (2001) provides a table of examples of how meta-analysis has been utilized in economics. The list covers things as divergent as benefits of endangered species (Loomis and White 1996) to union wage premiums (Jarrell and Stanley 1990).

Meta-analysis was first proposed in the 1970s as a method for systematically providing a quantitative summary of evidence across empirical studies. Applications of meta-analysis in the field of economics began in the late 1980s (Nelson and Kennedy 2009). In medical studies it has been found that use of a meta-analysis can provide clarity. When multiple studies provide different answers, and combined through the meta-analysis framework, a clear answer is then presented from the numerous different studies (Stanley 2001).

An environmental or natural resource economist would likely want to use meta-analysis to quickly summarize the values from various environmental studies for many reasons. First, meta-analysis could be used to determine the combined effect size from the numerous studies available. A second reason might be to determine the size of the wide variation in effect size from study-to-study or examine the causes of heterogeneity in the samples. Based on the results, suggestions for the improvement of primary data collection or model specification techniques would be provided. A third reason to conduct an environmental meta-analysis is to provide an in-sample estimate of predicted values of the dependent variable under a particular set of conditions. Fourth, an out-of-sample prediction might be desired as part of a benefit-transfer application. The EPA has characterized meta-analysis as the “most rigorous benefit transfer exercise” (Nelson and Kennedy 2009). Fifth, meta-analysis can be used to summarize the results of a single study that produced multiple estimates (Nelson and Kennedy 2009). Additionally meta-analysis can be used when time and money constraints are an issue. Meta-analysis can also be useful when a preliminary analysis is conducted to determine whether or not further investigation is warranted since collecting primary data is a time consuming and costly process. It also serves as a useful check on primary data estimates, to ensure that the value found is consistent with other work in the same field.

There are multiple possible goals to be accomplished by a meta-analysis (Nelson and Kennedy 2009); and they need not be in conflict with one another. One objective is to provide a "combined" estimate of the effect size, another objective is to discern the often wide study-to-study variation in effect sizes. Out of sample prediction can also be utilized, along with that benefit transfer is often at its most empirically rigorous when done through meta-analysis (Nelson and Kennedy 2009).

To date, a number of different tests have been conducted to determine of scope sensitivity is present in the researchers study. To our knowledge there has not been a recent study looking at the scope sensitivity through a meta-analysis. The most recent found was conducted in 1996 by Smith and Osborne. They looked at only 5 studies that used similar methods to examine visibility changes at National Parks (Smith and Osborne 1996). That paper is unique since it only used 5 papers it is felt that expanding the number of studies wider may provide more insight. Most research on the scope issue has been limited to extensive literature reviews, and finding that to date the results are inconclusive. This would present the opportunity for deeper and more quantified study of what causes sensitivity or lack thereof in the contingent valuation (CV) line of research – and a good use of the meta-analysis method.

Model

To test for factors that influence the presence or absence of scope a data set was constructed. The first stage in conducting the test was an internet and database search for CVM studies that tested for scope, culling from various searches of studies that have utilized CVM and tested for scope, while looking at studies that examined environmental resources. Future research might be profitable to expand the search to include CVM studies that looked at other goods. However, ensuring that the model would be tractable was critical.

Since the scope variable is the variable of interest, it was chosen as the dependent variable. A study could fail, have mixed results, or pass the scope test. In those studies that reported different subsamples, if the reason for a "mixed" result was two subsets, the subsets were broken apart and then treated as a "pass" or "fail."

Data and Methods

To closely examine the issue of scope in CVM, search of the literature found articles that utilized CVM to examine environmental resources and to test for scope. Limiting the search to only environmental goods helped ensure that the same variable was being compared – that is apples were being compared to apples. Of the approximately forty studies found a few were from the same article, and were separated into distinct studies.

The choice was made to limit the study to published articles since there are numerous articles that pass, fail and present mixed results for the scope test. Since the strength of the analysis is partly dependent on the strength of the data, and peer reviewed publication is the gold standard, limiting the study to published data seemed reasonable. While this does depart from the recommendation of Stanley (2001), it follows the Nelson and Kennedy (2008) point that low quality primary studies are likely to have increased heterogeneity and therefore are reasonable to be omitted.

In searching for the potential sources, extensive literature searches were done, making use of sources like google.scholar.com, Web of Science, and JSTOR. The key words used were: environmental, economic, nonmarket valuation, contingent valuation, scope, and resource valuation. From the articles that then came up, I narrowed it down to articles that did scope tests. Also, those papers that were found to be useful often cited other papers that I then tracked down for inclusion where possible. The complete list of papers is shown in Appendix A.

The chosen dependent variable is the scope variable. This follows the recommendation of Stanley (2001) that choosing the most logical summary statistic should be a common and comparable metric. Here all the studies utilized either report one of three outcomes with regard to scope: Fail, Mixed, or Pass. Coding that followed the common protocol of Fail = 0, Mixed = 1 and Pass = 2. Of the 42 distinct data points, there were 14 that failed, 18 that passed, and 10 that showed mixed results. Those studies that had a "mixed" result where possible were separated out between the part that passed and failed. Because of the coding of the dependent variable, ordered probit is appropriate as the statistical model.

The independent variables chosen to examine were sample size, mode, passive use and year. All the studies were provided a measure of Hicksian surplus, and if the study provided information. If the study provided information from both a revealed preference and stated preference result, only the stated preference information was used. While using the WTP values has, at first glance, some appeal, its usefulness is limited in this situation since the various natural resource values are so different. In addition the potential endogeneity exists because whether or not a study passes or fails depends on the WTP, and it is captured in the dependent variable of if the study passes or fails the scope test. Therefore the choice was made to not utilize that variable. In addition not every study reported the exact way the sample was split, so establishing the relevant numbers would not have been practical.

It was also tested to see if active users of the resource versus passive users of the resource were found to make a difference in the determining the presence of scope or not. Active users were assigned a 0, while passive users were assigned a 1.

The final two hypothesis tests are on if year of the study and if sample size will be positive and significant. The calculation of year value is taking the current year and subtracting the date of publication from that. For example: a study published in 1997 would have a year value calculated as: 2012 – 1997 = 15.

For sample size, the underlying assumption is that an increase in the number of respondents increases efficiency and would lead towards a higher sensitivity to scope. The null hypothesis would be that the coefficient on sample size is equal to zero, the alternative would be that it is not.

This is summarized in Table 5.1.

Table 5.1. Code Sheet

Variable	Coding	Numbers
Scope	Fail = 0 Pass = 2 Mixed =1	Fail = 14 Pass = 18 Mixed =10
Year Value	2012 – Year = Value	Mean = 15 Standard Deviation = 8.85 Min = 1 Max = 32
Mode	1 = In Person 2 = Mail 3 = Phone 4 = Mixed Mode	1 = 10 2 = 17 3= 8 4 = 5
Sample Size	Reported by Studies	Mean = 570 Standard Deviation = 389 Min = 69 Max = 1678
Passive Use	0 = Active 1= Passive	Active = 13 Passive = 29

Since meta-analysis is a combination of multiple empirical studies, the characteristics of the studies are unlikely to be exactly the same in each setting. The sample sizes are likely to vary from study to study, a method of dealing with the varying sample sizes is critical. To account for the "effect size" a common metric needs to be devised. Stanley (2001) points out that originally, effect size was defined as "the average outcome of the treatment group minus the average outcome of the control group, divided by the standard deviation of the control group." Effect size then allows for different and diverse studies to be compared directly on the same scale (Stanley 2001). Conversion to a standard normal scale that is not dependent on the degrees of freedom from the tests is critical to develop a meaningful meta-analysis.

Due to different primary sample sizes, different sample observations, and different estimation procedures, effect-size estimates are likely to be non-homogenous. Effect estimates with smaller variances are more reliable. When these variances are known, the modeler can account for this and provide a better estimate. However that is rarely the case in practice (Nelson and Kennedy 2009).

When combining different studies with various mechanisms of collecting and measuring information, the issue of sample heterogeneity is introduced. The varying effect-sizes from primary studies are “not all estimating the same population effect” (Nelson and Kennedy 2009). Controlling for this heterogeneity can be resolved in a number of ways. One method would be to throw out studies that are of low quality with particularly poor data or methodologies (Nelson and Kennedy 2009). Though given that often the reason for a meta-analysis is a lack of primary data on a given area or value, the researcher might not have the luxury of disregarding studies even those of poor quality. Stanley (2001) would recommend erring on the side of inclusion rather than exclusion. Since this study looks very narrowly at the scope test and how authors have defined how the study passes or not, as many studies as practical were included.

Another way to control this heterogeneity would be directly through meta-regression. Typically through the use of dummy variables for potential sources of heterogeneity (Nelson and Kennedy 2009), these dummy variables are then able to control for variables like income, location, and time period. If the sample size permits, meta-analysis on more homogenous subsamples is also a good method of controlling for heterogeneity.

A second less common method of controlling heterogeneity would be to model the sample population effect-size as random draws from a distribution. Thus each study then is utilized as estimating a different population effect size (Nelson and Kennedy 2009). This is often known as the random-effect-size model.

To summarize, when examining different population sizes, the analysis should address the heteroskedastic nature of the data through the use of appropriate econometric techniques. These methods could include weighted least-squares, multilevel models, and panel data methods (Nelson and Kennedy 2009).

If the researcher suspects that the effect-sizes are correlated, due to research using the same public data, time-series data issues, or two or more values reported from a single study, then steps need to be taken to control the correlation. When correlation is present and OLS is used for the regression analysis, robust standard errors are critical for inference. If a panel data set is constructed, random-effects estimation has many advantages and should be used unless a very strong case can be made against its usage (Nelson and Kennedy 2009). When grouping values together from a number of primary studies to develop clusters or panels, the grouping criterion is often not obvious. To show sensitivity to specification, empirical results from several relevant stratifications of the data should be reported (Nelson and Kennedy 2009).

Following that, several specifications of the models was tested. The regressions were carried out in Stata 10. Modeling a simple ordered probit model, not controlling for the author effect was utilized, shown in Table 2. To control for the author effect, a regression analysis utilizing a random effects ordered probit model with author as the identifying variable. That model is presented in Table 3. Finally a model was run that tested to see if the NOAA Blue Ribbon publication on how to conduct CVM studies had an effect. This is coded as 1 if study was done before the NOAA panel recommendations and zero after. This is presented in Table 4.

Table 5.2, Ordered Probit

Variable	Coefficient	Standard Error	Z	Prob > Z
Year	0.0688	0.0401	1.71	0.087
Sample Size	0.0004	0.0004	0.83	0.408
Passive Use	-0.0619	0.3864	-0.16	0.873
Mail	0.2612	0.4488	0.58	0.561
Phone	0.0511	0.5214	0.10	0.922
Mixed Mode	-0.1318	0.6095	-0.22	0.829

N=42. Likelihood Ratio $\chi^2(5)$ = 0.97 Prob >χ^2 = 0.9652. Pseudo R^2= 0.0452 The baseline of in-person was used.

Table 5.3, Random Effects Ordered Probit

Variable	Coefficient	Standard Error	Z	P > Z
Year	0.0864	0.2190	3.93	0
Sample Size	-0.0049	0.0022	-2.15	0.032
Passive Use	-4.3175	1.4256	-3.03	0.002
Mail	14.9547	3.1997	4.67	0
Phone	4.0884	2.7257	1.50	0.134
Mixed Mode	-7.4068	2.4099	-3.07	0.002

N=42. Likelihood Ratio $\chi^2(6)$ = 19.72 Prob >χ^2 = 0.0032. The baseline of in-person was used.

Table 5.4, Random Effects Probit with NOAA Effect

Variable	Coefficient	Standard Error	Z	P>Z
Year	7.7253	2.6992	2.69	0.007
NOAA Blue Ribbon	-29.3795	12.9046	-2.28	0.023
Sample Size	0.0088	0.01839	0.48	0.630
Passive Use	-7.2231	9.7103	-0.74	0.457
Mail	48.3062	9.2360	5.23	0
Phone	-0.7352	10.8232	-0.07	0.946
Mixed Mode	18.1241	503.9057	0.04	0.971

N=42. Likelihood Ratio $\chi^2(7)$ = 25.01 Prob >χ^2 = 0.0008. The baseline of in-person was used.

Examining the ordered probit model, the Pseudo R^2 value for the overall model is very poor, and none of the variables proved significant. Not controlling for heterogeneity indicates that there is a lot of noise that cannot be accounted for.

The random effects ordered probit model is significant; however some of the results are surprising. Year is significant and positive, meaning (recall how Year is coded as the difference from 2012) older papers had a greater likelihood of passing the scope test. Sample size is perhaps the most troubling result, as an increase in sample size should reduce the variance of an estimator and the WTP and hence increase the likelihood of passing the scope test. Yet here our results seem to indicate that an increase in sample indicates less probability of passing the scope criterion.

The final model while significant, likely suffers from a high degree of multi-collinearity. This is indicated by the extremely high coefficient values, and that Year and the NOAA Blue Ribbon dummy are related. The switching of sign on Sample Size, Phone, and Mixed Mode variables all indicate high correlation of those variables. The older the study, the more likely it was to have been published prior to the NOAA report. An attempt to run the model without the Year but including the NOAA Blue Ribbon dummy variable was made, but the model failed to converge.

These three regressions indicate that modeling quantitatively for scope through a meta-analysis is very sensitive to the model chosen. Ideally the results would be more robust to model. The most reasonable model specification is one of the two random effects ordered probit models, most likely without the NOAA Blue Ribbon dummy variable. The random effects models contain controls for heterogeneity, and the ordered probit is correct for the issue of an ordinal dependent variable.

Conclusion

In examining the issue of scope there is little that can be said conclusively. In this quantitative analysis of various studies that look at environmental resources, only one model provided results that indicate what would be a significant contributor to scope sensitivity. There are few significant variables that indicate what would lead to a particular study passing, or what would indicate a study failing, and none that are robust to model. By limiting the meta-analysis to those studies that were published in peer-reviewed journal articles, presumably the studies were conducted within the bounds of accepted standards of CVM survey collection methods, thereby meeting the critical survey validity tests.

Mode in the random effects probit model without the NOAA Dummy was significant. Phone was never significant while mail was in two of the models. However, in contrast to the NOAA Blue Ribbon Panel recommendation, it seemed that mail is a better method of detecting scope than in-person surveys. Mail may work better because the respondent can see visually by the use of maps and graphs the different proposed changes in the amount of the resource or the increment, and hence responses may be more likely to be sensitive to scope. By the scope criterion, it seems that it is possible to fail the scope test regardless of what mode you utilize. While there are other reasons for preferring the in-person interview method, the CV method need not be viewed as weaker even though collected through alternative modes. This could be an indication that mail is actually better than an in person survey. The in person survey may cause the respondent to provide what they think the enumerator wants to hear, while with a mail survey there is no such action.

Somewhat surprising is the impact of sample size. It either has no impact, or it has a negative impact for detecting scope. This could indicate that as long as there is a sufficiently representative sample, more respondents will not necessarily help promote scope sensitivity. Again, there are other reasons for wanting larger sample sizes that relate more towards narrowing of the confidence intervals for a given estimate. However, it would seem that when a sample is large enough to be representative, sample size does little in the way of contributing to passing the scope test.

For the researcher having to weigh the costs of in-person survey collection against higher response, or low cost methods like phone or mail, these results would indicate that based on the scope criterion, mail may be the preferred mode. This frees the researcher to base the decision more on the needed statistical benefits of the different modes instead.

This brings us to examine the recommendations of Smith (1996) which is that one of the three critical tests of CV studies is that they pass a scope test. While it would be ideal that respondents respond with a WTP that increases with as quantity or quality of resource, stated preferences can be hard to estimate. It is possible that respondents are willing to pay for the first unit, and they either see the following units as joint-production goods (for example, preserving the first butterfly species will also help preserve the next 4 butterfly species). It is also possible that there is poor survey design. Future research might be able to discover whether poor survey design is the culprit by examining the grey literature and utilizing this literature as a proxy for "bad survey design". However, even within the realm of published literature with "good survey design" that utilized the Dillman Method or in-person survey, scope insensitivity is a possibility.

It may be that the sensitivity to scope just depends. Scope sensitivity is not a sufficient condition for a completely valid CV study, so the reason why the converse would be true should be suspect as well. Thus our understanding of what are the key factors in designing a CVM survey to reflect scope requires more research.

WORKS CITED

Boyle, Kevin J. "Contingent Valuation in Practice." In *A Primer on Nonmarket Valuation*, by Patricia A. Champ, Kevin J. Boyle and Thomas C. Brown, 111-116. Norwell: Kluwer Academic Publishers, 2003.

Boyle, Kevin J., William F. Desvouges, F. Reed Johnson, Richard W. Dunford, and Sara P. Hudson. "An Investigation of Part - Whole Biases in Contingent-Valuation Studies." *Journal of Environmental Econoimcs and Management*, 1994: 64-83.

Brown, Thomas C. "Introduction to Stated Prefernce Methods." In *A Primer on Nonmarket Valuation*, by Patricia A. Champ, Kevin J. Boyle and Thomas C. Brown, 99-110. Norwell: Kluwer Academic Press, 2003.

Brown, Thomas C., and John W. Duffield. "Testing part-whole valuation effects in contingent valuation of instream flow protection." *Water Resources Research*, 1995: 2341-2351.

Carson, Richard. "Contingent Valuation Surveys and Tests of Insensitivity to Scope." In *Determining the Value of Non-Marketed Goods: Economic Psycholgical, and Policy Relevant Aspects of Contingent Valuation Methods*, by R.J. Koop, W. Pommerhene and N. Schwartz. Boston: Kluwer, 1997.

Carson, Richard T., and Robert C. Mitchell. "The issue of scope in contingent valuation studies." *American Journal of Agricultural Economics*, 1993: 1263-1267.

Desvousges, William H., F. Reed Johnson, Richard W. Dunford, Kevin J. Boyle, and K. Nicole Wilson. "Measuring Natural Resource Damages with contingent valuation: tests of validity and reliablity 2nd ed." In *Contingent Valuation: A Critical Assesment*, by Jenny A. Hausman, 91-159. Amsterdam: North-Holland, 1993.

Haab, Timothy C, and Kenneth E McConnell. "Valuing Environmental and Natural Resources." Northampton: Edward Elgar Publishing, Inc., 2002.

Heberlein, Thomas A., Matthew A. Wilson, Richard C. Bishop, and Nora Cate Schaeffer. "Rethinking the scope test as a criterion for validity in contingent valuation." *Journal of Environmental Economics and Management*, 2005: 1-22.

Hunt, Morton. "How Science Takes Stock." New York: Russell Sage Foundation, 1997.

National Oceanic and Atmospheric Administration (NOAA). "Natural Resource Damange Assesment Under the Oil Pollution Act of 1990." Federal Government, Federal Register, 1993.

Nelson, Jon P., and Peter E. Kennedy. "The Use (and Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assesment." *Environmental & Resource Economics*, 2009: 345-377.

Ojea, Elena, and Maria L. Loureiro. "Valuation of Wildlife: Revising Some Additional Considerations For Scope Tests." *Contemporary Economic Policy*, 2009: 236-250.

Pouta, Eija. "Sensitivity to scope of environmental regulation in contingent valuation of forest cutting practices in Finland." *Forest Policy and Economics*, 2005: 539-550.

Schulze, William D., Gary H. McClelland, Jeffery K. Lazo, and Robert D. Rowe. "Embedding and calibration in measuring non-use values." *Resource and Energy Economics*, 1998: 163-178.

Smith, V. Kerry. "Can Contingent Valuation Distinguish Economic Values for Different Public Goods?" *Land Economics*, 1996: 139-151.

Stanley, T.D. "Wheat from Chaff: Meta-Analysis as Quantitative Literature Review." *The Journal of Economic Perspectives*, 2001: 131-150.

Veisten, Knut, Hans Fredrik Hoen, Stale Navrud, and Jon Strand. "Scope insensitivity in contingent valuation of complex environmental amenities." *Journal of Environmental Management*, 2004: 317-331.

CHAPTER SIX: Closing Thoughts

To assist the manager in weighing different interests, it is often needed to have some sort of measure of value. When a quantitative measure of value can be incorporated into a benefit coast analysis it becomes much simpler for the manager to allocate scare resources to where they can make the most impact.

Recognizing that the price signal in markets is quite powerful, simulating that signal in a nonmarket setting also has great potential. How the nonmarket data is gained and organized provides the lens by which the values are examined. The work here demonstrated that there are ways of gaining a more efficient estimator for a relatively low cost to the researcher. It has also shown in examining the scope issue, that there are multiple issues that a CVM study needs to clear, and scope is perhaps not the best litmus test.

In the second chapter we saw that a combination of TCM with a DC CVM question increased the efficiency for the estimation of the cost estimate. Recognizing that often one of the most difficult parts of a useful study is getting people to respond to a survey, a single question tacked onto a travel cost survey is both feasible, and not overly burdensome on the respondent. In seeking to provide useful information for policy makers, the researcher needs to consider ways to increase the amount of data, while being mindful of the burden being placed on the respondent. Further, a simple follow-up question as was used in this instance provides a less hypothetical scenario, as it only changes a single attribute. While choice experiments are quite powerful, they still need to be used in a relevant context.

The third chapter explored the difference in respondents to the first and second bid amounts for a follow-up bid amount in a CVM study. It was found that there is a shift from the first to second bid amounts. This is perhaps intrinsic to the model. The respondent is learning and updating his information as it is presented, trying to determine the “correct” price. The unexpected result is that onsite and in-person did not alleviate this result which is often observed in household surveys. The gain in efficiency is still present and quite useful; this is not surprising as more information is provided.

The fourth chapter examining the scope effect in meta-analysis found inconclusive results. There was no single variable that indicated a study passing or failing across models. By limiting the meta-analysis to the studies that have passed the peer-reviewed standard, it seems that scope issues may receive more ink than is warranted.

Primarily this work has shown that increasing the amount of information in a study does increase the efficiency of the estimates, and a single extra question or a combining of SP and RP data can provide that needed information. The researcher seeking to improve the quality of their CVM study could be more productive gaining additional information from respondents rather than being overly concerned with scope issues in a CVM study.

APPENDIX A. List of included papers in Meta Regression

Author	Title	Year	Journal
Bowker & Didychuk	Estimation of the Nonmarket Benefits of Agricultural Land Retention in Eastern Canada	1994	Agricultural and Resource Economics Review
Boyle, Welsh, & Bishop	Starting Point Bias in Contingent Valuation Bidding Games	1985	Land Economics
Boyle, Desvousges, Johnson, Dunford, Hudson	An Investigation of Part-Whole Biases in Contingent-Valuation Studies	1994	Journal of Environmental Economics and Management
Brown, Layton, & Lazo	Valuing Habitat and Endangered Species	1994	Institute for Economic Research Discussion Paper
Buzby, Ready, Hu	Differences between Continuous and Discrete Contingent Valuation Estimates	1996	Land Economics
Bloomquist Whitehead	Resource Quality information and Validity of Willingness to Pay in Contingent Valuation	1998	Resource and Energy Economics
Eom & Larson	Improving environmental valuation estimates through consistent use of revealed and stated preference information	2006	Journal of Environmental Economics and Management
Berrens, Ganderton, Silva	Valuing the Protection of Minimal Instream Flows in New Mexico	1996	Journal of Agricultural and Resource Economics
Brown & Duffield	Testing Part-Whole Valuation Effects in Contingent Valuation of Instream Flow Protection	1995	Waters Resources Research
Schakde, Payne	How People Respond to Contingent Valuation Questions: A Verbal Protocol Analysis of Willingness to Pay for and Environmental Regulation	1994	Journal of Environmental Economics and Management

Walsh Loomis Gillman	Valuing Option, Existence, and Bequest Demand for Wilderness	1984	Land Economics
Streever, Callaghan-Perry, Searles, Stevens, Svoboda	Public Attitudes and Values for Wetland Conservation in New South Wales, Australia	1998	Journal of Environmental Management
Poe Giraud Loomis	Computational Methods for Measuring the Difference of Empirical Distributions	2005	American Journal of Agricultural Economics
Whitehead & Finney	Willingness to Pay for a Green Energy Program: A Comparison of Ex-Ante and Ex-post Hypothetical Bias Mitigation Approaches	2003	Journal of Cultural Economics
Pouta	Sensitivity to Scope of Environmental Regulation in Contingent Valuation of Forest Cutting Practices in Finland	2005	Forest Policy and Economics
Veisten, Hoen, Navurd, Strand	Scope Insensitivity in Contingent Valuation of Complex Environmental Amenities	2004	Journal of Environmental Management
Svedaster	Contingent Valuation of Global Environmental Resources: Test of Prefect and Regular Embedding	2000	Journal of Economic Psychology
Whitehead, Cherry	Willingness to pay for a Green Energy program: A Comparison of ex-ante and ex-post hypothetical bias mitigation approaches	2007	Resource and Energy Economics
Berrens, Bohara, Silva, Brookshire, McKee	Contingent Valuation Values for New Mexico Instream Flows: With tests of Scope, Group-Size Reminder and Temporal Reliability	2000	Journal of Environmental Management
Hite Hudson Intrapapong	Willingness to Pay for Water Quality Improvements: The Case of Precision Application Technology	2002	Journal of Agricultural and Resource Economics
Giraud Loomis Johnson	Internal and External Scope in Willingness-to-Pay Estimates for Threatened and Endangered Wildlife	1999	Journal of Environmental Management
Ready Berger Blomquist	Measuring Amenity Benefits from Farmland: Hedonic Pricing vs. Contingent Valuation	1997	Growth and Change

Nunes Schokkaert	Identifying the Warm Glow Effect in Contingent Valuation	2003	Journal of Environmental Economics and Management
Smith, Kemp, Savage, Taylor	Methods and Results from a New Survey of Values for Eastern Regional Haze Improvements	2005	Journal of Air and Waste Management Association
Whitehead	Ex Ante Willingness to Pay with Supply and Demand Uncertainty: Implications for valuing a Sea Turtle Protection Programme	1992	Applied Economics
Stevens Decoteau Willis	Sensitivity of Contingent Valuation to Alternative Payment Schedules	1997	Land Economics
Stanely	Local Perception of Public Goods: Recent Assessments of Willingness-to-Pay for Endangered Species	2005	Contemporary Economic Policy
Whitehead, Groothuis, Southwick, Foster-Turley	Measuring the Economic Benefits of Saginaw Bay Coastal Marsh with Revealed and Stated Preference Methods	2009	Journal of Great Lakes Research
Douglas & Taylor	The Economic Value of Trinity River Water	1999	Water Resource Development
Cooper Poe Bateman	The Structure of Motivation for Contingent Values: A Case Study of Lake Water Quality Improvement	2004	Ecological Economics
McDaniels, Gregory, Arvai, Cheunpagdee	Decision Structuring to Alleviate Embedding in Environmental Valuation	2003	Ecological Economics
Hanemann Loomis Kanninen	Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation	1991	American Journal of Agricultural Economics
Choe Whittington Lauri	The Economic Benefits of Surface Water Quality Improvements in Developing Countries: A Case Study of Davao, Philippines	1996	Land Economics
Dupont	CVM Embedding Effects When There are Active, Potentially Active and Passive Users of Environmental Goods	2003	Environmental and Resource Economics

Hoevengel	The Validity of the Contingent Valuation Method: Perfect and Regular Embedding	1996	Environmental and Resource Economics
Rollins & Lyke	The Case for Diminishing Marginal Existence Values	1998	Journal of Environmental Economics and Management
Heberlin, Wilson, Bishop, Schaeffer	Rethinking the Scope Test as a Criterion for Validity in Contingent Valuation	2005	Journal of Environmental Economics and Management
Veisten, Hoen, Navrud, Strand	Scope Insensitivity in Contingent Valuation of Complex Environmental Amenities	2004	Journal of Environmental Management
White, Gregory, Lindley, Richards	Economic Values of Threatened Mamals in Britian: A Case Study of the Otter *Lutra Lutra* and the Water Vole *Arvicola Terrestris*	1997	Biological Conservation