Quasi-Experimental Evidence in Valuing Nonmarket/Public Goods: Consistency and Behavior of Intertemporal Bids in Contingent Valuation Techniques

by

Douglas D. Ofiara

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1998

UMI Number: 9830747

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

This manuscript has been read and accepted for the graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

4/20/98 Michael Grossman

Date — Chair of Examining Committee

4/20/98 Linda N. Edwards

Date — Executive Officer

Michael Grossman (CUNY-GSUC)

Theodore Joyce (Bernard Baruch College)

Robert Kaestner (Bernard Baruch College)

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Quasi-Experimental Evidence in Valuing Nonmarket/Public Goods: Consistency and Behavior of Intertemporal Bids in Contingent Valuation Techniques

by

Douglas D. Ofiara

Advisor: Michael Grossman

The purpose of this investigation is to examine the behavior of CV (Contingent Valuation) bids over multiple time periods through repeated surveys (i.e., repeated games). Standard consumer choice models assume that consumers make informed, rational decisions in arriving at their mix of purchased goods and services subject to their household budget constraint. This decision process, a major premise of economic theory, is based on consumer experience and information acquired over a period of time. If this fundamental premise holds in contingent markets, and for the CV method, then, it is reasonable to expect that economic values (i.e., bids) produced from contingent markets can change over time or bidding periods. Some researchers in fact wonder if bids based on contingent markets will converge to some "true" value over time.

This the first study to rigorously examine whether individuals participating in contingent markets via CV surveys revise their bids if given a long time horizon, and whether individuals' behavior in contingent markets corresponds to that in standard consumer choice theory by using a variety of methods from statistical inference to econometrics and time series techniques. It is also the first study to examine if experimental economics techniques can be used in real-life settings (e.g., experiments of active users of a particular resource/activity based on repeated games), and if the resulting evidence can provide useful information in understanding agents' behavior in contingent markets, and in future advancements of the CV method. Evidence based on quasi-experimental data (repeated games of sport fishermen over a season involving up to 5 repeated surveys/games each corresponding to a month, yielding 5 bidding periods) suggests that experienced, active recreational participants do revise their CV bids over multiple time periods. This could be

due to their familiarity with the CV game over time, changes in their tastes and in repeated thinking about their true preferences, and in their success or failure in a particular month. Such findings suggest profound effects on future work in CV methods and in experimental methods that elicit preferences for nonmarket/public type goods.

Acknowledgements

A number of individuals assisted in making this dissertation a reality. First I want to thank my wife, Robin Merle for her belief, understanding and patience; without either a failed attempt would have resulted. I also thank my son and best buddy, Jonathan who missed his daddy many nights and weekends during this effort. A great deal of gratitude is extended to Michael Grossman who gave me a chance at CUNY and who assisted in my progress and in the dissertation; to the faculty in the Graduate school at CUNY, to Charles de Bartolome at NYU who taught me Public Economics one of my fields, and to my fellow graduate students among them Leonid Golubchin with whom I spent much time preparing for exams.

Contents

Page

Contents, Continued

List of Tables

Page

List of Tables, Continued

Page

List of Figures

Chapter 1 - Outline of Investigation

1.1. Introduction

A variety of goods possess characteristics similar to public goods, namely that one cannot prevent others from using the good, any individuals' consumption does not reduce the quantity available for others, and the good is provided at nominal or zero marginal cost, i.e., the characteristics of nonrivalry, nonexcludability, and public good pricing. Examples of such goods are public education, fire protection, clean air and water, and recreational activities involving consumptive activities (e.g., fishing, hunting) and nonconsumptive activities (bird watching, boating). Such environmental and recreational goods and services have been referred to as nonmarket goods. Both public goods and nonmarket goods share an additional characteristic in that these goods are not sold in any market setting creating a problem regarding their relative economic importance and value. Without price and quantity data normally used to assess the economic benefits and value of marketed goods using welfare techniques, one is left without techniques and approaches to assess public and nonmarket goods.

Economists in the areas of public economics and in experimental economics as well as in the applied fields of agricultural and resource economics have made great strides in the past three decades in their ability to assess monetary values for resources, activities, and environmental degradation that are normally not traded in formal markets (i.e., nonmarket goods) and/or share public good attributes. This is partly due to advances in theoretical and applied research concerning welfare measures (Freeman 1979, 1993, Johansson 1987, Randall and Stoll 1980, Willig 1976), partly due to the development and refinement of techniques that can be used to assess public/nonmarket goods (Cummings *et al.* 1986a, 1986b, Davis and Holt 1993, Freeman 1979, 1993, Mitchell and Carson 1989, Mueller 1989, Myles 1995, Smith 1979), and partly due to advances in experimental economic techniques (Davis and Holt 1993, Smith 1979, 1980, 1982, 1986). For example, policies designed to manage fishery resources now usually include a recreational component and the value associated with recreational fishing in decisions about catch allocation. Damages to recreational activities attributable to air and water pollution are now factored in environmental policy decisions. And damages to species not commercially important from spills of oil and hazardous substances have been recently considered in compensation litigation. Were it not for the effort involved in developing nonmarket valuation techniques many nonmarket goods and activities, such as sport fishing and species not commercially important, might still be zero-valued in public policy decisions.

Nonmarket good valuation techniques can be classified as direct or indirect methods and are based on observed market behavior, unobserved market behavior, or hypothetical behavior; the travel cost method and hedonic price approach are examples of direct approaches, the contingent valuation (CV) method is an indirect approach (Cummings *et al.* 1986a, Freeman 1979, 1993, Johansson 1993, Mitchell and Carson 1989). Although these techniques have been around for some time and have been continually refined, the CV method, remains controversial. This controversy is associated with the following characteristics of the CV method: 1) the method assesses values based on hypothetical data/responses, 2) the method obtains ex ante welfare measures rather than ex post welfare measures, and 3) the CV method is claimed to be the only method available to obtain nonuse values (Cummings *et al.* 1986b, Diamond and Hausman 1994, Freeman 1993, Hausman 1993, Mitchell and Carson 1989, U.S. DOC, NOAA 1993). In attempts to gain credibility for the CV method among peers, numerous research was conducted to examine limitations of the CV method, and comparisons with alternative valuation methods (Cummings *et al.* 1986b, Mitchell and Carson 1989). In addition, several conferences and workshops were convened to critically assess the current state of the CV method.

Researchers now acknowledge that there were faults with early comparison studies (Mitchell and Carson 1989). Comparison studies were developed to provide support of the validity and hence, credibility of the CV method. This early research proceeded on the notion that estimates of economic value from alternative nonmarket valuation methods were comparable and theoretically equal. It was reasoned that if estimates of economic value from the CV method were similar (statistically different) to estimates from alternative nonmarket valuation methods, this then provided evidence to support the validity of the CV method. Researchers subsequently learned that none of the alternative nonmarket valuation methods results in a superior measure of the theoretical construct (monetary value of preferences), and that researchers were often confused with the concept of validity. If one nonmarket valuation method resulted in a superior measure of the construct, then a basis for comparisons would exist, and one could examine how close alternative nonmarket valuation methods would be to the method that results in a superior measure. In early comparison studies, correspondence between CV methods and alternative valuation methods was interpreted as evidence of validity in the truest sense, criterion validity (i.e., whereby measures of economic value from alternative nonmarket valuation methods are neither a truer measure of the theoretical construct (monetary value of preferences) than one another), with that of convergent validity (i.e., correspondence between economic value estimates from alternative nonmarket valuation methods), a weak form of validity. Hence, one can only establish credibility of the CV method in the sense of convergent validity. This realization weakens efforts to establish credibility of the CV method via comparison studies.

The U.S. Environmental Protection Agency (EPA) commissioned a state-of-the-art assessment and conference of the CV method in 1983-84; the overall conclusion reached was that the CV method shows promise, but faces some real challenges (Cummings *et al.* 1986b). In 1992, prompted by the *Exxon Valdez* oil spill and the Oil Pollution Act of 1990, the National Oceanic and Atmospheric Administration (NOAA) was directed to assemble an expert panel to provide advice to NOAA and conduct a current assessment of the CV method with particular attention to nonuse values (specifically to assess whether the CV method is capable of providing estimates of lost nonuse values "that are reliable enough" for natural resource damage assessments [Portney 1994: 8]). The NOAA panel concluded that the CV method "can produce estimates reliable enough to be the starting point of a judicial process of damage assessment, including lost passive nonuse values (Ibid: 8, U.S. DOC, NOAA 1993). In addition, Exxon convened its' own conference and team of economists to provide a critical assessment of the CV method in 1992, with the conclusion that the CV method should not be used to assess environmental damages or in Cost-Benefit Analyses (Hausman 1993, Diamond and Hausman 1994). Then in 1994, the Department of Energy and U.S. EPA co-sponsored a workshop on the CV method to address its status on assessments of nonuse values and to develop a research agenda for future CV research (Bjornstad and Kahn 1996).

In recent times the *Exxon Valdez* oil spill, federal rules and procedures pertaining to Natural Resource Damage Assessments via CERCLA (Comprehensive Environmental Response, Compensation, and Liability Act of 1980 and 1986) and the Oil Pollution Act of 1990 have reopened the status of the CV method and created added controversies among the public, environmentalists, and economists (see Ofiara and Seneca 1998 for a summary of federal rulemaking). Legal proceedings concerning economic assessments of the environmental damage challenged economists to critically evaluate the CV method, and at the same time resulted in a split in the profession concerning the use and credibility of the CV method and its' value estimates (see Hausman 1993, Portney 1994, Diamond and Hausman 1994, 1993, Carson *et al.* 1996b, 1996c). The current debate involves two polar views; those that advocate the CV method and those that do not. While most of the criticism involves using the CV method for nonuse values, Diamond and Hausman do not recommend the CV method for use in Cost-Benefit Analysis, hence, the CV method is also suspect in estimating use values (see Hausman 1993).

The above progress, concerns, and controversy suggest the need for further experimentation and refinement of the CV method to provide a more rigorous and thorough assessment of the underlying principles of the CV method and properties of resulting CV bids. Although field research can contribute to this effort, it is felt that more insight could be gained from experimental economic techniques. These techniques could assist in efforts to explore the behavior of contingent markets, CV bids, and the CV method in general in order to refine and advance the CV method. As Mitchell and Carson (1989) note, most CV research has involved field applications (usually involving a one-time survey, and sometimes based on statistical experimental designs, e.g., split samples testing different payment vehicles or information) versus experiments conducted in laboratory-type settings as in experimental economic studies.

1.2. Introduction to the Contingent Valuation Literature

1.2.1. Overview of the Literature

The roots of the contingent valuation approach (CVA) can be traced to public choice theory (Mueller 1979, Freeman 1979). One aspect of public choice theory deals with the issue of what levels of public goods should be provided in order to achieve a socially acceptable criteria. Because public goods are not produced and sold in a market, researchers faced a problem in terms of evaluating present and proposed levels of public goods. As a result, a number of methods were developed to induce individuals to reveal their preferences towards public goods, (i.e., preference revelation). Freeman (1979) describes three basic approaches: 1) voting on referendums, 2) revealed preferences of the quantity of public goods, and 3) revealed preferences of the money value of public goods. Of these, the CVA falls under the third case. It involves directly asking individuals to state their willingness–to–pay (WTP) for a given amount, or given change in the amount of the public good.

The CVA has been applied to assess the benefits of preserving and improving air quality and water quality, and in valuing recreational activities such as sport fishing, hunting, visits to natural areas, and in assessing values for natural environments and regional estuaries such as the Chesapeake Bay. An extensive literature dating to 1974 exists and is reviewed in Mitchell and Carson (1989), other reviews are in Cummings *et al.* (1986b), Hausman (1993), Bjornstad and Kahn (1996), Willis and Corkindale (1996).

The CVA has evolved into a much more sophisticated technique, however, than simply asking WTP questions. The first level of CV studies basically asked WTP questions, e.g., what is the maximum dollar amount you would be willing to pay for a specific good?

(Bohm 1972, Davis 1963, 1964, Hammack and Brown 1974, Knetsch and Davis 1966, McConnell 1977). The next level of CV studies began a series of methodological developments concerning various aspects of the CVA, for example variations in question formats (iterative bidding questions, closed-ended question, use of payment card), more descriptive definitions of the hypothetical good (involving photographs), variations in payment vehicles (donations, taxes, utility fees), variations in information, variations in sequencing of questions and information (Brookshire *et al.* 1976, 1979, 1981, 1982, Loehman 1984, Loehman *et al.* 1981, 1982, Mitchell and Carson 1984, Randall *et al.* 1974. 1978, 1983, Rowe *et al.* 1980, Schulze *et al.* 1981, Schulze and Brookshire 1983, Thayer 1981).

Randall *et al.* (1974) is credited with advancing the iterative bidding question and Brookshire *et al.* (1976), Rowe *et al.* (1980) were early contributors to this effort. The work of Hanemann (1983, 1984a. 1984b, 1985) is credited with advancing the use of the closed-ended question format and use of referendum type CV models. The combined work of Brookshire *et al.* (1976, 1979, 1981. 1982), Loehman (1984), Loehman *et al.* (1981, 1982), Mitchell and Carson (1981, 1984) Randall *et al.* (1978. 1983), Rowe *et al.* (1980), Schulze *et al.* (1981), Schulze and Brookshire (1983) and Thayer (1981) has contributed to refinements in descriptions of the hypothetical good. payment vehicle, and relative importance of inherent biases associated with the CVA.

Another level of CV studies examined and compared the CVA with other nonmarket valuation approaches and with experimental economic techniques and assessment methods (Bishop and Heberlein 1979, 1986, Bishop *et al.* 1983, Brookshire *et al.* 1979. 1982, Coursey *et al.* 1987, Coursey and Schulze 1986, Desvousges *et al.* 1983. 1988. Knetsch and Sinden 1984, Loehman 1984. Schulze *et al.* 1981. Seller *et al.* 1985). The experimental economic studies involving CV are useful to note because some found evidence of convergence between WTP and WTA (Coursey *et al.* 1987), and that the CVA compared favorably with alternative methods (Brookshire *et al.* 1979, 1981, 1982, Cummings *et al.* 1986b. Randall *et al.* 1983. Schulze *et al.* 1981. Seller *et al.* 1985).

The present level of CV studies involves numerous comparisons, tests, refinements to reexamine many of the earlier concerns, biases, and subsequent concerns (Adamowicz *et al.* 1993, Boyle *et al.* 1994, 1996, Carson *et al.* 1996a, Cameron 1992, Desvousges *et al.* 1993, 1996, Diamond *et al.* 1993, Randall and Hoehn 1996, Ready *et al.* 1996, Schulze *et al.* 1996, Smith and Osborne 1996, Whitehead *et al.* 1995), to rebut claims that the CV method has no basis in measuring nonuse values or in CBA (Diamond and Hausman 1993, 1994), and to rebut some recommendations from a NOAA Panel on CVA (U.S. DOC, NOAA 1993) (see Carson *et al.* 1996b, 1996c, Bjornstad and Kahn 1996. Willis and Corkindale 1995 for recent summaries).

The major assumption of the CVA is that individuals *can* reveal their willingness–to–pay for nonmarket goods contingent on a hypothetical market transaction (Freeman 1979, 1993). The mechanics of the approach are as follows. The CVA is based on the premise of a realistically designed, though hypothetical market setting. An individual is asked to reveal his/her preference via a survey in the form of a bid (e.g., maximum amount willing to pay, minimum amount willing to accept) contingent on the availability of the good in question. Thus, it is as if an individual faces an experimental market and reacts in terms of consumption behavior to the changes specified by the researcher.

The basic types of valuation questions that have been used and currently are in use are referred to as open-ended CV questions, closed-ended CV questions, and iterative bidding CV questions. In open-ended CV questions the individual is asked for their maximum WTP and/or minimum WTA (willingness–to–accept) associated with a specific level of the hypothetical good. Iterative bidding CV questions start from a preset price and ask for these values (i.e., max WTP or min WTA) corresponding to a specific level of the good in an iterative manner. If the individual responds yes to the initial price, the price is increased by some increment and the question repeated. This process is repeated in an iterative manner until a no response is reached and the process is stopped. The last yes-response is taken to be the max-WTP or min-WTA amount. Closed-ended questions (or referendum or discrete CV questions as they are also known as) were first based on asking individuals to reveal their preference by voting (yes-no) for a specific level of the good if available at a predetermined price (Hanemann 1984a). Later versions involved splitting the sample into several equal parts, assigning a different random price for each group and then asking the individual to vote (yes-no) for a specific level of the good at the stated price, where all individuals faced the same quantity of the good but with different prices. This was an attempt to introduce price variation into closed-ended or referendum formats. The closed-ended or referendum data is evaluated using a number of discrete-choice models because the responses follow binary data (1,0 for yes,no responses). The work of Hanemann (1983, 1984a, 1984b, 1985) developed procedures for estimating welfare measures from discrete-choice models and was a major factor in the popularity of the referendum CV method. Issues that remain unresolved concern the superiority of open-ended question formats versus closed-ended question formats (Boyle *et al.* 1996. Loomis 1990. Willis 1995).

Advantages of the CVA are that it is the most versatile of the nonmarket valuation techniques in assessing the value of a wide range of goods. It allows for flexibility in the good to be valued. and it can be used to assess nonuse values (i.e.. existence values). The CVA has been used in a wide range of applications. for example it has been used to value improvements in environmental quality (clean air. clean water). aesthetic goods (visibility. scenic views), recreational activities and experiences (sportfishing. recreational beach use. boating), protection of endangered species. preservation of unique habitats and environments. public pest control. and health–risk issues (Cummings *et al.* 1986a. Mitchell and Carson 1989. Schulze *et al.* 1981, Bjornstad and Kahn 1996). In contrast. other nonmarket good valuation techniques, the Travel Cost Approach (TCA) and Household Production Approach (HPA) have been confined to value recreational activities and experiences. while the HPA has been used chiefly to value changes in environmental quality. and housing characteristics. The CVA then. is the most flexible of the techniques in its application and the only technique capable of assessing nonuse values such as existence values.

Disadvantages and shortcomings of the CVA regard its basic premise and the difficulties of obtaining valid responses. The major weakness of the approach, as previously mentioned, is whether values based on expected behavior can (ever) accurately reflect true behavior. (That is, can ex–ante values represent values of actual behavior — ex–post values?) This concern involves two separate issues — strategic bias and hypothetical bias. The first involves whether respondents will try to influence the outcome of survey results by not stating their "true" preferences. Individuals can understate or overstate their "true" preferences if they believe the response they give will actually influence the price they will have to pay, or if they believe whatever their response is they give will not have to pay for the good. The second concern, hypothetical bias results from the contention that because willingness–to–pay or willingness–to–accept amounts are not actually paid, but are hypothetical, individuals lack incentives to determine and reveal their "true" preferences. Research has shown that for both of these possible biases, if care is taken in designing the market, the good, and valuation questions so the survey is realistic and credible, these biases can be minimized. This will increase the likelihood that the revealed preferences obtained via such surveys will approach individuals' "true" preferences (Cummings *et al.* 1986b, Mitchell and Carson 1989, Randall *et al.* 1983, Schulze *et al.* 1981).

Other shortcomings of the CVA concern questions related to aspects of the survey design such as the framing of questions, payment mode, information, starting bids (i.e., initial preset prices), and interviewer biases. Again, research has shown that well designed surveys can minimize the effects of these problems on survey results (Cummings *et al.* 1986b, Mitchell and Carson 1989, Randall *et al.* 1983, Schulze *et al.* 1981).

In response to the major concern of the hypothetical nature of the CVA, research has compared the economic value estimates obtained by the CVA with those obtained from the indirect valuation methods, namely the TCA and HPA. Evidence shows that under appropriate applications the CVA produces economic values that compare well with the economic values obtained from indirect valuation methods (see Cummings *et al.* 1986b for a thorough discussion). However, as Mitchell and Carson (1989) realized in their assessment of the CVA, early comparison studies were at fault and researchers often confused various validity criteria measures. Present research that involves comparisons and tests of validity attempt to address the shortcomings of previous research. In addition, present concerns with the CVA involve whether revealed bids will vary over time given experience and familiarity with the CVA and public good in question.

1.2.2. CV Literature: Implications for Convergence over Time - Intertemporal Consistency

Research that has examined CV bids over time has been in the context of what is known as "test-retest" criteria and more recently "temporal reliability." The approaches are similar in that the former involves administering the same questionnaire to the same group of individuals at two different time periods (Mitchell and Carson 1989), and the latter to a different group of individuals (Reiling *et al.* 1990, Carson *et al.* 1995). In both approaches the difference between responses is claimed to be attributable to the true variance plus noise from the questionnaire, although the latter test appears to be a more valid test of the survey instrument.

Loomis (1989) conducted the first study that used the test-retest approach on the same group of individuals that spanned a one year period. The first survey was administered during April-May 1986, the second survey January-March 1987. Results found that the direction of change in mean CV bids over time between those surveyed (general population versus users) were contradictory; for the former group mean CV bids increased, while for users mean CV bids decreased over time. WTP bid equations were estimated over the two periods of the form $WTP_{t+1} = \beta_0 + \beta_1(WTP_t)$, essentially an AR(1) process. Results were consistent with convergent behavior for the group of users (i.e., $\beta_1<1$; estimated coefficients ranged from 0.502 to 0.6769), and with random walk behavior for the general population group (i.e., $\beta_1=1$: estimated coefficients ranged from 0.801 to 1.004). A further issue from this study is if groups of users of a particular resource provide different and more accurate expressions of monetary values than do groups of the general population.

Stevens *et al.* (1994) followed the test-retest approach of Loomis in a study of existence values for wildlife over two points in time that spanned a three year period (spring 1989 - 1992). They found that the average CV bid increased over time. An AR(1) model was also estimated and findings were consistent with convergent behavior (i.e., $\beta_1 < 1$: estimated coefficient was 0.61).

Other research has examined the temporal reliability version of the test, but cannot offer any evidence about CV bid revision behavior because different individuals were surveyed over time, and AR(1) type equations were not estimated (Reiling *et al.* 1990, Carson *et al.* 1995). Reiling *et al.* (1990) examined changes in WTP bids from two surveys separated by a two month period and found that mean WTP bids increased. Carson *et al.* (1995) administered a survey over two points in time that spanned a two year period. Findings based on one set of questions showed no significant change in mean WTP ($52.80 vs. $52.81), and based on a series of questions that involved a range of stated prices again showed no significant change ($54.23 vs. $54.02). However, each of these studies relied on samples of different individuals over the two time periods and little can be concluded about CV bid revision behavior. In general, it would appear that studies based on temporal reliability have little to offer regarding CV bid behavior over time.

Early studies that have examined temporal stability via the test-retest approach were based on correlation statistics and did not indicate the direction of change between periods (Jones-Lee *et al.* 1985, Loehman and Du 1982). Jones-Lee *et al.* (1985) examined the value of safety and administered the survey to a subset of the same respondents one month after the original survey. Results based on a matched-pair, signed-rank test did not indicate a significant difference over both time periods. Loehman and De (1982) readministered their mail survey to the same individuals (a small sample as a test, N=47) after a three week period. Correlations estimated between both periods ranged from 0.82 to 0.95; because correlations were published with positive signs it is assumed that bids increased over time.

Lastly, in another study (Kealy *et al.* 1988, 1990) conducted a resurvey of students after a two week period concerning a controlled-laboratory experiment involving a private good (Cadbury candy bar) and a public good (deacidification program for lakes in the New York Adirondacks). Two separate analyses examined "stability" over time. The first analysis involved the private good only and found that both open-ended and referendum question formats resulted in no significant difference in paired-differences over the two time periods, suggesting that responses were stable over time (Kealy *et al.* 1988). The second analysis involved a test of the equality of the regression model (estimated probit model) over both time periods and examination of the correlation over time for both goods. Findings showed that the probit models were not significantly different over time for both goods. Results of the second test yielded correlations of 0.70 for the private good and 0.66 for the public good, and the researchers concluded that the results indicated a "high level of reliability" but that the nature of the good did not matter (Kealy *et al.* 1990: 256).

In sum, the inherent design of the test-retest approach can only yield limited information about the process of CV bid revisions over time, and general conclusions based on these types of studies cannot be made. The test-retest approach implicitly assumes all other factors such as tastes and experience with the contingent market remain constant over time, an assumption that is at odds with the standard consumer choice theory. Another problem is that only two time periods are considered and it generally takes more than three data points to produce some evidence of trend. Hence, it remains unknown if individuals would continue to revise their CV bids given a longer time horizon as evident in laboratory-type experiments, and if this process of revision would yield an individual's "true" preferences.

1.2.3. CV Literature: Implications for Differences Between WTP and WTA Measures

An observation from early applied research has been that a large and significant difference between WTP and WTA measures exist. A number of studies have replicated this finding (Adamowicz *et al.* 1993, Banford *et al.* 1977, Bishop and Heberlein 1979, Coursey *et al.* 1987, Coursey and Schulze 1986, Cummings *et al.* 1986b, Gordon and Knetsch 1979, Hammack and Brown 1974, Knetsch 1989, Knetsch and Sinden 1984, Mitchell and Carson 1989, Schulze *et al.* 1981, Shogren *et al.* 1994, Sinclair 1976). Both field-applications and laboratory experiments have confirmed this result (Coursey *et al.* 1987, Mitchell and Carson 1989, Schulze *et al.* 1981, Shogren *et al.* 1994). This result has in turn become a concern among researchers and with the CVA, because standard economic theory predicts that WTP and WTA measures should coincide given small income effects; with zero income effects ordinary (Marshallian) demand curves and compensated (Hicksian) demand curves will coincide (Just *et al.* 1982, Layard and Walters 1978, Randall and Stoll 1980, Willig 1976).

From early research findings of empirical discrepancies between WTP and WTA indicated WTA was from 3 to 5 times larger than WTP (see Cummings *et al.* 1986b for a summary). Hammack and Brown (1974) first showed empirical differences between mean values of WTP and WTA of a four fold difference for duck hunting (WTP-S247 vs. WTA-S1044). Banford *et al.* (1977) found differences in means of almost three times for recreational fishing and over four times for postal services (WTP-S43 vs. WTA-S120; WTP-S22 vs. WTA-S93). Another early study found that median values differed by a factor of 20 (WTP-S35 vs. WTA-S700) for recreational fishing (Sinclair 1976). Brookshire *et al.* (1980) in a study of elk hunting found a difference of 2.6 times between mean values (WTP-S54 vs. WTA-S142). Bishop and Heberlein (1979) and Bishop *et al.* (1983) in a study of goose hunting found a difference of almost five times (WTP-S21 vs. WTA-S101).

Implications of such differences have caused some to recommend that WTP estimates be adjusted upwards or that WTA be adjusted downwards, and that field applications should concentrate on WTP measures even though in some cases WTA measures are theoretically correct. Following these early studies that noted differences, advancements in research considered techniques that could either account for such differences or could test whether WTP and WTA measures might converge given repeated CV trials. One of the first tests of convergence involved an experimental economic technique used by Coursey *et al.* (1987) in a laboratory setting of a bitter tasting good. The study used an open-ended CV question as the starting point, then an iterative bidding CV question, followed by a question based on a Vickery auction mechanism. Results indicated initial mean WTP and WTA bids were significantly different as well as for mean values for the iterative bidding part, but for mean bids from the Vickery auction differences were not significant (however this finding of no significance was found to be sensitive to data outliers, specifically inclusion of an outlier). Hence, Coursey *et al.* (1987) conclude that WTP and WTA bids converged over repeated trials, and that a factor in this convergence was repeated trials which caused the respondents to become familiar and learn from the experiment.

Brookshire and Coursey (1987) follow this general theme in a study to examine if WTP and WTA will converge if similar techniques are applied in the field to the general public as opposed to a laboratory setting that typically involves students. They repeatedly emphasize that the process of repetition in a CV experiment can induce learning behavior and that bids could change as individuals become more familiar with the CV technique and experiment. The field application involved a CV question with a payment card that established the starting point, followed by an iterative bidding CV question to allow for revisions in CV bids. A Smith auction was also used in a separate field application and in a laboratory experiment for comparison purposes. The public good was a local park. Results found that differences remained significant for the CV technique, that an auction mechanism repeated twice also produced significant differences, and that the auction instrument in the laboratory with 5 trials (or repetitions) produced differences that were smaller in magnitude. The researchers conclude that repetitions can induce learning behavior but may be more difficult in field settings.

In two recent papers, Adamowicz *et al.* (1993) and Shogren *et al.* (1994) examine and empirically test plausible reasons for discrepancies in WTP and WTA measures following research of Hanemann (1991) and Randall and Stoll (1980). Hanemann (1991) argued that due to the degree of substitutability between goods, differences will occur between WTP and WTA measures. Where a good has many substitutes, WTP and WTA should coincide, and where few substitutes exist differences should result even in the absence of income effects. Adamowicz *et al.* (1993) conducted two laboratory experiments involving students. The first involved a movie ticket as the good and used open-ended CV questions. Results found that WTA and WTP were significantly different from each other. The second experiment involved a hockey ticket as the good and used closed-ended CV questions. Results indicated that differences between WTP and WTA remained significant, but that when a substitute-alternative good, i.e., viewing/listening event of the hockey game, was included differences between WTP and WTA became smaller although still significant: that is, the presence of a substitute caused differences to decrease. Although both experiments showed differences in WTP and WTA to be present and significant. the researchers attribute the presence of a substitute to the good to account for differences to become smaller and provides some evidence in favor of Hanemann's hypothesis.

Shogren *et al* (1994) conducted three experiments for three different goods. a brand-name candy bar, a lunch with a typical probability of being contaminated with one of five common pathogens versus a screened lunch. and a coffee mug. An auction mechanism (Vickery auction) was used in all experiments involving repeated trials to allow for learning. Results of the candy bar experiment representing a private good with many substitutes indicated that mean differences between WTP and WTA were not significant after 5 repetitions and that mean values were quite similar with all the adjustment from the WTA measure; differences were significant in the initial trial. For the second experiment, the contaminated lunch representing a nonmarket good with little substitution. Mean differences between WTP and WTA were significant in the initial trial and remained significant even after 17-20 repetitions. Even after elimination of outliers. mean WTA was 2 to 6 times larger than WTP. In a third experiment. a coffee mug served as the good (a good with many substitutes) and 10 repetitions were conducted. Findings indicated that differences were present in the initial trial but decreased over time, whereby differences were not significant over repeated trials; this result occurred with the full data and without outliers. The overall conclusion is that for a good with many substitutes, mean WTP and WTA converged and differences became negligible and insignificant, and for a good with few or little substitutes, differences in WTP and WTA persisted over many repetitions and remained significantly different. These results provide support for Hanemann (1991) hypothesis regarding the role of substitution effects in WTP and WTA differences.

Overall research that has examined differences in WTP and WTA has evolved from the early studies that merely noted empirical differences to examination of possible reasons why differences occur. General findings from this research suggest that learning may form a critical part of the explanation and repeated trials can foster this process. As one becomes more familiar with the valuation technique and good, individuals should revise CV bids. Whether or not convergence between WTP and WTA results may depend on the degree of substitutability of the good as proposed by Hanemann (1991) and firmly supported by Shogren *et al.*'s research but weakly supported by Adamowicz *et al.*'s research. Hence, it is suspected that convergence may be achieved for goods with many substitutes versus few substitutes, but more research is necessary to confirm such a finding. Implications for the present investigation concern the issue of convergence given repeated trials over time, and differences between WTP and WTA measures.

1.2.4. CV Literature: Implications for Treatment of $0 Bids, Outliers, and Influential Observations

Researchers involved with CV studies have noted for some time that CV data can exhibit noise and variability and may require some type of adjustment and/or treatment (Rowe *et al.* 1980, Desvousges *et al.* 1983, Desvousges *et al.* 1987, 1996, Mitchell and Carson 1989). At the same time CV bids can be sensitive to truncation and/or data adjustment (Coursey *et al.* 1987, Gregory and Furby 1987, Hausman 1993, Desvousges *et al.* 1996, Mitchell and

Carson 1989, Schulze *et al.* 1996).

A variety of techniques to adjust the data were suggested in early research, from removing observations that were 10 standard deviations beyond the mean value on both sides of the range or distribution, to other ad hoc techniques (Rowe *et al.* 1980, Desvousges *et al.* 1983, Smith and Desvousges 1986). Jones-Lee *et al.* (1985) for example in a study of the value of safety, removed upper-tail responses and estimated "trimmed means" for WTP measures. Even recent research has continued to use fairly ad hoc treatment methods. Loehman *et al* (1994) noted the potential problem of \$0 bids, nonvarying bids, and other protest bids: these problematic bids were removed for analysis. Shogren *et al* (1994) examined differences between WTP and WTA and eliminated the highest and lowest bids. Adamowicz *et al.* (1993) removed outliers that were more than 2 standard deviations from mean values. Whitehead *et al.* (1995) eliminated non-responses and the highest bids both classified as outliers in addition to \$0 protest bids in a study of use of an estuary in North Carolina.

Desvousges *et al.* (1983) was the first study to use and advocate the use of a statistical technique to identify influential observations, which includes extreme value observations (i.e., outliers or aberrant observations) and observations that exert an unusually large amount of influence on estimated regression equations, but are not extreme values, developed by Belsley *et al.* (1980). The technique developed by Belsley *et al.* (1980) is based on studentized residual measures from early research on statistical outliers and a new measure to examine the relative influence each observation has on estimated parameters in a regression model. Influential observations found by Desvousges *et al.* (1983) ranged from extreme value observations (outliers) including extremely high bids and extremely low bids and \$0 bids along with observations throughout this range (see Smith and Desvousges 1986: 101). This technique was also suggested by Mitchell and Carson (1989) to use in CV studies. In addition, Mitchell and Carson (1989) also suggest the use of robust regression techniques, a suggestion that has not been commonly applied in CV studies.

The Belsley *et al.* (1980) influential diagnostic technique has seen use in recent studies. In a study of moose hunting Boyle *et al* (1996) removed protest bids, bids in excess of 25% of income, and outliers identified by the Belsley *et al.* (1980) technique. Boyle *et al.* (1994) in a study of waterfowl protection also used the method of Belsley *et al.* (1980) to remove outliers from analysis.

Implications from treatment of outliers and data adjustments for the current investigation include examination of a variety of subsets of outliers and/or suspect observations to determine whether results are robust over such subsets. For example, responses of $0 bids, bids that exhibit noise (i.e., bids that oscillate between zero and nonzero bids over time), extreme value bids (both high and low bids), and bids identified as influential bids via the Belsley *et al.* technique are all of potential concern in this investigation.

1.3. Rationale of Investigation

The basis for this research follows. Standard consumer choice models assume that consumers make informed, rational decisions in arriving at their mix of purchased goods and services subject to their household budget constraint. This decision process, a major premise of economic theory, is based on consumer experience and information acquired over a period of time (e.g., evaluating alternatives, comparison shopping, advice from others, changes in tastes, prices, incomes, etc.). If this fundamental premise holds in contingent markets and for the CV method, then it is reasonable to expect that economic values (i.e., CV bids) produced from contingent markets can change over time or bidding periods. Some researchers in fact wonder if bids based on contingent markets will converge to some "true" value over time (Cummings *et al.* 1986b).

In the case of public-type goods most (if not all) individuals seldom go through an exercise of thinking about their preferences expressed in monetary terms. Such a process for private goods involves an extensive learning process. Components of this process are: 1) an understanding of the market and in some cases market conditions (e.g., availability of some large durable goods -- cars, houses -- i.e., supply conditions) can influence choice; 2) an understanding of quality and various attributes (e.g., durability, warranties) can influence preferences and choice; and 3) an understanding of competitive products and pricing behavior can also influence choice. This process is learned, revised, and developed over a period of time. A CV experiment essentially asks an individual to go through a similar process, possibly for the first time for public type goods, usually as a one-time process, and in an incredibly short amount of time (e.g., Mitchell and Carson 1989 note that the detail and description required in a CV question can range from 10 to 30 minutes to ask in a CV survey [compared to 1 to 2 minutes for a public opinion question version, 118]). This can introduce many problems, among which the most important is that the respondent may not have revealed his/her true preferences. As a result, it is possible that there are some issues that may have been overlooked or not treated adequately in past research concerning the CV method. Only recently has research begun to notice and emphasize the importance of learning through multiple repetitions in CV applications (Adamowicz *et al.* 1993. Brookshire and Coursey 1987, Shogren *et al.* 1994).

Research that has examined CV bids over time has been in the context of what is known as "test-retest" criteria and more recently "temporal reliability." To date. two studies have used the test-retest approach on the same group of individuals over time (Loomis 1989, Stevens *etal.* 1994). Loomis (1989) found that the direction of change in mean CV bids over time between those surveyed (general population vs. users) were contradictory: for the former group mean CV bids increased, while for users mean CV bids decreased over time. From an estimated AR(1) type WTP bid equations over both periods, results were contradictory again; results were consistent with convergent behavior for users. and with random walk behavior for the general population group. Stevens *et al.* (1994) followed a similar approach and found that mean bids of existence value were revised upwards over two time periods, and from an estimated AR(1) type WTP bid equation results suggested that WTP bids exhibited convergence behavior over time.

In general, the inherent design of the test-retest approach can only yield limited information about the process of CV bid revisions over time, and general conclusions based on these types of studies cannot be made. The test-retest approach implicitly assumes all other factors such as tastes and experience with the contingent market remain constant over time, an assumption that is at odds with the standard consumer choice theory. Another problem is that only two time periods are considered, and it generally takes more than three data points to produce some evidence of trend. Hence, it remains unknown if individuals would continue to revise their CV bids given a longer time horizon as evident in laboratory-type experiments, and if this process of revision would yield an individual's "true" preferences.

The proposed investigation is unique in several ways. It will rigorously examine whether individuals participating in contingent markets via CV surveys revise their bids over time if given a long time horizon. It also will examine if experimental economics techniques can be used in field settings (e.g., experiments of active users of a particular resource/activity based on repeated games), and if the resulting evidence can provide useful information in understanding agents' behavior in contingent markets, and in future advancements of the CV method. A study of this nature can have critical implications for future CV studies, both field and experimental, and could initiate further investigations into the CV method and its theoretical basis.

1.4. Research Objectives

This investigation uses quasi-experimental economics research techniques involving repeated surveys (i.e., repeated games) over multiple time periods to examine various aspects of the CV method and resulting CV bids (WTP and WTA). Specific objectives will examine:

1) if agents' behavior in contingent markets corresponds to agents' behavior from which standard consumer choice theory is based on by using a variety of methods from statistical inference to econometrics and time series techniques as follows:

1.1) to determine if mean values of CV bids are statistically different over multiple time periods;

1.2) to apply time series techniques to panel data to determine if CV bids converge over time (i.e., if the process describing CV bids converges), diverge or follows a random walk;

1.3) to examine the empirical effect of subsets of outliers and suspect observations on convergence results using Chow tests, and whether results are robust to these subsets; and

1.4) to determine if WTP bids and WTA bids converge towards one another.

This research will contribute to efforts involved with 1) consistently/ reliability and behavior of intertemporal CV bids, 2) treatment of outliers and $0 bids over multiple time periods, 3) discrepancy between WTP and WTA bids, and 4) the use of time-series techniques in panel level data.

Chapter 2 - Overview of the Nonmarket Experiment and Research Methods

2.1. Introduction

In this Chapter a description of the quasi-experiment is discussed and the research methods.

2.2. Description of the Experiment

The data this investigation is based on comes from repeated surveys of about 200 active recreational fishermen that voluntarily agreed to participate in a unique study of the value of recreational fishing and of sport caught fish throughout the 1990 fishing season in New Jersey. The study was conducted by a joint project between the New Jersey Marine Fisheries Administration (NJMFA) and Rutgers University. The main purpose of this study was to obtain bids of WTP and WTA in open-ended and close-ended contingent valuation question formats over multiple time periods. A time period was to represent a month.

Meetings took place with most of the recreational fishermen prior to the survey process, to explain the study and their involvement in the study. The survey technique used in this study was based on mail survey techniques, and questionnaires were sent out on a monthly basis beginning in April-May and ending in November collecting data on their fishing activity in the previous month. Up to five repeated surveys (i.e., repeated games) were completed for any one individual resulting in up to 5 bidding periods. Questions regarding the marginal valuation of a specific fishing trip (e.g., shark trip, bluefin tuna trip, striped bass trip, etc.) and that of a specific fish caught (e.g., striped bass, bluefish, fluke, mako shark, yellowfin tuna, etc.) followed various formats for several commonly targeted fish (i.e., fluke, bluefish, striped bass, yellowfin tuna, and mako shark): open-ended questions of WTP, WTA: open-ended questions based on different types of information: closed-ended questions where dollar amounts specified were randomly varied over five groups within each month (the closed-ended questions pertained to a specific fishing trip, a specific fish caught, and specific fish caught based on various size dimensions).

Responses to the economic value questions depended on whether the angler took a targeted trip for any one of the five commonly targeted fish during the month in question. If the angler did take such a targeted trip during a month, then they proceeded onto the valuation questions and entered responses for WTP and/or WTA. It was hoped that anglers would revise their bids over time when they became more familiar with the questions. contingent valuation technique. and fishing behavior.

2.3 Research Methods

2.3.1. Tests of Convergence - Statistical Inference Methods

Statistical inferences involved tests of hypotheses based on paired t-tests and Duncan's multiple range tests. Hypotheses were developed to examine: 1) the equality of the variances over CV bids that contained zeros (full data set) and CV bids without zeros (subset of data without S0's); 2) the equality of paired means across the full data and the subset of data without S0 bids ; 3) the equality of paired means (i.e.. mean CV bids across any two successive months) and the equality of means over all time periods. These types of statistical inferences can provide preliminary evidence regarding convergence over time.

For tests of the equality of variances over the full data and a subset of data without S0 bids. the following test was used:

H_0: $\sigma_f^2 = \sigma_s^2$. versus not. The test statistic is:

$F = s_f^2 / s_s^2$. and the decision criteria is to reject H_0

if $F \leq F(\alpha/2)\,(n_f-1,\ n_s-1)$ or if $F \geq F(1-\alpha/2)\,(n_f-1,\ n_s-1)$. where the

f-subscript refers to the full data set and the s-subscript the data subset without S0's.

Paired t-tests of the equality over the full data and a subset of data without S0 bids were as follows:

H_0: $\mu_f = \mu_s$, versus not. The test statistic with unequal variances over both data sets is:

$t = (X_f - X_s) / (s_f^2/n_f + s_s^2/n_f)^{1/2}$, and the decision criteria is to reject H_0

if $t \geq |\,SMT\,|$, where $SMT = (w_f t_f + w_s t_s) / (w_f + w_s)$, where $w_f = s_f^2/n_f$

and $w_s = s_s^2/n_f$, and $t_f = t_{(1-\alpha/2)(n_f-1)}$ and $t_s = t_{(1-\alpha/2)(n_s-1)}$.

Similar tests over successive time periods follow where the subscripts f,s are replaced with t, t+1.

Formal tests of the equality of the regression model across the full data and subset of data were also examined based on tests similar to tests of structural change of the data (Greene 1993) and Fomby *et al.* (1984). These tests are sometimes referred to as Chow-tests following Chow (1969) and examine whether or not the subset of data without S0 bids should be pooled and treated as a full data set. The formal tests are of the form:

$F = [(SSE_r - SSE_u)/k] / [SSE_u/(n_1 + n_2 - 2k)]$, where

SSE_r = sum of squared error for the pooled or restricted model (i.e., model for full data set), SSE_u = sum of squared error for the unrestricted model ($SSE_u = SSE_1 + SSE_2$), k is the number of estimated parameters (including a intercept term if necessary), n is the number of observations, and the subscripts refer to the 1st data subset - 1, and the 2nd data subset - 2. In the test of the subset of S0 bids, n_1 is the number of observations for the set without S0 bids, n_2 the subset with S0 bids, and both sum to the full data set. The decision criteria is to reject the null hypothesis that the subsets are equal if $F \geq F(\alpha)\,(k, n_1 + n_2 - 2k)$, and conclude do not pool the data (the subsets are significantly different). In the case where insufficient observations exist (i.e., $k > n_1$), the following test is based on Fisher

(1970):

$$F = [(SSE_r - SSE_u)/n_1] / [SSE_u/(n_2 - k)] \text{, where}$$

SSE_u is now the sum of squared error for the model associated with the second data set (i.e., $SSE_u = SSE_2$). The decision criteria is to reject the null hypothesis that the subsets are equal if $F \geq F(\alpha)\ (n_1,\ n_2 - k)$, and conclude do not pool the data (the subsets are significantly different).

2.3.2. Convergence in Time-Series Techniques

Methods based on time series techniques are more sophisticated and can provide more insight and perhaps more conclusive evidence regarding convergence. Various autoregressive (AR) models were estimated to determine if CV bids (i.e., WTP amounts and WTA amounts) in previous time periods are good predictors of bids in current periods and in future periods, and if CV bids converge over time (that is, if the process describing CV bids over time converges, etc.). Such models are as follows:

(1.1) AR(1) model: $BID(t) = \beta_1 BID(t-1) + e(t)$,

(1.2) AR(2) model: $BID(t) = \beta_1 BID(t-1) + \beta_2 BID(t-2) + e(t)$,

(1.3) AR(3) model: $BID(t) = \beta_1 BID(t-1) + \beta_2 BID(t-2) + \beta_3 BID(t-3) + e(t)$,

(1.4) AR(4) model: $BID(t) = \beta_1 BID(t-1) + \beta_2 BID(t-2) + \beta_3 BID(t-3) + \beta_4 BID(t-4) + e(t)$,

where BID refers to CV bids (WTP or WTA) associated with a specific time period, the ß's are the parameter estimates, and e(t) is the error term assumed to be white noise (i.e., mean of zero, finite variance, with error terms uncorrelated across time periods). As specified, these AR models do not contain a drift term, i.e., a constant coefficient. This investigation examines the above AR models as well as AR models with drift and trend terms where appropriate.

Regarding the AR(1) model, evidence in favor of convergence is if $|\beta_1| < 1$, evidence of a random walk is if $|\beta_1| = 1$, and if the process diverges $|\beta_1| > 1$(Hamilton 1994, Judge *et al.* 1988). Higher order AR processes become more complicated quickly; for the AR(2) process and above the condition for convergence is if the roots (i.e., characteristic roots) of the polynomial equation associated with the appropriate AR model lie inside the unit circle; specifically if the length or modulus of the roots are within the unit circle, i.e., less than one. For example in the following AR(2) model:

(2.1) $y_t = a + \beta_1 y_{t-1} + \beta_1 y_{t-2} + e_t$, expressed in lag notation as

(2.2) $(1 - \pi_1 L - \pi_2 L^2)y_t = a + e_t$, where L is the lag operator.

The condition for convergence of the AR(2) process is met if the roots of

(2.3) $(1 - \pi_1 z - \pi_2 z^2) = 0$ lie outside the unit circle,

or if the roots of the polynomial (or characteristic equation) obtained from factoring the AR(2) model expressed in lag operator form, i.e.:

(2.4) $(\lambda^2 - \varphi_1\lambda - \varphi_2) = 0$ lie inside the unit circle.

For the purposes of this investigation the criteria for convergence will refer to the roots of the polynomial equation associated with the appropriate AR model, e.g., equation (2.4), and if all of the roots lie inside the unit circle the process is said to converge. To determine and express the polynomial or characteristic equation associated with a given difference equation is easier then it seems. Following Shone (1997), for a difference equation of the form, i.e., homogeneous difference equation:

(3.1) $y(t) = ay(t-1) + by(t-2)$, or

(3.2) $y(t) - ay(t-1) - by(t-2) = 0$,

the associated characteristic equation can be expressed as:

(3.3) $x^2 - ax - b = 0$, i.e., in the form of a quadratic equation. In general we have:

(3.4) $y(t) = ay(t-1) + by(t-2) + cy(t-3) + \ldots + ny(t-n)$, or

(3.5) $y(t) - ay(t-1) - by(t-2) - cy(t-3) - \ldots - ny(t-n) = 0$,

(3.6) $x^n - ax^{n-1} - bx^{n-2} - \ldots - (n-n+1)x^{n-n+1} - n = 0$.

For AR models that contain a drift term (i.e., a constant or intercept), a non-homogeneous difference equation. Shone (1997) emphasizes that one can only determine the characteristic roots associated with the homogeneous part. Consider the following AR(2) model, a non-homogeneous difference equation:

(3.7) $y(t) = ay(t-1) + by(t-2) + c$, or

(3.8) $y(t) - ay(t-1) - by(t-2) = c$.

The homogeneous part of the above AR(2) model refers to all but the constant term (c), i.e., the part on the left-hand of equation (3.8). Finding the characteristic roots associated with an AR(2) model with drift then becomes similar to the process for an AR(2) model without drift where the drift term is ignored.

Using mathematical computational software packages such as Maple or Mathematica it becomes a straightforward process to solve the characteristic equation and find all the roots.

In the case that all roots are real, the issue becomes trivial, but in the case of real and complex roots, i.e., $z = a + bi$, where $i=(-1)^{1/2}$ one must examine the unit circle in real and imaginary space, usually graphed with the real part of the root a on the x-axis and the imaginary part of the root b on the y-axis, and determine whether the complex root lies within the unit circle. This is accomplished as follows, the length of the vector from the origin to the x,y coordinate of the root is determined from by the length of the hypotenuse from the Pythagoras theorem, that is $|z| = (a^2 + b^2)^{1/2}$, where $|z|$ denotes the length of the vector of the root. If $|z|$ is <1 the root lies within the unit circle.

2.3.3. Treatment of Subsets of Suspect Observations and Detection of Influential Observations

In general the treatment of subsets of observations with suspect CV bids, e.g., $0 bids, extreme value CV bids (both large and small bids), and influential observations will be examined based on the tests outlined on structural change. This involves tests of the equality of the regression model associated with a specific data subset to determine if the data should be pooled or not. The previous discussion in Section 2.3.1. has already covered these tests in detail.

The technique of Belsley *et al.* (1980) consists of up to five diagnostic statistics: 1) DFBETA, a measure of the strength each observation has on estimated parameters of the regression model; 2) DFFIT, a measure of the strength each observation has on the predictions of the regression model; 3) h_i, a measure of leverage of each observation on the regressors; 4) COVRATIO, a measure of the influence of each observation on the variance-covariance matrix; and 5) RSTUDENT, a measure of the influence of each observations' residual. The central approach in all these measures is to isolate the *i*th observation and reestimate the regression model to examine the observations' influence. This is sometimes examined as a ratio of a measure without the *i*th observation to the measure based on the full data, whereby the observation will have weak effects/influence if the ratio is close to one, and significant influence if different than one (e.g., the measures of COVRATIO, etc.). In addition, in some cases these ratios are weighted or scaled by a measure of variance. This is relevant for DFBETA, and DFFITS. The following discussion treats these diagnostic statistics in detail.

Leverage

A statistic that becomes important in many of the diagnostic statistics that follow comes from the least-squares projection matrix, commonly referred to as the hat matrix:

4.1) $H = X(X'X)^{-1}X'$.

The usefulness of the hat matrix is in determining the fitted or predicted values from least-squares, i.e., $y = X\mathbf{b} = Hy$. The statistic based on this hat matrix represents the diagonal elements of the hat matrix and is defined as:

4.2) $h_i = x_i(X'X)^{-1}x_i'$,

where x_i represents the *i*th observation or row of the X matrix (a vector of 1xk dimension). This statistic serves as a measure of leverage, the influence of the *i*th observation on the regressors (i.e., independent variables). The hat statistic will become large for observations that are very different or removed from the values of the bulk of the data that pertain to the independent observations (i.e., the X-data). Belsley *et al.* (1980) recommend values of leverage that exceed 2(k/T) (k is the number of parameters and T observations), that is observations where

4.3) $h_i > 2(k/T)$,

are indicative of leverage points and these observations should be examined in more detail to determine whether the observation is questionable and influential. The diagnostic statistics that follow aid in this decision.

Studentized Residuals

Residuals have long been a traditional means to examine the data and regression model. The traditional popularity of the residual in the detection of outliers comes from the desire to consider observations that result in large errors, the same is true with statistical measures designed to detect large errors. The problems of heteroscedasticity and autoregression have been detected from residuals. Scatter plots have also been used. However, both methods suffer from a lack of statistical rigor in determining limits that observations must exceed to qualify as outliers, traditionally these methods were used in a somewhat ad hoc manner. Similar comments hold concerning the problems of heteroscedasticity and autoregression, and research has produced superior statistical measures and tests to detect these conditions. The development and advancement of diagnostic statistical measures such as those statistics refined by Belsley *et al.* (1980) brings to the detection and identification of influential observations the statistical rigor that was lacking in earlier approaches and fills a void in this area.

Judge *et al.* (1988) emphasize that residuals of the OLS (ordinary least squares) model are not always a good means of identifying observations with large error terms because the OLS procedure weights extreme errors heavily which can yield residuals moderately large in absolute value. Such a problem can be overcome by considering the *i*th residual of the OLS model, that is the residual of of the OLS model with the ith observation deleted. The use of standardized residuals has been suggested by researchers (Belsley *et al.* 1980), that is a ratio of the residual without the ith observation divided or standardized by its variance $var(e_i) = \sigma^2(1 - h_i)$. The standardized residual becomes:

4.4) $e_s = e_i / [\sigma(1 - h_i)^{1/2}]$.

Belsley *et al.* (1980) recommend using s(i) the estimated standard deviation of the OLS model without the ith observation giving the studentized residual:

4.5) $e_i^* = e_i / [s(i)(1 - h_i)^{1/2}]$ or $[y_i - x_i\mathbf{b}(i)] / \{s(i)[1 - x_i(X'X)^{-1}x_i']^{1/2}\}$.

Belsley *et al.* (1980) suggest that values of e_i^* where

4.5.a) $|e_i^*| \geq 2$ is of potential concern.

DFBETA

DFBETA represents a measure of the strength each observation has on estimated parameters of the regression model. This measure is defined as the difference between the estimated parameters of the full model, **b**, and the estimated parameters of the model without the *i*th observation, **b**(i), or **b** - **b**(i), expressed as:

4.6) $\mathbf{b} - \mathbf{b}(i) = [(X'X)^{-1}x_i'e_i] / [1 - h_i]$.

The magnitudes of the elements in the vector **b** - **b**(i) [a kx1 vector, with k parameters] depend on units of measurement and it becomes necessary to scale/standardize this vector by a standard deviation measure (Belsley *et al.* 1980, Judge *et al.* 1988). Consider the jth element of **b** - **b**(i), $\mathbf{b}_j - \mathbf{b}_j(i)$. This is then scaled by an appropriate standard deviation estimate of $\mathbf{b}_j$, the square root of the jth diagonal element of the estimated variance without the ith observation, $s(i)^2\ (X'X)^{-1}$. One can then express (4.6) in terms of this standard deviation. Define a_{jj} as the jth diagonal element of $(X'X)^{-1}$, and c_{ij} as the i,jth element in

$(X'X)^{-1}X' = C$. DFBETA can be expressed as follows:

$$4.7)\quad DFBETA = [\mathbf{b}_j - \mathbf{b}_j(i)] / [s(i)a_{jj}^{1/2}]$$

$$= [c_{ij}e_i] / [s(i)a_{jj}^{1/2}]\,[1 - h_i]$$

$$= [c_{ij}e_i^*] / [a_{jj}(1 - h_i)]^{1/2}.$$

The usefulness of introducing the leverage statistic, h_i, and the studentized residuals, e_i^*, now becomes meaningful and aids in the interpretation of the DFBETA statistic. When the studentized residual is large, DFBETA will become large. In addition as the leverage statistic increases so too will DFBETA. Hence, these three statistics move in the same direction given an aberrant or influential observation, and all three statistics should cross verify one another. However, what remains concerns criteria about what is meant by "large" DFBETA statistics. Belsley *et al.* (1980) recommend a "size-adjusted cutoff" that could be suggestive of an influential observation where

$$4.8)\quad |DFBETA_{ij}| > 2/T^{1/2}.$$

Any such observations involve further investigation.

DFFITS

DFFITS is a measure of the strength each observation has on the predictions of the regression model. To derive this expression premultiply both sides of (4.6) by x_i, which gives us

4.9) $x_i\mathbf{b} - x_i\mathbf{b}(i) = [x_i(X'X)^{-1}x_i e_i] / [1 - h_i] = [h_i e_i] / [1 - h_i]$.

Scaling (4.9) by the standard deviation of y_i by the estimate $s(i)[x_i(X'X)^{-1}x_i']^{1/2}$ or $\sigma(i)h_i^{1/2}$ results in a diagnostic measure of each observation on the predicted model:

4.10) $\text{DFFITS} = [x_i\mathbf{b} - x_i\mathbf{b}(i)] / [s(i)h_i^{1/2}]$

$= [h_i / (1 - h_i)] [e_i / s(i)h_i^{1/2}]$

$= h_i e_i / [(1 - h_i) s(i)h_i^{1/2}$

$= h_i^{1/2} e_i / [s(i)(1 - h_i)]$

$= [h_i^{1/2} / (1 - h_i)^{1/2}][e_i / [s(i)(1 - h_i)^{1/2}]$

$= [h_i / (1 - h_i)]^{1/2} e_i^*$.

Again, large leverage statistics and/or studentized residuals will result in large values of DFFITS, hence, all four statistics (h_i, e_i^*, DFBETA, and DFFITS) can serve as cross-verification of one another and all move in the same direction. Size-adjusted cutoffs recommended by Belsley *et al.* (1980) are

4.10.a) $|\text{DFFITS}_i| > 2(k/T)^{1/2}$.

Individual observations that result in such values are suggestive of providing greater than usual influence on the predicted model and warrant further investigation.

COVRATIO

COVRATIO, consisting of the ratio of the determinants of the full variance-covariance matrix and the variance-covariance matrix with the *ith* observation/row deleted is a measure of the influence of each observation on the variance-covariance matrix. This ratio statistic is:

4.11) $\text{COVRATIO} = \det \{s(i)^2[X(i)'X(i)]^{-1}\} / \det [s^2(X'X)^{-1}]$.

They note that $\det [X(i)'X(i)]^{-1}\} / \det [(X'X)^{-1}] = 1/(1 - h_i)$. From substitution into (4.11) one obtains:

4.12) $\text{COVRATIO} = s(i)^2 / s^2(1/(1 - h_i)$.

A series of substitutions that incorporate the studentized residual and a process of simplifying the expression Belsley *et al.* (1980) obtain a boundary condition for the covariance ratio as:

4.13) $\text{COVRATIO} \sim [1 - (3k/T)]-1 \sim [1 + (3k/T)]$. Hence, the suggest that any

value of COVRATIO where

4.13.a) $|\text{COVRATIO}| \geq 3k/T$ is a rough indicator of potentially troublesome observations.

This set of diagnostic statistics are firmly rooted in statistical theory and provide the rigor that was once lacking in examinations of outliers and suspect observations.

Chapter 3. Results: Characteristics of the Data

3.1. Characterization of the Data: Overview and Statistical Inferences

This overview is meant to present a visual examination of the mean CV bid series corresponding to a variety of categories, to give some idea of direction of the mean-bids over time, and to discuss results of the statistical inferences. In this study the nonmarket good represents fishing trips specific to the targeted species striped bass, bluefish, and fluke. All other species of fish were not as popular and were only reported sporadically. Although the generic good represents a targeted fishing trip, these species are different enough to classify as unique goods, and possibly may not serve as substitutes to one another. In addition, several categories that represented different characteristics of the angler and fishing behavior were examined to determine if these characteristics had any influence on CV bid behavior.

The categories are as follows: "All Individuals" - all individuals that submitted bids for a particular specie in each time period (i.e., month), "Took Targeted Trips" - only those individuals that took at least one targeted trip for the specie they submitted a bid for in a particular month, "Most Recent Trip" - those individuals that took their most recent trip for the specie they submitted a bid for in a particular month, and "Trip & Catch $\geq$ 1" - a subset of individuals contained in the category "Most Recent Trip," where it includes those individuals where their most recent trip was for the specie they submitted a bid for *and* caught at least one of these fish in a particular month. Each category is basically a subset of the preceding category so in general the number of observations (i.e., individuals that submitted bids) declines across the categories.

In general, the monthly data exhibit another characteristic related to the abundance of a particular specie of fish associated with migratory behavior, and in turn fishing activity. For fluke, the number of trips and catch is high during the May-August period, and declines afterwards, hence the number of observations associated with CV bids is high in the months of July and August and decline after. For bluefish, abundance and fishing activity is fairly consistent throughout the summer months and early fall months. And for striped bass abundance and fishing activity is fairly consistent throughout the summer months and begins to increase later in the Fall.

3.2. Implications of $0 CV Bids

3.2.1. First-Level Statistical Inferences

As emphasized in the discussion of the literature the issue and treatment of $0 CV bids is a concern in CV studies. In a study such as the present one, the implication of $0 CV bids has different implications because most of the literature has examined how socioeconomic factors affect CV bids in a "bid equation" so $0 bids show up on the left hand side of the equation. In the case of many $0 bids the data may be described as following a censored distribution and estimation of a Tobit model is appropriate. However in the case where $0 CV bids appear on both sides of the regression model, one case is a possible concern. The case where an individual gives a $0 bid in one period followed by a nonzero bid in a consecutive period followed by a $0 bid and then followed by a nonzero bid can introduce noise, i.e., more highly variable data than one would face in a historical analysis of stock prices for example where all observations contain a nonzero value for price. The case of CV bids where all responses are zero is of little concern, because in a regression model these observations have no impact on the estimated parameter but do have impacts on specific measures that are adjusted by the number of observations, e.g., means, standard deviation, t-statistics.

An examination of graphs of the mean CV bid series will illustrate the effect of inclusion and elimination of $0 bids. Figures 1.1 to 1.8 correspond to striped bass. In each of the four categories, the series including $0 bids (the full data set) and the series without $0 bids move together over time. Mean CV bids without $0 bids always lie above the mean bids based on the full data and both series appear to correspond to each other exhibiting similar behavior. It appears that inclusion of $0 bids dampens extreme movement. If one ignores the last period mean bid in the graphs of the CV bid series for bluefish the series also move

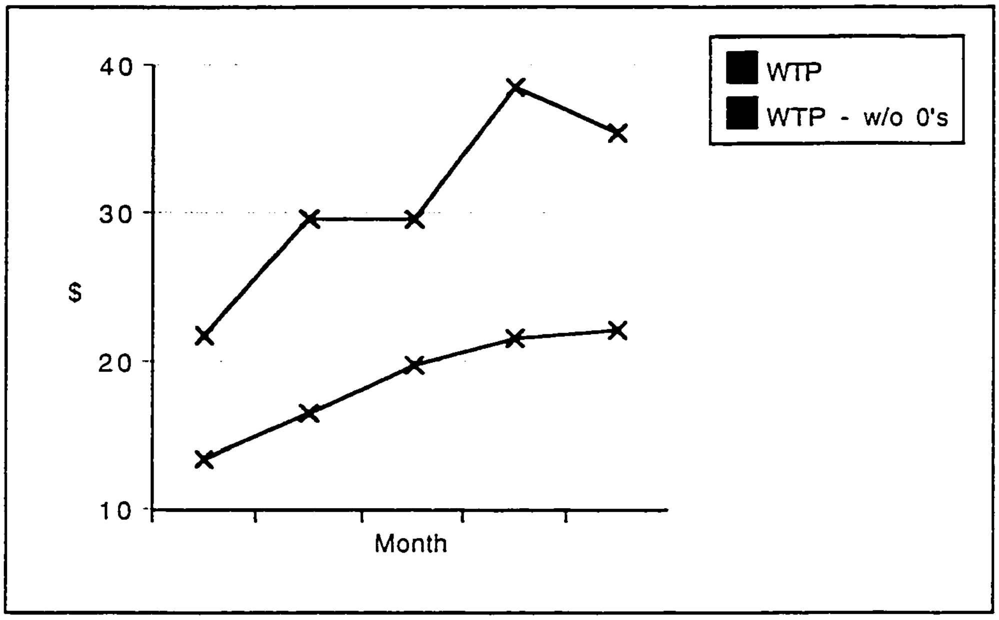


Figure 1.1. WTP - All Individuals: Striped Bass

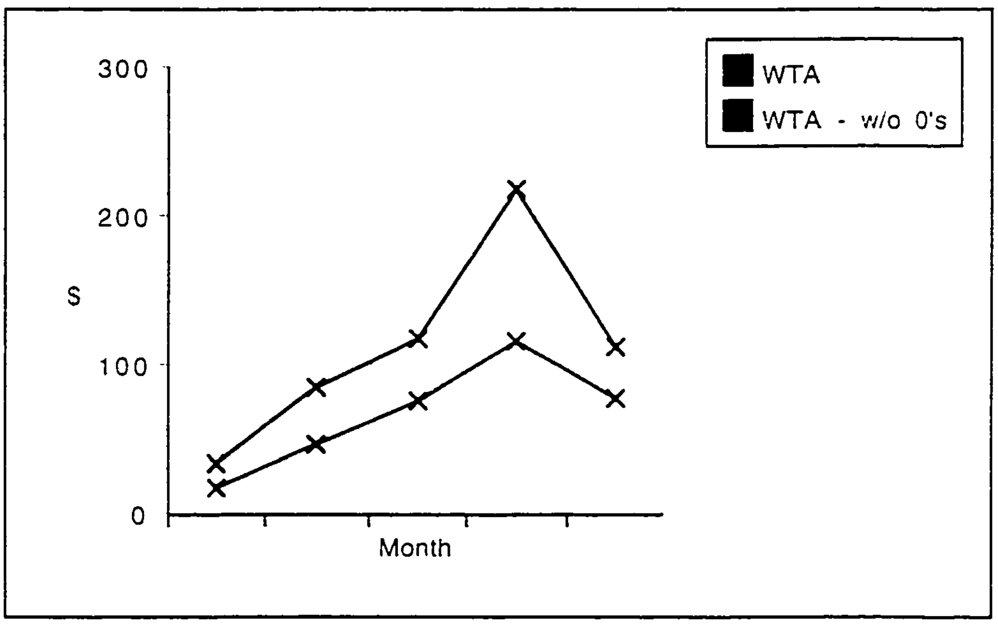


Figure 1.2. WTA - All Individuals: Striped Bass

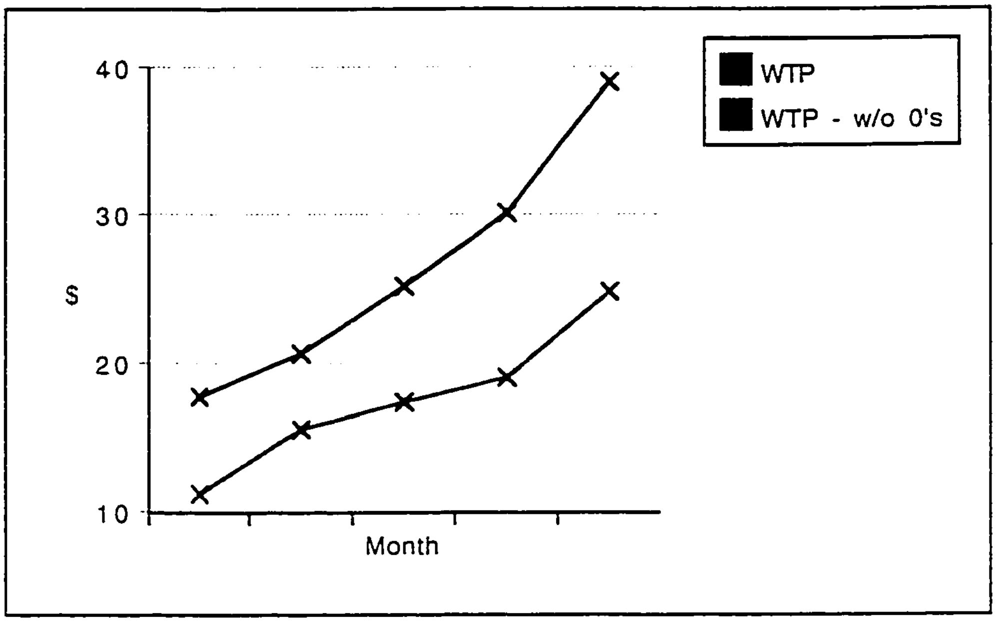


Figure 1.3. WTP - Took Targetted Trip: Striped Bass

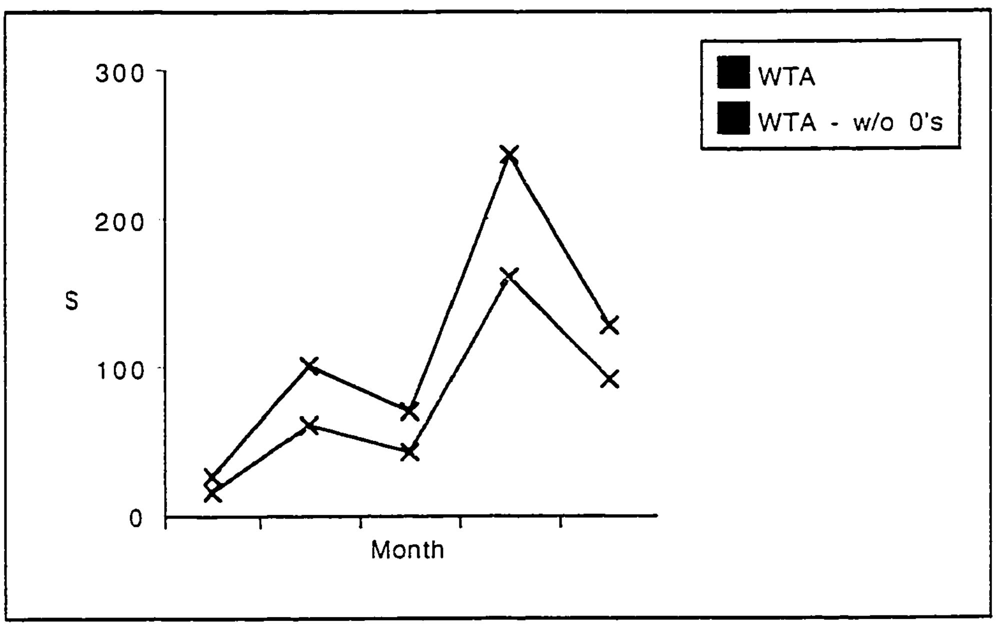


Figure 1.4. WTA - Took Targetted Trip: Striped Bass

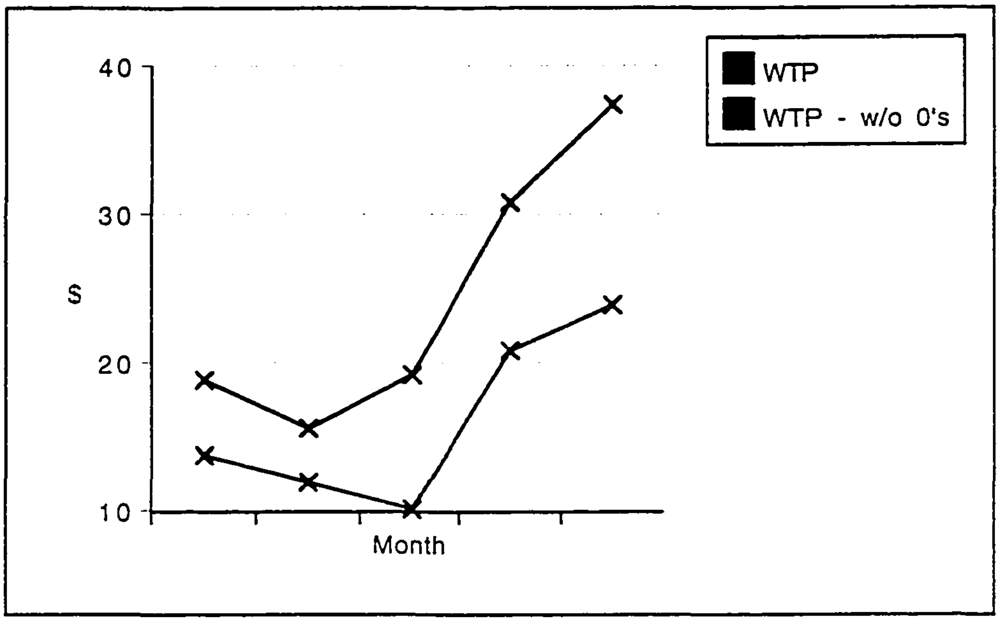


Figure 1.5. WTP - Most Recent Trip: Striped Bass

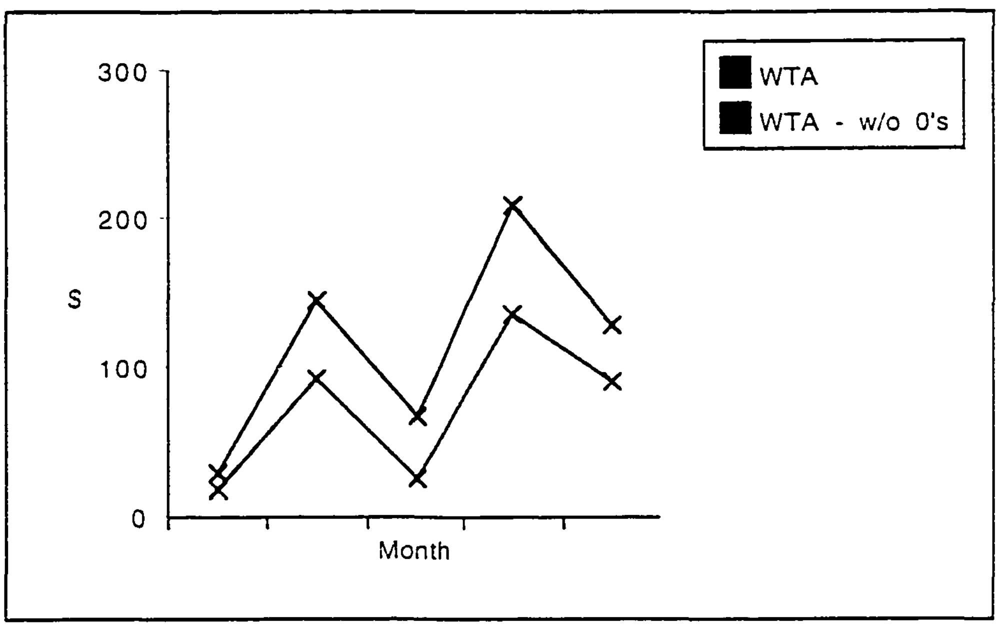


Figure 1.6. WTA - Most Recent Trip: Striped Bass

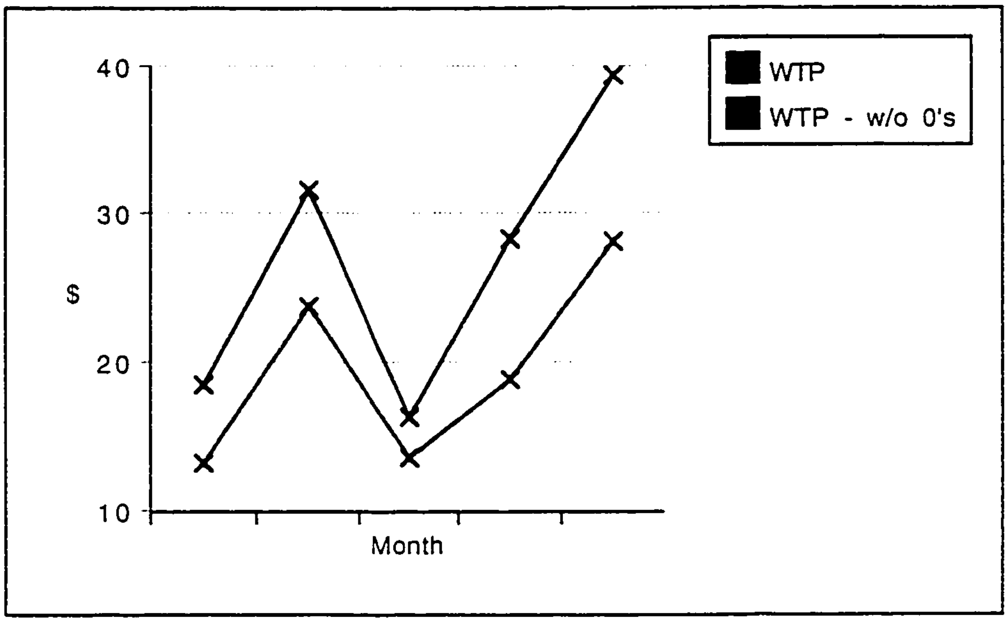


Figure 1.7. WTP - Trip & Catch >= 1: Striped Bass

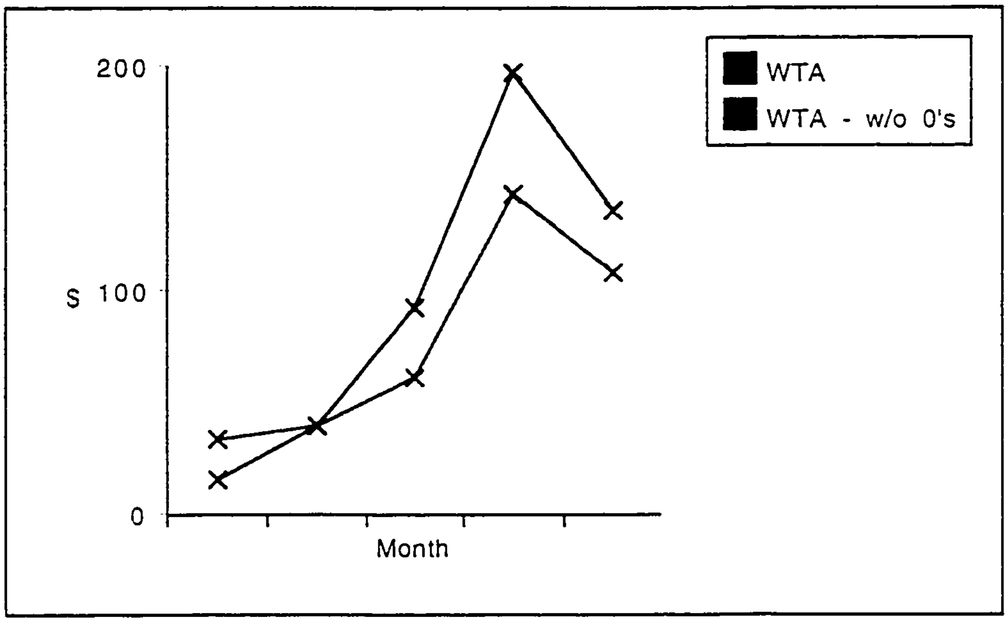


Figure 1.8. WTA - Trip & Catch >= 1: Striped Bass

together (see Figures 2.1 to 2.8). Similar behavior follows in the figures associated with fluke (Figures 3.1 to 3.8). As demonstrated in the diagrams the mean values of CV bids without $0 bids always lie above the mean values that correspond to the full data set.

Statistical inferences based on tests of equal variances of the two series, the series with $0 bids and the series without $0 bids, indicated that variances were not significantly different between series based on a F-test for most of the cases tested (2 WTP series out of 60 possible categories, and 6 out of 60 categories for WTA bids). Tables 3.1 to 3.3 contains summary results. Hypothesis tests of equal means across both series were also conducted and results indicated that in several categories mean WTP bids were significantly different based on a t-test, results were spotty for mean WTA bids; 21 mean WTP bids were significantly different at the .10 level out of 60 possible categories, whereas 4 of 60 mean WTA bids were significant (Tables 3.1 to 3.3).

3.2.2. Second-Level Statistical Inferences

Second level statistical inferences were conducted that involved testing the equality of the regression model associated with subsets of the data with and without $0 CV bids to determine more conclusively if the data should be pooled or not. For this analysis various AR models were estimated and the test of equality was based on a Chow-test. Results for the AR(1) models indicated that in 4 out of 16 possible categories for WTP bids, the null hypothesis that both AR models are equal versus not was rejected (Table 3.4). Results corresponding to the WTA AR(1) models only showed 1 of 16 categories that was significant. Hence results are not overwhelmingly convincing that the two series are different and that they should not be pooled. Similar results were obtained from tests based on AR(2), AR(3) and AR(4) models; significant findings were sparse (Tables 3.4 and 3.5). On the basis of these tests one could conclude that the occurrence of $0 CV bids in this investigation does not appear to have strong effects on CV bid behavior over time.

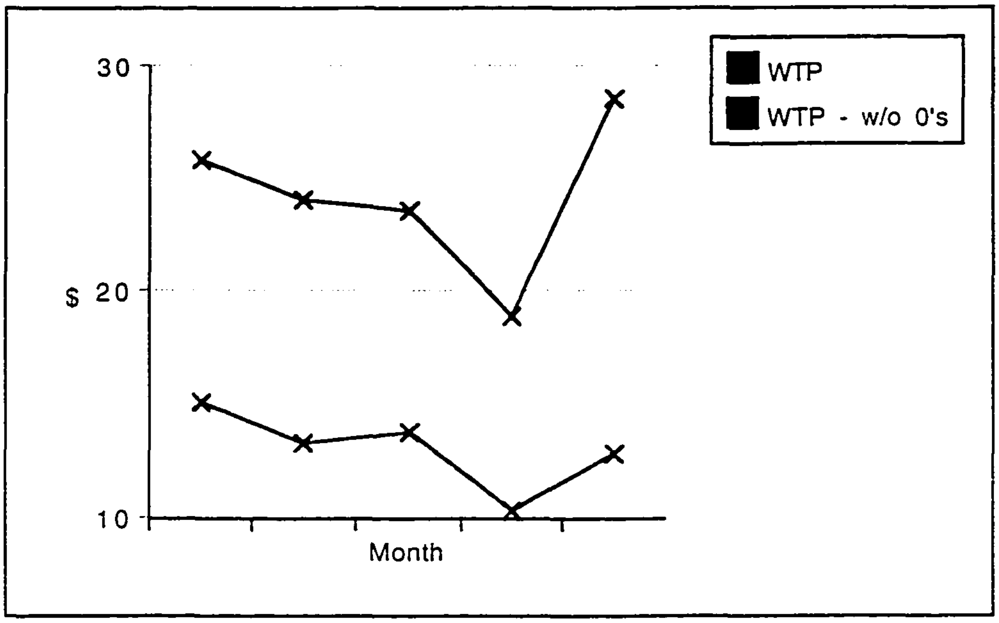


Figure 2.1. WTP - All Individuals: Bluefish

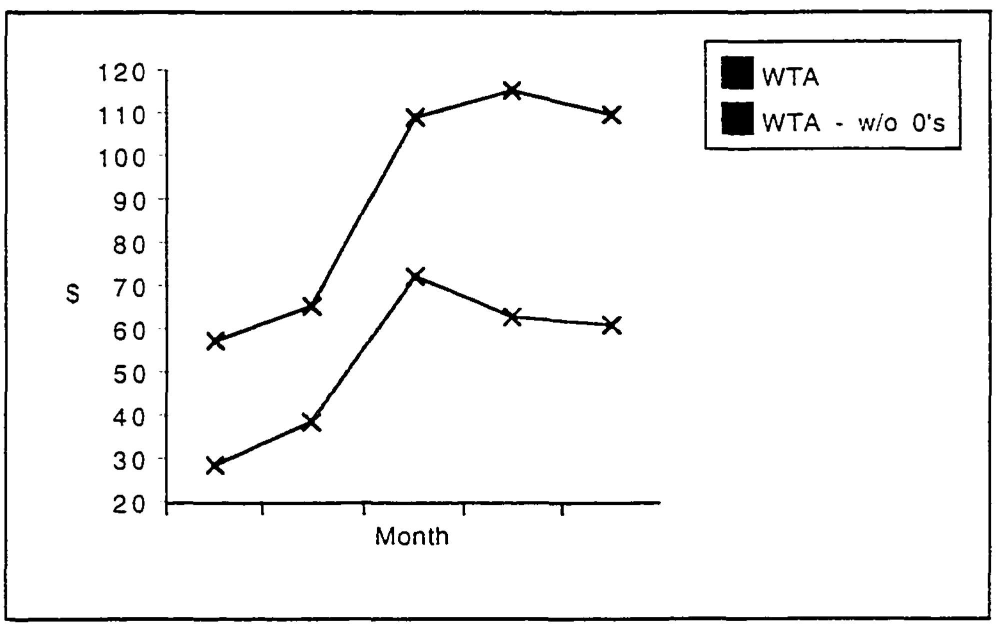


Figure 2.2. WTA - All Individuals: Bluefish

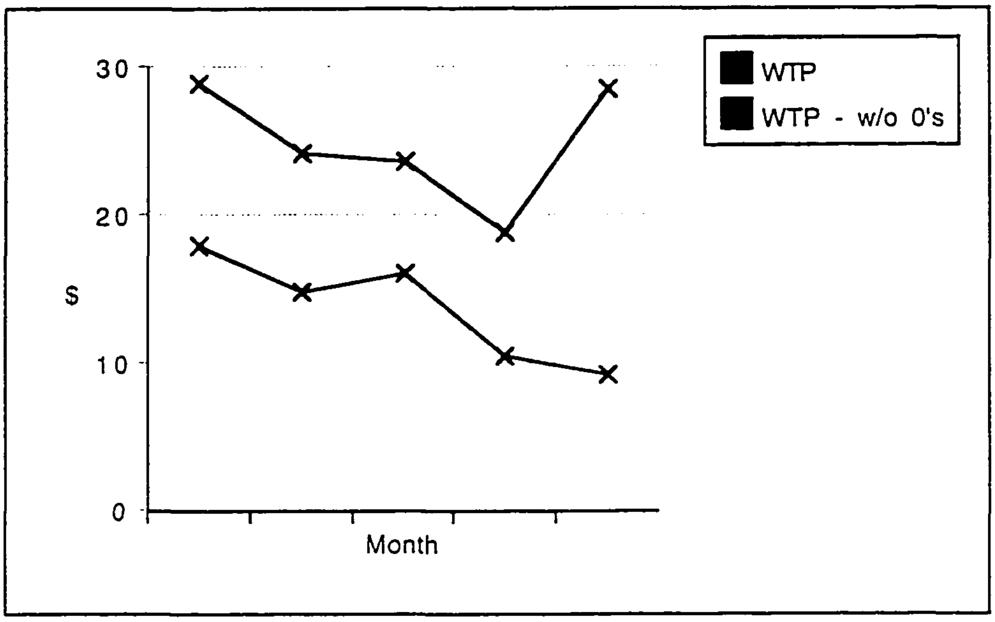


Figure 2.3. WTP - Took Targetted Trip: Bluefish

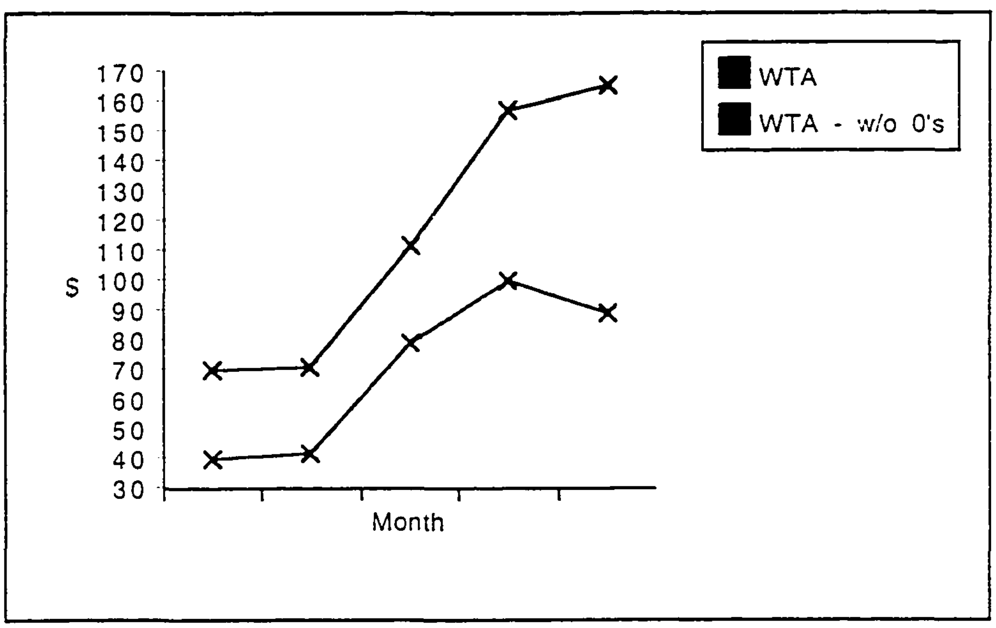


Figure 2.4. WTA - Took Targetted Trip: Bluefish

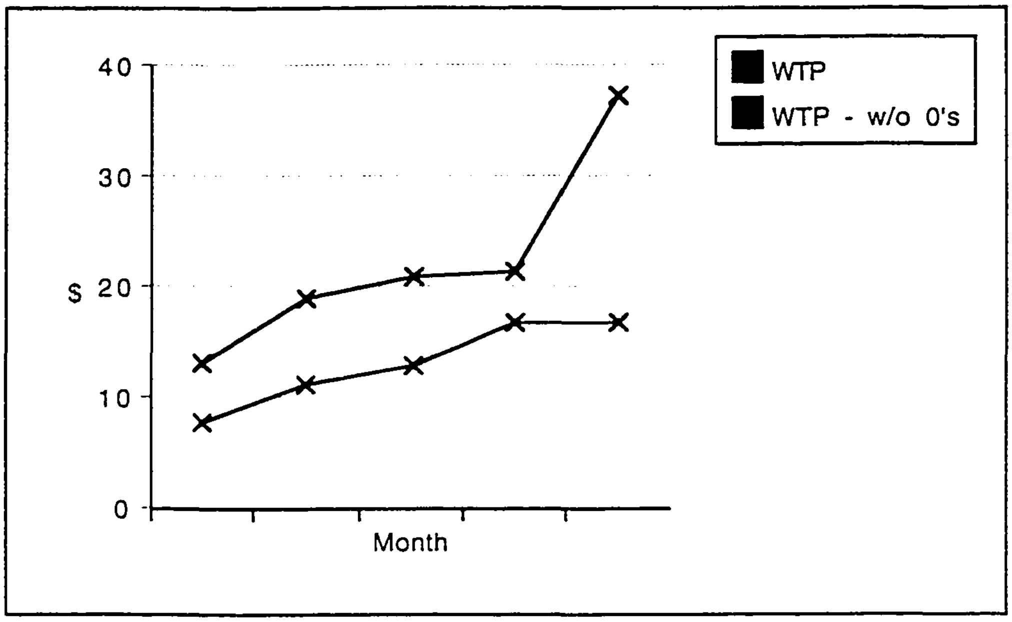


Figure 2.5. WTP - Most Recent Trip: Bluefish

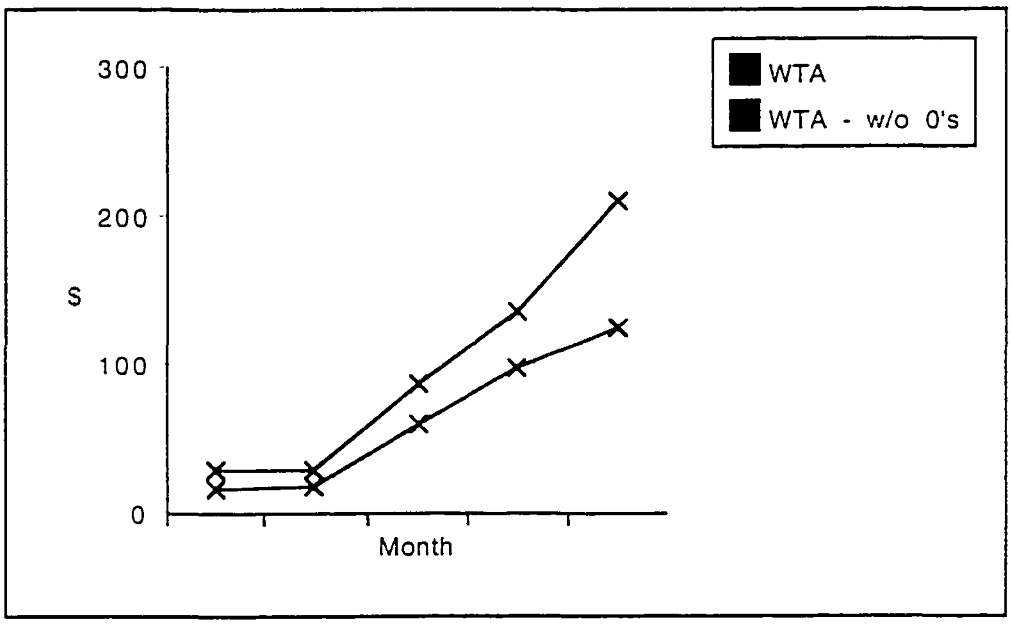


Figure 2.6. WTA - Most Recent Trip: Bluefish

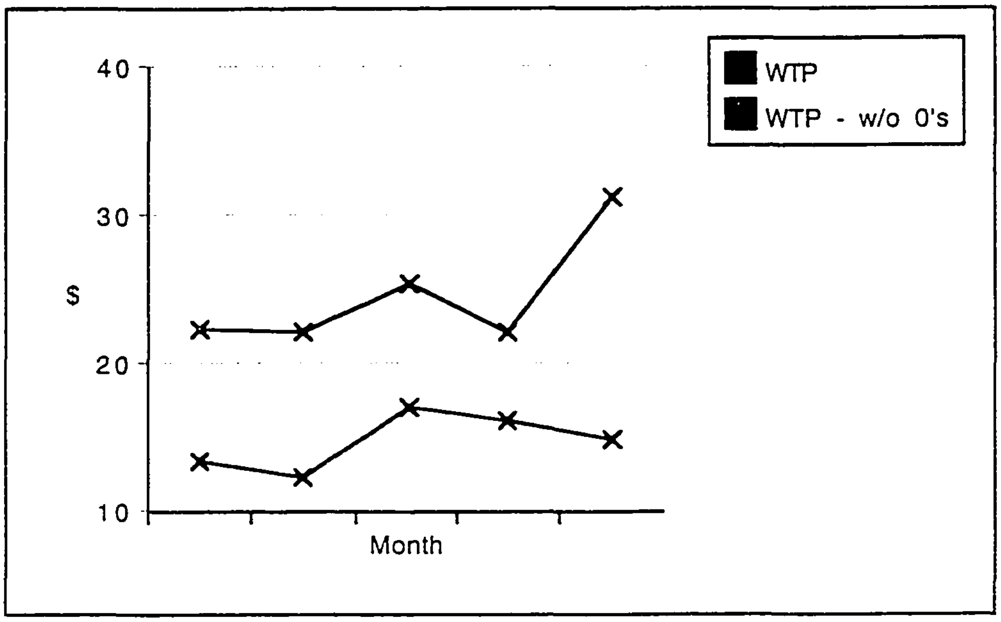


Figure 2.7. WTP - Trip & Catch >= 1: Bluefish

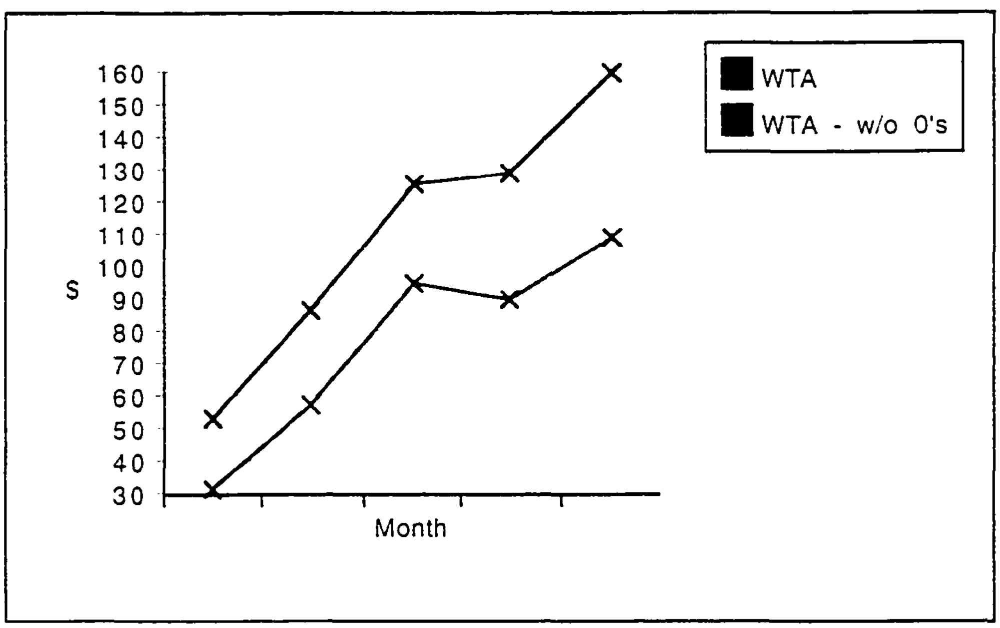


Figure 2.8. WTA - Trip & Catch >= 1: Bluefish

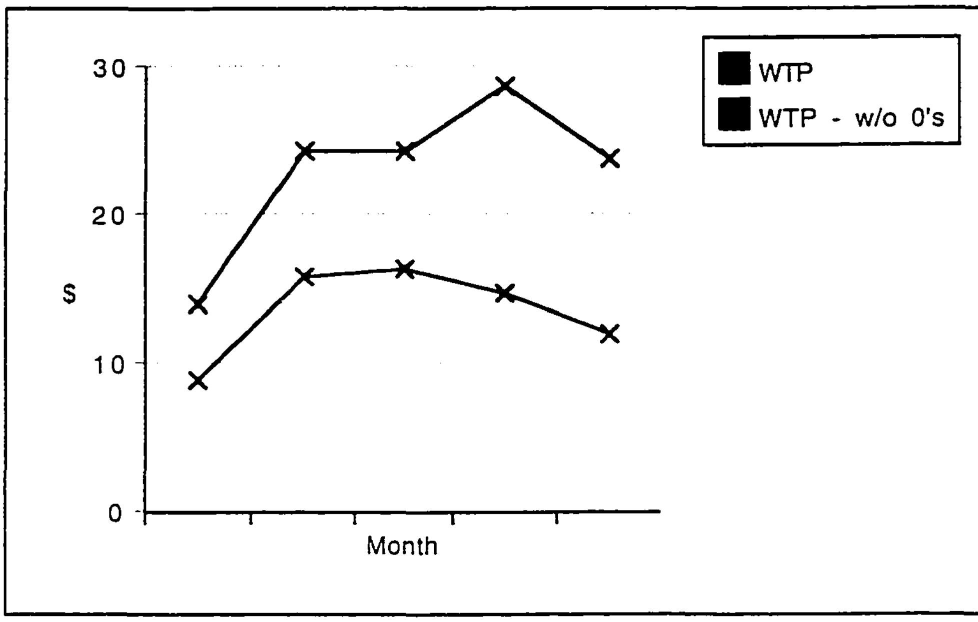


Figure 3.1. WTP - All Individuals: Fluke

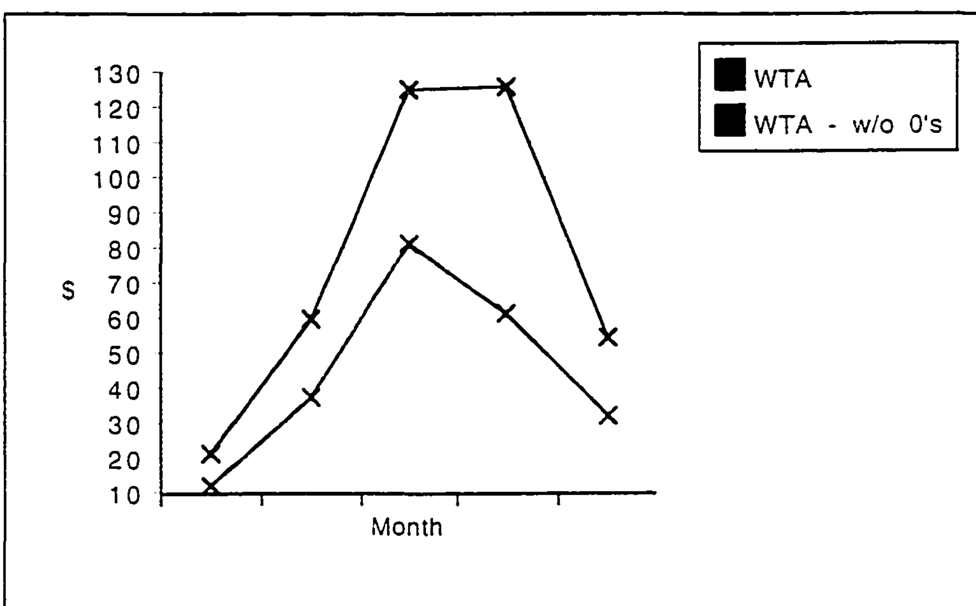


Figure 3.2. WTA - All Individuals: Fluke

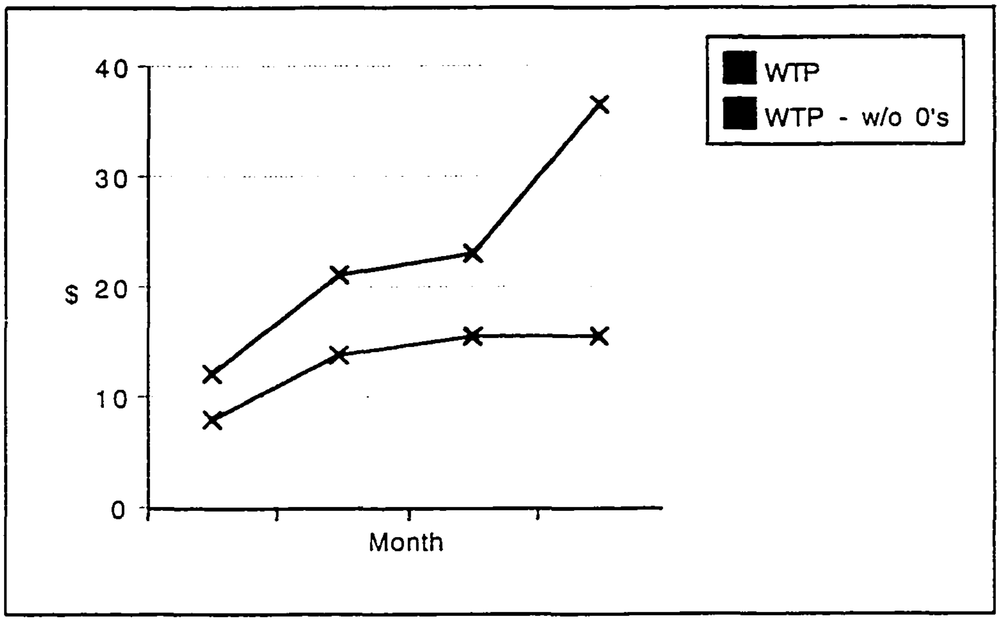


Figure 3.3. WTP - Took Targetted Trip: Fluke

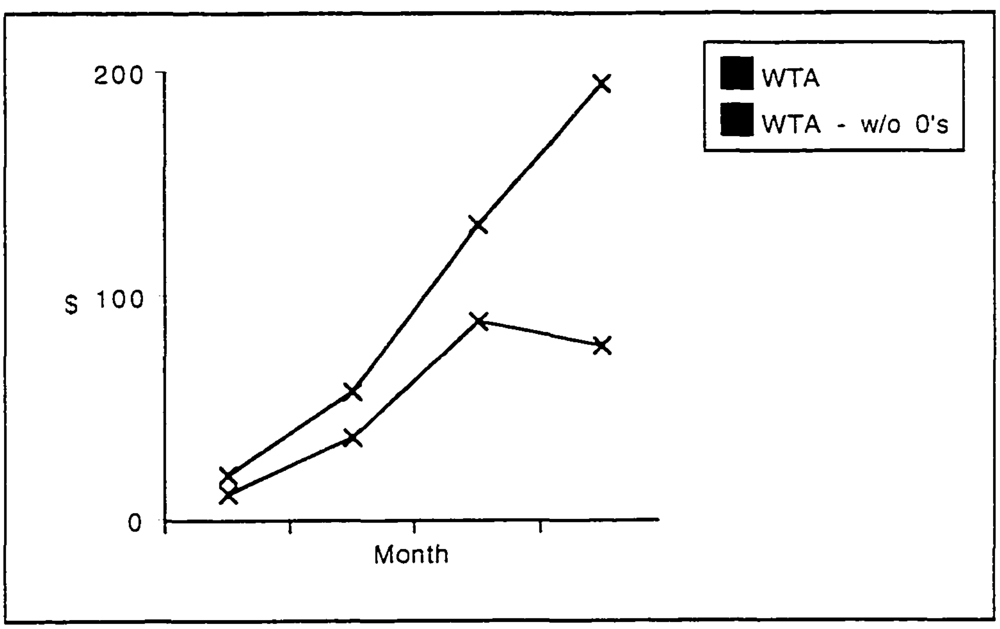


Figure 3.4. WTA - Took Targetted Trip: Fluke

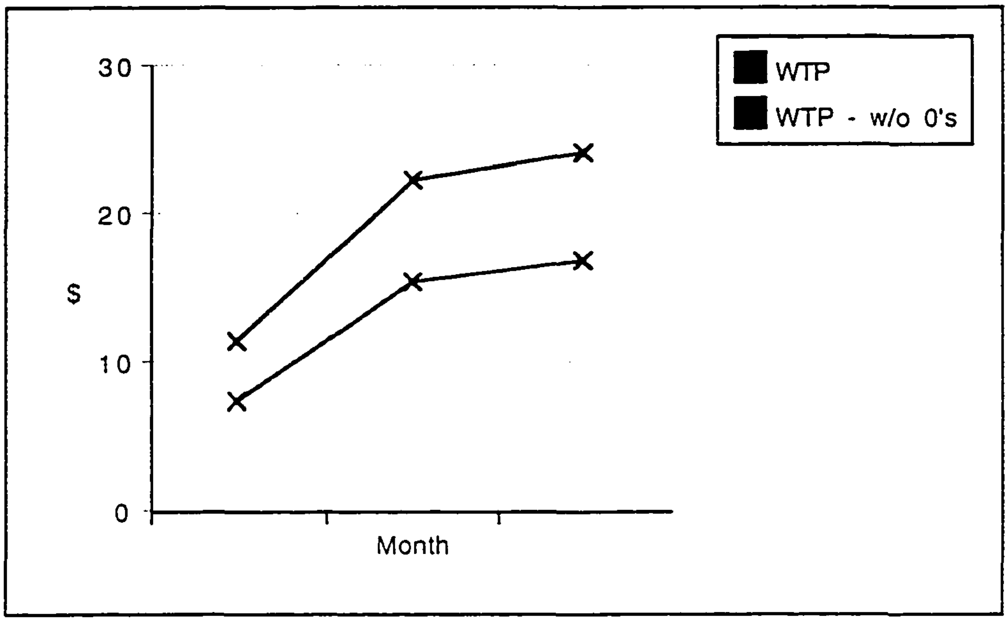


Figure 3.5. WTP - Most Recent Trip: Fluke

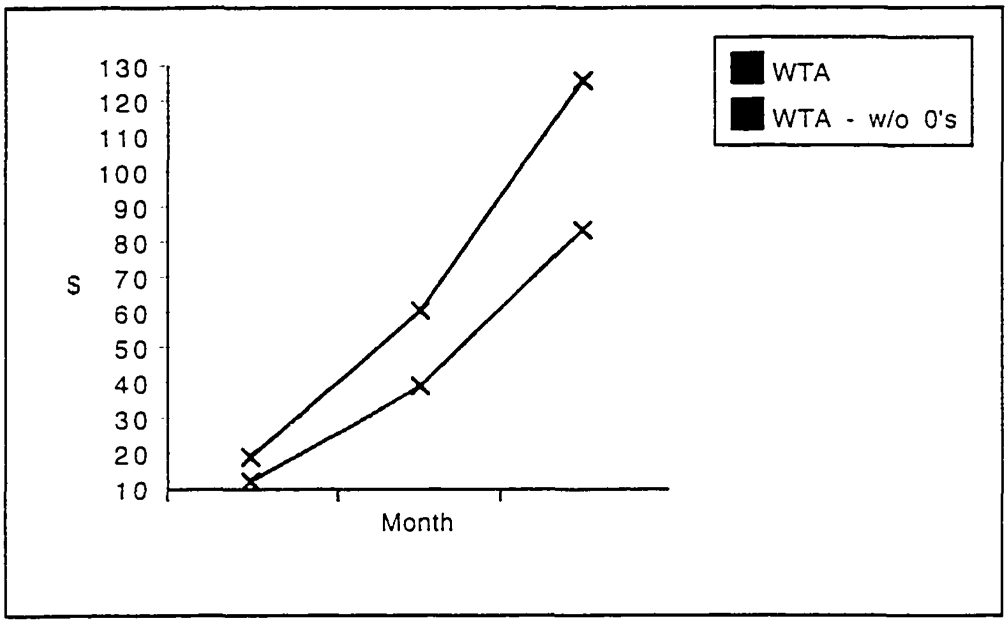


Figure 3.6. WTA - Most Recent Trip: Fluke

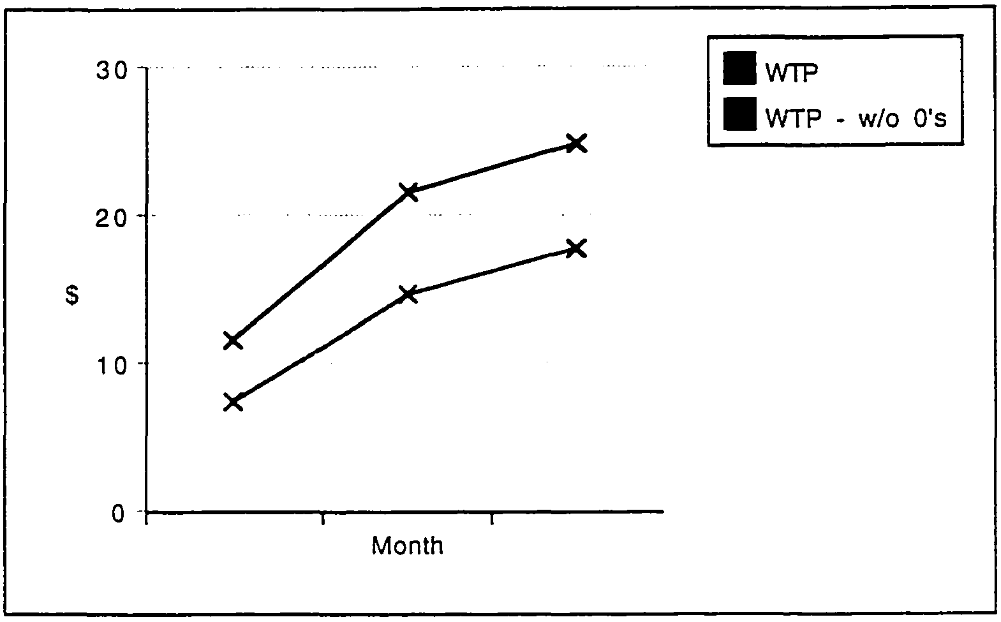


Figure 3.7. WTP - Trip & Catch >= 1: Fluke

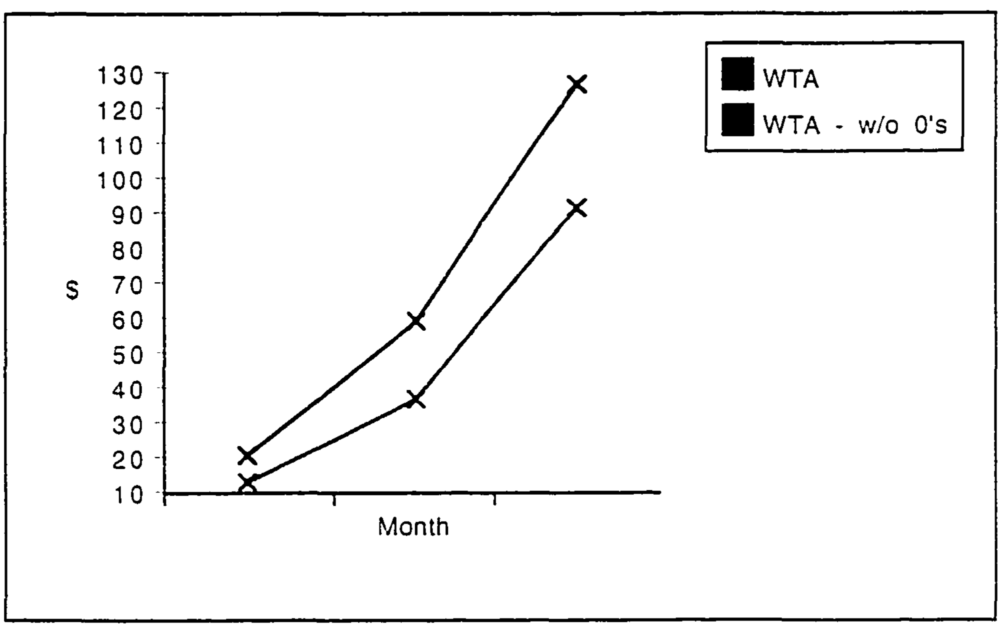


Figure 3.8. WTA - Trip & Catch >= 1: Fluke

Table 3.1. Results of First-Level Statistical Inferences: Hypothesis Tests of Equal Variances and Equal Means Across Subsets of S0 CV Bids - Striped Bass.

	July	August	September	October	November
	---------- F-tests of equal variances ----------				
WTP bids:					
All individuals	-	-	-	**	-
Took targeted trips	-	-	-	-	-
Most recent trip	-	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-
WTA bids:					
All individuals	-	**	-	**	-
Took targeted trips	-	-	-	-	-
Most recent trip	-	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-
	---------- t-tests of equal means ----------				
WTP bids:					
All individuals	*	**	**	**	-
Took targeted trips	-	**	-	**	-
Most recent trip	-	**	-	**	-
Trip & catch ≥ 1	-	**	-	**	-
WTA bids:					
All individuals	-	**	-	**	-
Took targeted trips	-	**	-	**	-
Most recent trip	-	**	-	**	-
Trip & catch ≥ 1	-	**	-	**	-

* Represents significance at .05 level.
** Represents significance at .10 level.

Table 3.2. Results of First-Level Statistical Inferences: Hypothesis Tests of Equal Variances and Equal Means Across Subsets of S0 CV Bids - Bluefish.

	July	August	September	October	November
	F-tests of equal variances				
WTP bids:					
All individuals	-	-	-	-	**
Took targeted trips	-	-	-	-	-
Most recent trip	-	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-
WTA bids:					
All individuals	*	-	-	*	**
Took targeted trips	-	-	-	-	-
Most recent trip	-	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-
	t-tests of equal means				
WTP bids:					
All individuals	**	**	**	**	**
Took targeted trips	-	**	-	-	**
Most recent trip	**	**	-	-	-
Trip & catch ≥ 1	-	**	-	**	-
WTA bids:					
All individuals	-	-	-	-	-
Took targeted trips	-	-	-	-	-
Most recent trip	-	**	-	-	-
Trip & catch ≥ 1	-	-	-	-	-

* Represents significance at .05 level.
** Represents significance at .10 level.

Table 3.3. Results of First-Level Statistical Inferences: Hypothesis Tests of Equal Variances and Equal Means Across Subsets of S0 CV Bids - Fluke.

	July	August	September	October	November
	--------------------- F-tests of equal variances ---------------------				
WTP bids:					
All individuals	-	-	-	-	-
Took targeted trips	-	-	-	-	-
Most recent trip	-	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-
WTA bids:					
All individuals	-	-	-	**	-
Took targeted trips	-	-	-	-	-
Most recent trip	-	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-
	---------------------- t-tests of equal means ----------------------				
WTP bids:					
All individuals	**	**	-	**	-
Took targeted trips	**	**	-	-	**
Most recent trip	**	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-
WTA bids:					
All individuals	**	-	-	-	-
Took targeted trips	-	-	-	-	-
Most recent trip	-	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-

* Represents significance at .05 level.
** Represents significance at .10 level.

Table 3.4. Tests of Equality of AR(1) and AR(2) Models: Inclusion of $0 Bids.

	AR(1) Models			
	Aug.-July	Sept.-Aug.	Oct.-Sept.	Nov.-Oct.
Striped Bass:				
WTP	-	-	**	-
WTA	-	-	-	-
Bluefish:				
WTP	*	-	-	*
WTA	-	-	-	-
Fluke:				
WTP	-	-	-	*
WTA	**	-	-	-

	AR(2) Models		
	Sept.= Aug. + July	Oct. = Sept. + Aug.	Nov. = Oct. + Sept.
Striped Bass:			
WTP	-	-	-
WTA	-	-	-
Bluefish:			
WTP	*	-	-
WTA	*	-	-
Fluke:			
WTP	*	-	-
WTA	-	-	-

* Represents significance at .01 level.
** Represents significance at .05 level.

Table 3.5. Tests of Equality of AR(3) and AR(4) Models: Inclusion of $0 Bids.

	AR(3) Models	
	Oct. = Sept. + Aug. + July	Nov. = Oct. + Sept. + Aug.
Striped Bass:		
WTP	**	-
WTA	-	**
Bluefish:		
WTP	*	-
WTA	-	-
Fluke:		
WTP	*	-
WTA	-	-

	AR(4) Models
	Nov. = Oct. + Sept. + Aug. + July
Striped Bass:	
WTP	-
WTA	-
Bluefish:	
WTP	*
WTA	*
Fluke:	
WTP	-
WTA	-

* Represents significance at .01 level.
** Represents significance at .05 level.

3.2.3. Descriptive Statistics of CV Bids Over Time

The purpose of this next section is to present and discuss estimated mean CV bids over time and examine tests of the equality of mean bids over consecutive time periods based on paired t-tests and equality of mean bids over all time periods. Associated with striped bass mean WTP bids ranged from S13.40/trip to S22.14/trip over all individuals, while mean WTA bids ranged from S18.59/trip to S78.42/trip (Table 3.6). Without S0 bids, these estimates ranged from S21.85/trip to S35.50/trip for WTP and S34.61/trip to S113.27/trip for WTA (Table 3.7). In almost all categories the estimated mean CV bid increased consistently over time (i.e., over each month) (Table 3.6). Overall mean WTP-bids increased from S13.40 to S22.14 (All Individuals), S11.29 to S24.97 (Took Targeted Trips), S13.33 to S24.07 (Most Recent Trip), and from S13.33 to S28.12 (Trip & Catch ≥ 1). Of these categories the greatest increase in bids over time corresponded for those individuals where a striped bass trip was their most recent trip *and* caught at least one striped bass. It seems reasonable to expect that participation and success in catching a fish will directly influence an individuals preferences as reflected in WTP-bid behavior.

As previously mentioned mean CV bids did not differ significantly for most cases. Similar figures are contained in Tables 3.8 to 3.11, for bluefish and fluke. In most all cases mean CV bids increased over time, except for WTP bids for the categories "All Individuals," and "Took Targeted Trips" for bluefish (Tables 3.8 and 3.9).

Hypothesis tests of equal variances across any two consecutive months were conducted and findings revealed that for striped bass variances were significantly different in 26 out of 32 possible cases and that this result was supported based on the subset without $0 bids (Table 3.12). Results for fluke also exhibited a high number of categories with significant differences (Table 3.12). Results for bluefish were opposite (Table 3.12). Hypothesis tests of equal means across consecutive months were conducted with little significant difference found for striped bass CV bids and bluefish CV bids (Tables 3.13). For fluke, mean CV bids were found to be significantly different between July and August, i.e., Bid(t) versus Bid(t+1), and this result carried over to the subset without $0 bids (Table

Table 3.6. Estimates of Mean CV Bids (Bids≥0) for Striped Bass by Month.

Specie: Category	July	August	September	October	November
WTP bids:					
All individuals	13.40	16.49	19.74	21.56	22.14
	(22.07)	(27.34)	(26.53)	(60.62)	(45.32)
	75	72	69	75	93
Took targeted trips	11.29	15.59	17.41	19.17	24.97
	(14.48)	(21.37)	(27.81)	(25.23)	(49.51)
	44	32	32	41	75
Most recent trip	13.83	12.00	10.11	20.93	24.07
	(14.54)	(15.13)	(13.57)	(27.96)	(47.41)
	30	17	19	28	67
Trip & Catch≥1	13.33	23.75	13.67	18.89	28.12
	(15.15)	(26.26)	(11.25)	(25.58)	(52.90)
	18	4	6	18	49
WTA bids:					
All individuals	18.59	47.67	76.41	116.84	78.42
	(35.56)	(137.42)	(188.34)	(321.13)	(169.63)
	67	67	71	75	91
Took targeted trips	16.24	61.94	44.38	162.57	93.61
	(27.65)	(178.38)	(54.66)	(390.33)	(187.33)
	41	31	32	42	72
Most recent trip	17.33	93.53	25.00	136.66	91.08
	(28.56)	(237.12)	(50.25)	(402.56)	(188.72)
	27	17	19	29	65
Trip & Catch≥1	15.87	40.00	61.67	142.94	108.16
	(34.63)	(17.32)	(77.57)	(464.79)	(212.13)
	15	3	6	18	49

Note: Estimated values are means, standard deviations in parentheses, and the number of observations.

Table 3.7. Estimates of Mean CV Bids (Bids>0) for Striped Bass by Month.

Specie: Category	July	August	September	October	November
WTP bids:					
All individuals	21.85	29.68	29.61	38.50	35.50
	(24.75)	(30.97)	(27.66)	(77.23)	(53.22)
	46	40	46	42	58
Took targeted trips	17.75	20.79	25.32	30.23	39.02
	(14.67)	(22.42)	(30.53)	(25.90)	(57.44)
	28	24	22	26	48
Most recent trip	18.86	15.69	19.20	30.84	37.51
	(13.88)	(15.57)	(13.20)	(29.16)	(54.92)
	22	13	10	19	43
Trip & Catch≥1	18.46	31.67	16.40	28.33	39.37
	(14.91)	(25.66)	(10.11)	(26.83)	(59.12)
	13	3	5	12	35
WTA bids:					
All individuals	34.61	86.32	117.94	219.08	113.27
	(42.60)	(176.61)	(223.99)	(451.58)	(194.31)
	36	37	46	40	63
Took targeted trips	27.75	101.05	71.00	243.86	129.62
	(31.57)	(221.05)	(53.72)	(459.14)	(209.99)
	24	19	20	28	52
Most recent trip	29.25	144.55	67.86	208.58	128.69
	(32.28)	(286.09)	(64.67)	(486.06)	(213.78)
	16	11	7	19	46
Trip & Catch≥1	34.00	40.00	92.50	197.92	135.90
	(45.60)	(17.32)	(78.89)	(542.45)	(230.16)
	7	3	4	13	39

Note: Estimated values are means, standard deviations in parentheses, and the number of observations.

Table 3.8. Estimates of Mean CV Bids (Bids≥0) for Bluefish by Month.

Specie: Category	July	August	September	October	November
<u>WTP bids</u>:					
All individuals	15.13	13.33	13.79	10.33	12.81
	(30.48)	(23.11)	(22.27)	(18.76)	(29.55)
	87	94	89	79	67
Took targeted trips	18.09	14.97	16.11	10.63	9.31
	(34.46)	(24.94)	(24.22)	(17.70)	(19.42)
	64	69	65	48	39
Most recent trip	7.76	11.11	12.97	16.74	16.75
	(10.36)	(15.99)	(19.83)	(21.77)	(26.67)
	34	36	37	23	20
Trip & Catch≥1	13.35	12.29	17.13	16.15	15.00
	(23.83)	(20.69)	(26.52)	(21.97)	(41.91)
	37	38	40	26	23
<u>WTA bids</u>:					
All individuals	28.53	38.82	72.65	63.13	61.21
	(92.79)	(94.15)	(178.72)	(184.69)	(177.74)
	89	88	89	75	70
Took targeted trips	39.63	42.16	78.85	100.18	88.93
	(109.52)	(106.22)	(188.69)	(234.52)	(226.64)
	62	64	65	44	41
Most recent trip	16.79	17.71	59.89	97.62	125.00
	(38.06)	(23.05)	(114.68)	(232.82)	(287.15)
	33	35	38	21	22
Trip & Catch≥1	31.59	57.18	95.15	90.00	108.80
	(87.47)	(134.44)	(189.11)	(223.43)	(272.11)
	37	38	41	23	25

Note: Estimated values are means, standard deviations in parentheses, and the number of observations.

Table 3.9. Estimates of Mean CV Bids (Bids>0) for Bluefish by Month.

Specie: Category	July	August	September	October	November
WTP bids:					
All individuals	25.80	24.10	23.62	18.98	28.60
	(36.29)	(26.63)	(24.88)	(22.03)	(38.99)
	51	52	52	43	30
Took targeted trips	28.95	24.02	24.35	17.59	22.69
	(39.94)	(28.00)	(26.24)	(19.98)	(25.13)
	40	43	43	29	16
Most recent trip	13.20	19.05	20.87	21.39	37.22
	(10.53)	(16.70)	(21.71)	(22.54)	(28.84)
	20	21	23	18	9
Trip & Catch≥1	22.45	22.24	25.37	22.11	31.36
	(27.56)	(23.66)	(28.94)	(23.05)	(57.44)
	22	21	27	19	11
WTA bids:					
All individuals	57.70	65.69	109.59	115.49	109.87
	(126.10)	(115.42)	(210.58)	(238.55)	(227.77)
	44	52	59	41	39
Took targeted trips	70.20	71.00	111.41	157.43	165.73
	(138.96)	(130.79)	(216.64)	(279.72)	(290.69)
	35	38	46	28	22
Most recent trip	29.16	29.52	87.54	136.67	211.54
	(46.85)	(23.18)	(130.18)	(267.80)	(352.80)
	19	21	26	15	13
Trip & Catch≥1	53.14	86.92	125.84	129.38	160.00
	(109.17)	(158.66)	(209.02)	(260.36)	(319.94)
	22	25	31	16	17

Note: Estimated values are means, standard deviations in parentheses, and the number of observations.

Table 3.10. Estimates of Mean CV Bids (Bids≥0) for Fluke by Month.

Species: Category	July	August	September	October	November
WTP bids:					
All individuals	8.89	15.85	16.29	14.64	11.94
	(14.23)	(23.29)	(35.77)	(26.39)	(21.65)
	119	106	87	47	34
Took targeted trips	8.09	13.89	15.69	15.71	-
	(11.03)	(20.71)	(36.37)	(29.99)	-
	106	90	72	14	-
Most recent trip	7.42	15.54	16.87	20.00	-
	(10.16)	(23.15)	(41.82)	(44.72)	-
	79	59	53	5	-
Trip & Catch≥1	7.51	14.79	17.86	24.17	-
	(9.51)	(22.60)	(44.68)	(38.78)	-
	76	63	49	6	-
WTA bids:					
All individuals	12.00	37.90	81.69	61.70	32.26
	(22.73)	(83.32)	(206.28)	(194.61)	(82.41)
	112	90	86	47	34
Took targeted trips	11.71	37.92	89.00	78.00	-
	(22.17)	(85.38)	(221.08)	(256.10)	-
	98	77	70	15	-
Most recent trip	12.07	39.48	84.06	200.00	-
	(22.35)	(100.68)	(221.17)	(447.21)	-
	71	48	51	5	-
Trip & Catch≥1	12.71	36.54	91.59	155.71	-
	(22.75)	(97.19)	(226.06)	(372.73)	-
	69	52	47	7	-

Note: Estimated values are means, standard deviations in parentheses, and the number of observations.

Table 3.11. Estimates of Mean CV Bids (Bids>0) for Fluke by Month.

Specie: Category	July	August	September	October	November
WTP bids:					
All individuals	13.93	24.35	24.43	28.67	23.88
	(15.73)	(25.06)	(41.57)	(31.20)	(25.77)
	76	69	58	24	17
Took targeted trips	12.26	21.19	23.06	36.67	-
	(11.54)	(22.38)	(42.22)	(37.64)	-
	70	59	49	6	-
Most recent trip	11.49	22.37	24.16	-	-
	(10.65)	(24.91)	(48.42)	-	-
	51	41	37	-	-
Trip & Catch≥1	11.65	21.67	25.00	48.33	-
	(9.60)	(24.52)	(51.33)	(44.81)	-
	49	43	35	3	-
WTA bids:					
All individuals	21.68	59.84	125.45	126.09	54.85
	(26.95)	(98.47)	(245.30)	(265.89)	(102.43)
	62	57	56	23	20
Took targeted trips	20.50	58.40	132.55	195.00	-
	(26.14)	(100.43)	(259.65)	(395.31)	-
	56	50	47	6	-
Most recent trip	19.48	61.13	126.09	-	-
	(25.80)	(120.47)	(262.02)	-	-
	44	31	34	-	-
Trip & Catch≥1	20.88	59.38	126.62	-	-
	(26.14)	(118.92)	(261.86)	-	-
	42	32	34	-	-

Note: Estimated values are means, standard deviations in parentheses, and the number of observations.

Table 3.12. Results of First-Level Statistical Inferences: Hypothesis Tests of Equal Variances Across Time - Striped Bass. Bluefish. and Fluke.

	July-Aug.		Aug.-Sept.		Sept.-Oct.		Oct.-Nov.	
	Full	w/o 0's	Full	w/o 0's	Full	w/o 0's	Full	w/o 0's
Striped Bass:								
WTP bids:								
All individuals	***	-	***	-	*	-	**	*
Took targeted trips	**	**	-	-	-	*	*	*
Most recent trip	-	-	-	-	*	**	*	**
Trip & catch ≥ 1	-	-	***	-	***	***	*	*
WTA bids:								
All individuals	*	*	**	-	*	*	*	*
Took targeted trips	*	*	*	*	*	*	*	*
Most recent trip	*	*	*	*	*	*	*	*
Trip & catch ≥ 1	-	-	**	***	*	***	*	*
Bluefish:								
WTP bids:								
All individuals	**	***	-	-	-	-	*	*
Took targeted trips	**	**	-	-	**	-	-	-
Most recent trip	**	**	-	-	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-	-	*	-
WTA bids:								
All individuals	-	-	*	*	-	-	-	-
Took targeted trips	-	-	*	*	-	-	-	-
Most recent trip	**	*	*	*	*	*	-	-
Trip & catch ≥ 1	**	-	-	-	-	-	-	-
Fluke:								
WTP bids:								
All individuals	*	*	*	*	**	-	-	-
Took targeted trips	*	*	*	*	-	-	NA	NA
Most recent trip	*	*	*	*	NA	NA	NA	NA
Trip & catch ≥ 1	*	*	*	*	NA	NA	NA	NA
WTA bids:								
All individuals	*	*	*	*	-	-	*	*
Took targeted trips	*	*	*	*	-	-	NA	NA
Most recent trip	*	*	*	*	NA	NA	NA	NA
Trip & catch ≥ 1	*	-	*	-	*	-	NA	NA

NA refers to not available, too few observations were present.
* Represents significance at .01 level.
** Represents significance at .05 level.
*** Represents significance at .10 level.

Table 3.13. Results of First-Level Statistical Inferences: Hypothesis Tests of Equal Means Across Time - Striped Bass, Bluefish, and Fluke.

	July-Aug.		Aug.-Sept.		Sept.-Oct.		Oct.-Nov.	
	Bid>=0	Bid>0	Bid>=0	Bid>0	Bid>=0	Bid>0	Bid>=0	Bid>0
Striped Bass:								
WTP bids:								
All individuals	-	-	-	-	-	-	-	-
Took targeted trips	-	-	-	-	-	-	-	-
Most recent trip	-	-	-	-	**	-	-	-
Trip & catch ≥ 1	-	-	-	-	-	-	-	-
WTA bids:								
All individuals	**	**	-	-	-	-	-	-
Took targeted trips	-	-	-	-	**	**	-	-
Most recent trip	-	-	-	-	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-	-	-	-
Bluefish:								
WTP bids:								
All individuals	-	-	-	-	-	-	-	-
Took targeted trips	-	-	-	-	-	-	-	-
Most recent trip	-	-	-	-	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-	-	-	-
WTA bids:								
All individuals	-	-	-	-	-	-	-	-
Took targeted trips	-	-	-	-	-	-	-	-
Most recent trip	-	-	**	**	-	-	-	-
Trip & catch ≥ 1	-	-	-	-	-	-	-	-
Fluke:								
WTP bids:								
All individuals	**	**	-	-	-	-	-	-
Took targeted trips	**	**	-	-	-	-	NA	NA
Most recent trip	**	**	-	-	-	NA	NA	NA
Trip & catch ≥ 1	**	**	-	-	-	NA	NA	NA
WTA bids:								
All individuals	**	**	**	**	-	-	-	-
Took targeted trips	**	**	**	**	-	-	NA	NA
Most recent trip	**	**	-	-	-	NA	NA	NA
Trip & catch ≥ 1	**	-	-	-	-	NA	NA	NA

NA refers to not available, too few observations were present.
* Represents significance at .05 level.
** Represents significance at .10 level.

3.13).

Tests of the equality of all means over time was conducted based on a Duncan's Multiple Range test. Not surprisingly results indicated few significant differences over all means (Tables 3.14 - 3.16).

3.3. Implications of Various Subsets of Suspect Observations

The results that examined the implications of various subsets of outliers are presented for WTP models associated with striped bass. Subsets of outliers involved observations where CV bids 1) were large relative to income, 2) were either small relative to income or exhibited a pattern of a zero bid followed by a nonzero bid, followed by a zero bid and then a nonzero bid (e.g., \$0, \$10, \$0, \$10) - these observations were treated as noise whereby they could introduce excessive variability into the data, 3) a combination of large and noise, or 4) were identified by the technique developed by Belsley *et al.* (1980). These subsets of outliers were tested similar in spirit as tests of structural change in the data to determine whether data can be pooled versus not based on Chow-tests. Such tests were first used to examine the issue of structural change in data over time (Greene 1993). The specific procedure was outlined in Chapter 2.

Results are contained in Tables 3.17 and 3.18. AR(1) models were first estimated and equality of the regression model was tested for the various subsets of suspect observations. Results indicated that for AR(1) models the data subsets "Noise," "Large," and "Large & Noise" were significantly different for the October and September time periods. AR(2) model results showed that the data subset "Large" corresponding to large CV bids was significant for all possible time periods (Table 3.17). The data subset of large CV bids was also found to be significant in AR(3) and AR(4) models (Table 3.18). In addition, observations identified using the technique of Belsley *et al.* (1980) were significant for one of the AR(3) models and for the AR(4) models. Although significant in the AR(3) and AR(4) models, "Influence" did not contain a significant subset of CV bids in the case of

Table 3.14. Results of First-Level of Statistical Inferences: Hypothesis Tests of Equality of All Mean CV Bids (Bids≥0 & Bids>0) Across Successive Months: Striped Bass.

Specie: Category	July	August	September	October	November
			BIDS≥0		
WTP bids:					
All individuals	13.40a	16.49a	19.74a	21.56a	22.14a
Took targeted trips	11.29a	15.59a	17.41a	19.17a	24.97a
Most recent trip	13.83a	12.00a	10.11a	20.93a	24.07a
Trip & Catch≥1	13.33a	23.75a	13.67a	18.89a	28.12a
WTA bids:					
All individuals	18.59a	47.67ab	76.41bc	116.84c	78.42bc
Took targeted trips	16.24a	61.94a	44.38a	162.57b	93.61ab
Most recent trip	17.33a	93.53a	25.00a	136.66b	91.08ab
Trip & Catch≥1	15.87a	40.00a	61.67a	142.94a	108.16a
			BIDS>0		
WTP bids:					
All individuals	21.85a	29.68a	29.61a	38.50a	35.50a
Took targetted trips	17.75a	20.79ab	25.32ab	30.23ab	39.02b
Most recent trip	18.86a	15.69a	19.20a	30.84a	37.51a
Trip & Catch≥1	18.46a	31.67a	16.40a	28.33a	39.37a
WTA bids:					
All individuals	34.61a	86.32a	117.93a	219.08b	113.27a
Took targetted trips	27.75a	101.05a	71.00a	243.86b	129.62ab
Most recent trip	29.25a	144.55a	67.86a	208.58a	128.69a
Trip & Catch≥1	34.00a	40.00a	92.50a	197.92a	135.90a

Note: A Duncan's Multiple Range test was used to test equality of means across successive months. Means with different letters across months are significantly different from one another at the .10 level.

Table 3.15. Results of First-Level of Statistical Inferences: Hypothesis Tests of Equality of All CV Bids (Bids≥0 & Bids>0) Across Successive Months: Bluefish.

Specie: Category	July	August	September	October	November
			BIDS≥0		
WTP bids:					
All individuals	15.13a	13.33a	13.79a	10.33a	12.81a
Took targetted trips	18.09a	14.97a	16.11a	10.63a	9.31a
Most recent trip	7.76a	11.11a	12.97a	16.74a	16.75a
Trip & Catch≥1	13.35a	12.29a	17.13a	16.15a	15.00a
WTA bids:					
All individuals	28.53a	38.82ab	72.65b	63.13ab	61.21ab
Took targetted trips	39.63a	42.16a	78.85a	100.18a	88.93a
Most recent trip	16.79a	17.71a	59.89ab	97.62b	125.00b
Trip & Catch≥1	31.59a	57.18a	95.15a	90.00a	108.80a
			BIDS>0		
WTP bids:					
All individuals	25.80a	24.10a	23.62a	18.98a	28.60a
Took targetted trips	28.95a	24.02a	24.35a	17.59a	22.69a
Most recent trip	13.20	19.05	20.87	21.39	37.22
Trip & Catch≥1	22.45a	22.24a	25.37a	22.11a	31.36a
WTA bids:					
All individuals	57.70a	65.69a	109.59a	115.49a	109.87a
Took targetted trips	70.20a	71.00a	111.41a	157.43a	165.73a
Most recent trip	29.16a	29.52a	87.54a	136.67ab	211.54b
Trip & Catch≥1	53.14a	86.92a	125.84a	129.38a	160.00a

Note: A Duncan's Multiple Range test was used to test equality of means across successive months. Means with different letters across months are significantly different from one another at the .10 level.

Table 3.16. Results of First-Level of Statistical Inferences: Hypothesis Tests of Equality of All CV Bids (Bids≥0 & Bids>0) Across Successive Months: Fluke.

Specie: Category	July	August	September	October	November
			BIDS≥0		
WTP bids:					
All individuals	8.89a	15.85a	16.29a	14.64	11.94
Took targetted trips	8.09a	13.89a	15.69a	15.71	-
Most recent trip	7.42a	15.54a	16.87a	20.00a	-
Trip & Catch≥1	7.51a	15.03ab	18.17ab	24.17b	-
WTA bids:					
All individuals	12.00a	37.90ab	81.69c	61.70bc	32.26ab
Took targetted trips	11.71a	37.92ab	89.00b	78.00b	-
Most recent trip	12.07a	39.48b	46.67b	-	-
Trip & Catch≥1	12.67a	36.04a	91.59ab	155.71b	-
			BIDS>0		
WTP bids:					
All individuals	13.93a	24.35ab	24.43ab	28.67b	23.88ab
Took targetted trips	12.26a	21.19ab	23.06ab	36.67b	-
Most recent trip	11.49a	22.37ab	24.16b	-	-
Trip & Catch≥1	11.65a	22.37ab	25.65b	-	-
WTA bids:					
All individuals	21.68a	59.84ab	125.45b	126.09b	54.85a
Took targetted trips	20.50a	58.40ab	132.55bc	195.00c	-
Most recent trip	19.48a	61.13b	71.47b	-	-
Trip & Catch≥1	20.63a	57.88a	123.09b	-	-

Note: A Duncan's Multiple Range test was used to test equality of means across successive months. Means with different letters across months are significantly different from one another at the .10 level.

Table 3.17. Tests of Equality of AR(1) and AR(2) Models for WTP Series Across Subsets of Outliers - Striped Bass.

	AR(1) Models			
	Aug.-July	Sept.-Aug.	Oct.-Sept.	Nov.-Oct.
w/o Noise	-	-	*	-
w/o Large	-	-	*	-
w/o Large + Noise	-	-	**	-
w/o Influence	-	-	-	-

	AR(2) Models		
	Sept.= Aug. + July	Oct. = Sept. + Aug.	Nov. = Oct. + Sept.
w/o Noise	-	-	-
w/o Large	*	*	**
w/o Large + Noise	*	-	-
w/o Influence	-	-	-

* Represents significance at .01 level.
** Represents significance at .05 level.

Table 3.18. Tests of Equality of AR(3) and AR(4) Models for WTP Series Across Subsets of Outliers - Striped Bass.

	AR(3) Models	
	Oct. = Sept. + Aug. + July	Nov. = Oct. + Sept. + Aug.
w/o Noise	*	-
w/o Large	*	**
w/o Large + Noise	*	-
w/o Influence	*	-

	AR(4) Models
	Nov. = Oct. + Sept. + Aug. + July
w/o Noise	-
w/o Large	**
w/o Large + Noise	-
w/o Influence	**

* Represents significance at .01 level.
** Represents significance at .05 level.

AR(1) and AR(2) models.

This evidence implies that for the purposes of this investigation extreme value observations that corresponded to large WTP bids introduced a significant amount of variability in the CV bid data. and that other types of observations proved relatively inconsequential. Such evidence suggests that CV bid data may contain particular characteristics that are unique to any one study and that it is useful to search for this characteristic.

3.4. Implications of Alternative AR Models

A concern in empirical research regards how robust the results are across alternative models. This usually focuses on functional form (linear versus log models) and could include alternative types of data structure (e.g.. time-series models associated with successive time periods versus pooled time-series data) as well as specific AR model forms (AR models with versus without a drift term. trend term, or both). In this section the latter two issues are examined for all three target species (striped bass, bluefish, and fluke) and whether or not the data includes $0 CV bids. An F-test was used to examine whether a restricted AR model (i.e.. an AR model without one of the following terms: drift, trend, or lagged independent variables or some combination) represented the true model versus an unrestricted model (a model with lagged independent variables, drift and trend terms. i.e., a full model).

Concerning AR(1) models estimated over succesive time periods, referred to as "successive AR models", the null hypothesis was that a restricted model (i.e., AR(1) without drift, AR(1) with only a constant term--drift) was true versus an unrestricted model (AR(1) model with drift) was true. Results for the WTP series indicated that over all three species 11 cases out of 12 possible cases for bids with $0's rejected the null that the AR process follows a drift term alone for both linear and log models (Tables 3.19 and 3.20). This result was fairly robust for bids without $0's. Results were mixed concerning whether or not a drift term should be included in the AR models for the WTP series; 6 of 12 cases

Table 3.19. Tests of Restricted Models: Successive AR(1) Linear Models.

	Aug. = July R1	Aug. = July R2	Sept. = Aug. R1	Sept. = Aug. R2	Oct. = Sept. R1	Oct. = Sept. R2	Nov. = Oct. R1	Nov. = Oct. R2
				WTP				
Striped Bass:								
Bids ≥ 0	-	*	*	*	-	*	*	*
Bids > 0	***	-	**	**	**	*	*	*
Bluefish:								
Bids ≥ 0	*	-	**	*	-	*	-	*
Bids > 0	**	-	**	**	-	*	-	*
Fluke:								
Bids ≥ 0	*	**	-	*	-	*	-	*
Bids > 0	*	***	-	*	-	-	-	*
				WTA				
Striped Bass:								
Bids ≥ 0	-	-	***	-	-	***	***	*
Bids > 0	-	-	-	-	-	-	-	**
Bluefish:								
Bids ≥ 0	***	-	-	*	-	*	-	*
Bids > 0	-	-	-	*	-	*	-	*
Fluke:								
Bids ≥ 0	***	***	-	*	-	*	-	*
Bids > 0	-	-	-	*	-	*	-	*

Note: The unrestricted AR(1) model is $y(t) = \mu + \beta y(t-1) + u(t)$, where μ is a drift term. R1 refers to the restricted AR(1) model: $y(t) = \beta y(t-1) + u(t)$ and R2 refers to the restricted AR(1) model: $y(t) = \mu$. Null hypothesis is that the restricted model is true versus the alternative hypothesis that the unrestricted model is true.

* Represents significance at .01 level.

** Represents significance at .05 level.

*** Represents significance at .10 level.

Table 3.20. Tests of Restricted Models: Successive AR(1) Log Models.

	Aug. = July R1	Aug. = July R2	Sept. = Aug. R1	Sept. = Aug. R2	Oct. = Sept. R1	Oct. = Sept. R2	Nov. = Oct. R1	Nov. = Oct. R2
	WTP							
Striped Bass:								
Bids ≥ 0	***	*	**	*	-	*	-	*
Bids > 0	-	**	**	*	-	***	**	*
Bluefish:								
Bids ≥ 0	*	*	-	*	-	*	-	*
Bids > 0	*	-	**	*	**	*	**	*
Fluke:								
Bids ≥ 0	*	*	**	*	-	*	-	*
Bids > 0	*	**	***	*	-	-	-	**
	WTA							
Striped Bass:								
Bids ≥ 0	-	*	*	*	-	*	**	*
Bids > 0	**	-	-	**	-	-	-	*
Bluefish:								
Bids ≥ 0	*	*	*	*	-	*	-	*
Bids > 0	*	-	***	*	-	*	-	*
Fluke:								
Bids ≥ 0	*	*	**	*	-	*	***	*
Bids > 0	*	*	*	*	-	*	**	*

Note: The unrestricted AR(1) model is $y(t) = \mu + \beta y(t-1) + u(t)$, where μ is a drift term. R1 refers to the restricted AR(1) model: $y(t) = \beta y(t-1) + u(t)$ and R2 refers to the restricted AR(1) model: $y(t) = \mu$. Null hypothesis is that the restricted model is true versus the alternative hypothesis that the unrestricted model is true.

* Represents significance at .01 level.

** Represents significance at .05 level.

*** Represents significance at .10 level.

indicated it should for bids with S0's and 8 of 12 for bids without S0's for linear models. results for log models were similar. Specific results for individual species tends to reflect these findings. no one specie dominated the distribution of rejected null occurrences.

Regarding the WTA series the null that the process follows a drift term alone was rejected for a majority of cases: the log model indicated more rejections compared to the linear model (Tables 3.19 and 3.20). For this series the inclusion of S0 bids increased the number of rejected nulls. Considering the null of a lagged independent variable only (without drift) results were mixed: more instances of rejections occurred for log models compared to linear models. 7 of 12 for bids with S0's and 6 without S0's compared to 4 instances for linear models with S0's and none when S0 bids were dropped (Tables 3.19, 3.20).

Overall. findings support that the successive AR models do not follow a drift term alone for both WTP and WTA series, but findings were mixed whether the WTP series follows an AR process with a drift term, and for the WTA series the process seems to follow an AR process without a drift term.

This analysis was extended to examine if results would differ based on pooled time series data. The unrestricted model was that the AR process follows an AR(1) model with drift and trend terms (where trend terms were represented by dummy variables associated with each month a bid occurred) compared to the null that a restricted model was true. Four types of restricted models were compared, R1 an AR(1) model without trend, R2 an AR(1) model without trend and drift terms, R3 a model with drift alone, and R4 a model with drift and trend. Both a typical AR(1) model is examined and an AR(1) model based on first differences, a $\Delta y(t)$ AR model (i.e., $y(t) - y(t-1) = \mu + \beta y(t-1) + DT + u(t)$, where μ is the drift term and DT represents a trend component).

For the WTP series associated with the typical AR model (y(t)-model) the latter two restricted models (R3 and R4) are rejected in all cases for both linear and log specifications and whether or not $0 bids are included: the series when pooled does not follow an AR process with a drift term alone nor an AR process with only drift and trend components (Table 3.21 and 3.22). This result also holds for the WTA series across linear and log specifications and over bids with or without $0's. Results petaining to the $\Delta y(t)$ linear models only indicate that for striped bass and bluefish both the R3 and R4 restricted models are rejected across subsets of $0 bids for the WTP series and only for striped bass for the WTA series. Results of log models did not support rejection of the null.

Pertaining to the other restricted AR-$y(t)$ models (R1-with drift but without trend, and R2-without drift and trend) findings over all three species indicate that the null associated with the R1 restricted model can not be rejected in all but one case (WTA-Bluefish with $0 bids). This is strong evidence that a trend component or a seasonal component is not present in this CV data. Results are identical for the AR-$\Delta y(t)$ models. In addition, results indicate that functional specification can influence whether an AR model should contain a drift and trend term; for all but one case (WTP-log model, striped bass without $0's) the null is rejected for the log specification with the results robust over subsets of $0 bids for both WTP and WTA series and over AR-$\Delta y(t)$ models (Table 3.22). However, for linear models the null is rejected only for Bluefish with $0 bids for the WTP series and is rejected for both striped bass and bluefish without $0 bids for WTP. Concerning the WTA series for linear models, the null is rejected for bluefish and fluke associated with bids that include $0's but only rejected for fluke without $0 bids for WTA (Table 3.21). This evidence implies that once the drift term is dropped from an AR model without trend, the null is rejected for log specifications and is mixed for linear specifications; log models require a drift term for both WTP and WTA series but for linear models whether a drift term should be included is mixed. If one accepts the notion that CV bids may approach some "true" value other than $0 for a long-run equilibrium then a drift term is necessary, which implies that the mean value of the dependent variable will be non-zero over time.

In general the results based on these model comparisons indicate that 1) successive time

Table 3.21. Tests of Restricted Models: Pooled AR(1) Linear Models.

	y(t)				Δy(t)			
	R1	R2	R3	R4	R1	R2	R3	R4
	WTP							
Striped Bass:								
Bids ≥ 0	-	-	*,a	*,a	-	-	*,a	*,a
Bids > 0	-	*	*,b	*,a	-	***	*,b	*,a
Bluefish:								
Bids ≥ 0	-	*	*,a	*,a	-	*	*,a	*,a
Bids > 0	-	*	*,b	*,a	-	*	*,c	*,a
Fluke:								
Bids ≥ 0	-	-	*,a	*,a	-	-	-	-
Bids > 0	-	-	*,a	*,a	-	-	-	-
	WTA							
Striped Bass:								
Bids ≥ 0	-	-	*,c	*,a	-	-	*,a	*,a
Bids > 0	-	-	-	*,b	-	*	*,a	*,a
Bluefish:								
Bids ≥ 0	-	**	*,a	*,a	-	**	-	-
Bids > 0	-	-	*,a	*,a	-	-	-	-
Fluke:								
Bids ≥ 0	-	*	*,a	*,a	-	*	-	-
Bids > 0	-	***	*,a	*,a	-	***	-	-

Note: The unrestricted AR(1) model: $y(t) = \mu + \beta y(t-1) + DT + u(t)$, where μ is a drift term. DT refers to trend terms, here dummy variables associated with each month (the drift term is dropped when all dummy variables are included). R1 refers to the restricted AR(1) model: $y(t) = \mu + \beta y(t-1) + u(t)$, R2 refers to: $y(t) = \beta y(t-1) + u(t)$, R3 refers to: $y(t) = \mu + u(t)$, and R4 refers to: $y(t) = DJ + DA + DS + DO + u(t)$, a model with drift and trend terms. Null hypothesis is that the restricted model is true versus the alternative hypothesis that the unrestricted model is true.

* Represents significance at .01 level based on a t-test.
** Represents significance at .05 level based on a t-test.
*** Represents significance at .10 level based on a t-test.
[a]Represents significance at .01 level based on a Dickey-Fuller test for unit roots.
[b]Represents significance at .05 level based on a Dickey-Fuller test for unit roots.
[c]Represents significance at .10 level based on a Dickey-Fuller test for unit roots.

Table 3.22. Tests of Restricted Models: Pooled AR(1) Log Models.

	y(t)				Δy(t)			
	R1	R2	R3	R4	R1	R2	R3	R4
	WTP							
Striped Bass:								
Bids ≥ 0	-	**	*,a	*,a	-	*	-	-
Bids > 0	-	-	*,a	*,a	-	***	-	-
Bluefish:								
Bids ≥ 0	-	*	*,a	*,a	-	*	-	-,c
Bids > 0	-	*	*,a	*,a	-	**	-	-
Fluke:								
Bids ≥ 0	-	*	*,a	*,a	-	*	-	-
Bids > 0	-	*	*,a	*,a	-	*	-	-
	WTA							
Striped Bass:								
Bids ≥ 0	-	*	*,a	*,a	-	*	-	-
Bids > 0	-	*	*,a	*,a	-	*	-	-
Bluefish:								
Bids ≥ 0	**	*	*,a	*,a	**	*	***,-	-
Bids > 0	-	*	*,a	*,a	-	**	-	-
Fluke:								
Bids ≥ 0	-	*	*,a	*,a	-	*	-	-
Bids > 0	-	*	*,a	*,a	-	*	-	-

Note: The unrestricted AR(1) model: $y(t) = \mu + \beta y(t-1) + DT + u(t)$, where μ is a drift term. DT refers to trend terms. here dummy variables associated with each month (the drift term is dropped when all dummy variables are included). R1 refers to the restricted AR(1) model: $y(t) = \mu + \beta y(t-1) + u(t)$. R2 refers to: $y(t) = \beta y(t-1) + u(t)$. R3 refers to: $y(t) = \mu + u(t)$, and R4 refers to: $y(t) = DJ + DA + DS + DO + u(t)$, a model with drift and trend terms. Null hypothesis is that the restricted model is true versus the alternative hypothesis that the unrestricted model is true.

* Represents significance at .01 level based on a t-test.
** Represents significance at .05 level based on a t-test.
*** Represents significance at .10 level based on a t-test.
[a]Represents significance at .01 level based on a Dickey-Fuller test for unit roots.
[b]Represents significance at .05 level based on a Dickey-Fuller test for unit roots.
[c]Represents significance at .10 level based on a Dickey-Fuller test for unit roots.

periods and pooled time periods are in agreement that an AR model does not follow a constant or drift term alone: 2) strong support was found that neither a trend nor seasonal component was present in this CV data, and hence, a trend component was not important in AR models of the WTP and WTA series. However, whether or not an AR model should include a drift term was found to be mixed when AR models were based on successive time periods, while results based on pooled data indicated that a drift term is necessary for log models but mixed for linear models. Inclusion of a drift may make sense from another perspective, whether or not CV bids approach a non-zero value over time. If so a drift term should be included.

Chapter 4 - Results and Implications of Intertemporal Convergence

4.1. Introduction

In this section results from an examination of convergence over time of the WTP and WTA series is presented. The results are further examined for robustness across subsets of the data without S0 bids versus the full data set, as well as for subsets of suspect observations.

4.2. Convergence of WTP and WTA Series and Implications of $0 Bids

Various AR models were estimated to examine convergence of WTP and of WTA series over time. AR(1) models were first estimated. Convergence criteria associated with an AR(1) models is based on the estimated parameter of an AR(1) model with or without a drift term. If the estimated parameter is less than one the series is said to converge over time and if greater than one the series is said to diverge. A testable hypothesis was developed to examine the issue of convergence in AR(1) models, i.e., the null hypothesis was formed to test that the estimated parameter is equal to one versus greater than one. These tests were based on a t-test although for typical time-series data a t-test is not appropriate because the test statistics are nonstandard (or that the error term is nonstationary) and a Dickey-Fuller test is commonly used (Enders 1995). In this investigation it is believed that the panel-type time-series data do not contain inherent structure common in typical time-series data namely seasonality and trend components, and that a t-test is appropriate given this characteristic (as a check Dickey-Fuller tests were conducted).

Results are contained in Table 4.1. For the WTP series parameter estimates were less than one in 9 of 12 cases, greater than 1 in 2 of 12 cases and equal to 1 in 1 case. Results indicated that the estimated parameter was significantly different and smaller than one in 7 out of 12 possible cases (all species across all consecutive two-period time periods for AR(1) model without drift), and significantly different and greater than one in 2 of 12

Table 4.1. Convergence Results of AR(1) Models for WTP Series and WTA Series: Full Data Set.

	Aug. = July	Sept. = Aug.	Oct. = Sept.	Nov. = Oct.
	WTP			
Striped Bass:				
AR(1) w/o drift	(<1)***,d	(<1)*,a	(>1)*,a	(<1)*,a
AR(1) w drift	(<1)*,a	(<1)*,a	(>1)*,a	(<1)*,a
Bluefish:				
AR(1) w/o drift	(<1)*,a	(<1)*,a	(<1)*,a	(=1.007)
AR(1) w drift	(<1)*,a	(<1)*,a	(<1)*,a	(=.999)
Fluke:				
AR(1) w/o drift	(<1)**,c	(>1)*,a	(<1)	(<1)
AR(1) w drift	(<1)*,a	(>1)*,a	(<1)	(<1)
	WTA			
Striped Bass:				
AR(1) w/o drift	(>1)	(<1)*,a	(<1)***,d	(<1)*,a
AR(1) w drift	(>1)	(<1)*,a	(<1)**,d	(<1)*,a
Bluefish:				
AR(1) w/o drift	(=.925)	(>1)*,d	(=1.0431)	(<1)*,a
AR(1) w drift	(<1)	(>1)	(=1.0139)	(<1)*,a
Fluke:				
AR(1) w/o drift	(>1)	(>1)*,a	(<1)	(<1)*,a
AR(1) w drift	(<1)	(>1)**,c	(<1)	(<1)*,a

Note: Numbers in parentheses represent the absolute value of the magnitude of the parameter estimate, ß as in the following AR(1) model: $y(t) = \beta y(t-1) + u(t)$, relative to 1. Test is if $\beta=1$ versus not.

* Represents significance at .01 level based on a t-test.
** Represents significance at .05 level based on a t-test.
*** Represents significance at .10 level based on a t-test.
[a]Represents significance at .01 level based on a Dickey-Fuller test for unit roots.
[b]Represents significance at .025 level based on a Dickey-Fuller test for unit roots.
[c]Represents significance at .05 level based on a Dickey-Fuller test for unit roots.
[d]Represents significance at .10 level based on a Dickey-Fuller test for unit roots.

cases. This result was also obtained when a drift term was included. Furthermore the finding was robust across the subset without S0 bids: the majority of cases involving WTP models exhibited estimates significantly different and smaller than one (Table 4.2). For these series the process converges based on these results. However, for WTP associated with October, September time periods for striped bass and September, August time periods for fluke the process diverges.

For the WTA series, parameter estimates were less than one for 6 of 12 cases, greater than one in 4 of 12 cases and equal to 1 in 2 of 12 cases (Table 4.1). Estimated parameters were found to be significantly different and smaller than one in 5 of 12 cases, and significantly different and greater than one in 1 case. Results were similar in the case of nonzero bids (Table 4.2). Overall evidence appears to favor convergence of WTP series (more cases exhibited convergent behavior) as opposed to WTA series which exhibited less convergent behavior on the basis of AR(1) models.

The tests of hypotheses (ß=1) based on Dickey-Fuller (D-F) tests tended to confirm the t-test results; all the same results and significance levels for the WTP series, and 2 cases out of 48 possible cases in the WTA series (full data set and subset without S0 bids combined) declined in significance by one-level (e.g., from .05 to .10 level). Both changes occurred for an AR(1) model with a drift term where critical values for a D-F test increased over those of an AR(1) without drift term. Also in one case, AR(1) model without drift the slope parameter for the WTA series (i.e., bluefish, September=August, full data set) becomes significant.

The similarity between both tests (t-test and D-F test) is probably surprising given the time-series literature. Traditional time-series analysis is based on historical prices, financial instruments, exchange rates, etc. These series contin inherent structure such as seasonality and/or trend effects that could result from macroeconomic events such as cyclical changes. It is suspected that these factors contribute to specific structure of the error term that renders OLS and traditional statistical tests inappropriate due to nonstandard distributions in the

Table 4.2. Convergence Results of AR(1) Models for WTP Series and WTA Series: Without $0 Bids.

	Aug. = July	Sept. = Aug.	Oct. = Sept.	Nov. = Oct.
	WTP			
Striped Bass:				
AR(1) w/o drift	(<1)	(<1)[*,a]	(>1)[*,a]	(<1)[*,a]
AR(1) w drift	(<1)[**,c]	(<1)[*,a]	(>1)[*,a]	(<1)[*,a]
Bluefish:				
AR(1) w/o drift	(<1)	(<1)[*,a]	(<1)[*,a]	(>1)
AR(1) w drift	(<1)[***,c]	(<1)[*,a]	(<1)[*,a]	(>1)
Fluke:				
AR(1) w/o drift	(<1)[**,b]	(>1)[*,a]	(<1)	(=1.0663)
AR(1) w drift	(<1)[*,a]	(>1)[*,a]	(<1)	(=.9459)
	WTA			
Striped Bass:				
AR(1) w/o drift	(>1)	(<1)[**,c]	(<1)	(<1)[*,a]
AR(1) w drift	(<1)	(<1)[*,a]	(<1)[***,d]	(<1)[*,a]
Bluefish:				
AR(1) w/o drift	(=1.0774)	(>1)	(=1.0694)	(<1)[*,a]
AR(1) w drift	(<1)	(>1)	(=.9976)	(<1)[**,d]
Fluke:				
AR(1) w/o drift	(>1)	(>1)[*,a]	(<1)	(<1)[*,a]
AR(1) w drift	(<1)	(>1)[**,c]	(<1)	(<1)[*,a]

Note: Numbers in parentheses represent the absolute value of the magnitude of the parameter estimate, ß as in the following AR(1) model: $y(t) = \beta y(t-1) + u(t)$, relative to 1.

* Represents significance at .01 level based on a t-test.
** Represents significance at .05 level based on a t-test.
*** Represents significance at .10 level based on a t-test.
[a]Represents significance at .01 level based on a Dickey-Fuller test for unit roots.
[b]Represents significance at .025 level based on a Dickey-Fuller test for unit roots.
[c]Represents significance at .05 level based on a Dickey-Fuller test for unit roots.
[d]Represents significance at .10 level based on a Dickey-Fuller test for unit roots.

standard error term. The series in which this investigation is based on does not contain this inherent structure because the data are panel data, represents welfare measures - not prices, and do not represent a long time series as does traditional time series data. It is believed that this welfare series are free of seasonality, trend and other such structure inherent in traditional time series data as shown in Chapter 3. This showed that the restricted AR(1) model without trend and/or drift could not be rejected compared to an unrestricted AR(1) model with drift and trend terms in all cases (across species, functional form, and subsets of S0 bids). This is strong evidence that there is no trend nor seasonality associated the CV data used in this investigation.

Examination of convergent behavior followed based on AR(2) models. In this case the roots associated with the characteristic equation of an estimated AR(2) model were derived based on the logic of the quadratic equation using the mathematical computational package, Maple. Where complex roots were derived the procedure for finding the length of the vector from the origin to the point represented by the root described in Chapter 2 was conducted. Results of the relative magnitude of these roots are contained in Table 4.3. Evidence indicates that both WTP series and WTA series exhibit relatively similar convergent behavior, 5 of 9 cases for WTP had estimated roots less than one, and 6 of 9 cases for WTA had estimated roots less than one. Results were robust over the subset without S0 bids.

Next an AR(3) model was estimated and the respective roots were derived for the associated characteristic equation. Results for the WTP series were mixed, 3 of 6 cases had estimated roots less than one, and 4 of 6 cases for the WTA series exhibited roots greater than one (Table 4.4). Results of AR(4) models were less encouraging, 3 of 3 possible cases exhibited divergence behavior for the WTP series and only 1 of 3 cases for the WTA series (Table 4.5). It is possible that some aspect of data variability is causing these results. The next step involved examining convergence behavior using AR models for various subsets of suspect observations.

Table 4.3. Convergence Results of AR(2) Models for WTP Series and WTA Series: Implications of S0 Bids.

	Sept.= Aug. + July	Oct. = Sept. + Aug.	Nov. = Oct. + Sept.
Striped Bass:			
WTP:			
w SO	R<1	R>1	R<1
w/o SO	R<1	R>1	R<1
WTA:			
w SO	R>1	R<1	R<1
w/o SO	R>1	R<1	R>1
Bluefish:			
WTP:			
w SO	R<1	R<1	R>1
w/o SO	R>1	R<1	R>1
WTA:			
w SO	R>1	R<1	R<1
w/o SO	R>1	R<1	R<1
Fluke:			
WTP:			
w SO	R>1	R>1	R<1
w/o SO	R>1	R>1	R>1
WTA:			
w SO	R>1	R<1	R<1
w/o SO	R>1	R<1	R<1

Note: R refers to the roots of the characteristic equation associated with an appropriate AR model. The relative magnitude of all roots are indicated relative to the length of the radius of the unit circle, e.g., R<1 indicates that all roots lie inside the unit circle.

Table 4.4. Convergence Results of AR(3) Models for WTP Series and WTA Series: Implications of S0 Bids.

	Oct. = Sept. + Aug. + July	Nov. = Oct. + Sept. + Aug.
Striped Bass:		
WTP:		
w S0	R<1	R<1
w/o S0	R>1	R<1
WTA:		
w S0	R>1	R>1
w/o S0	R>1	R>1
Bluefish:		
WTP:		
w S0	R<1	R>1
w/o S0	R<1	R>1
WTA:		
w S0	R>1	R<1
w/o S0	R>1	R>1
Fluke:		
WTP:		
w S0	R>1	R>1
w/o S0	R>1	R>1
WTA:		
w S0	R>1	R<1
w/o S0	R>1	R<1

Note: R refers to the roots of the characteristic equation associated with an appropriate AR model. The relative magnitude of all roots are indicated relative to the length of the radius of the unit circle, e.g., R<1 indicates that all roots lie inside the unit circle.

Table 4.5. Convergence Results of AR(4) Models for WTP Series and WTA Series: Implications of S0 Bids.

	Nov. = Oct. + Sept. + Aug. + July
Striped Bass:	
WTP:	
w S0	R>1
w/o S0	R>1
WTA:	
w S0	R<1
w/o S0	R<1
Bluefish:	
WTP:	
w S0	R>1
w/o S0	R>1
WTA:	
w S0	R>1
w/o S0	R>1
Fluke:	
WTP:	
w S0	R>1
w/o S0	R>1
WTA:	
w S0	R>1
w/o S0	R>1

Note: R refers to the roots of the characteristic equation associated with an appropriate AR model. The relative magnitude of all roots are indicated relative to the length of the radius of the unit circle, e.g., R<1 indicates that all roots lie inside the unit circle.

4.3. Convergence Behavior of WTP Bids and Implications of Suspect Observations

In this section the issue of convergence is examined for various subsets of suspect observations identified earlier. A testable hypothesis was examined, whether the estimated parameter equal one versus not for AR(1) WTP bid models for striped bass. Results indicate that based on the full data 3 of 4 consecutive WTP series exhibits convergent behavior: 3 of 4 AR models resulted in estimated parameters less than one and these were significantly different than one (Table 4.6). This result is identical for the data subset that corresponds to "Noise," however, the data subset of large WTP bids results in convergent behavior of AR models for all consecutive time periods. Surprisingly the influence data subset exhibited mixed results.

Hypothesis tests based on unit root tests (i.e., ß=1) were also conducted for the AR(1) models. Only 2 changes occurred when hypothesis tests were based on a D-F test compared to t-test results (i.e., September=August period for the subset without "influence" and for the November=October period for the subset without "large & noise" Table 4.6). Hence, results of the unit root tests based on a D-F test were almost exact compared to those based on t-tests. Again it is suspected that this similarity is because the CV data does not contain inherent structure as does traditional time-series data.

Findings from the AR(2) models and their characteristic equations indicates that convergence behavior is robust over all data subsets except for the "influence" subset, although the magnitudes and/or length of some roots change and reverse over various subsets (Table 4.6). All indicate that in 2 of 3 different time periods WTP behavior represented convergence behavior. Results of the AR(3) models and associated characteristic equation were more encouraging where results indicated that for both time periods the WTP series converges (Table 4.7). This result was robust over the data subsets except for the "Noise" subset. Finally the AR(4) model and roots from the characteristic equation indicate convergence behavior for the full data set, and the data set without "large" WTP bids (Table 4.7). Overall the results based on the Belsley *et al.*

Table 4.6. Convergence Results of AR(1) and AR(2) Models for WTP Series for Striped Bass: Implications of Subsets of Outliers.

	AR(1) Models			
	Aug. = July	Sept. = Aug.	Oct. = Sept.	Nov. = Oct.
Full data set	(<1)***,d	(<1)*,a	(>1)*,a	(<1)*,a
w/o Noise	(<1)	(<1)*,a	(>1)*,a	(<1)*,a
w/o Large	(<1)**,c	(<1)*,a	(<1)*,a	(<1)*,a
w/o Large + Noise	(>1)***,d	(<1)	(<1)	(<1)**,b
w/o Influence	(>1)	(<1)*,d	(=.9748)	(>1)**,c

	AR(2) Models		
	Sept.= Aug. + July	Oct. = Sept. + Aug.	Nov. = Oct. + Sept.
Full data set	R<1	R>1	R<1
w/o Noise	R<1	R>1	R<1
w/o Large	R>1	R<1	R<1
w/o Large + Noise	R>1	R<1	R<1
w/o Influence	R>1	R<1	R>1

Note: For AR(1) models, numbers in parentheses represent the absolute value of the magnitude of the parameter estimate, ß as in the following AR(1) model: $y(t) = ßy(t-1) + u(t)$, relative to 1. Test is if ß=1 versus not. In addition for AR(2) models, R refers to the roots of the characteristic equation associated with an appropriate AR model. The relative magnitude of all roots are indicated relative to the length of the radius of the unit circle, e.g., R<1 indicates that all roots lie inside the unit circle.

* Represents significance at .01 level based on a t-test.

** Represents significance at .05 level based on a t-test.

*** Represents significance at .10 level based on a t-test.

[a]Represents significance at .01 level based on a Dickey-Fuller test for unit roots.

[b]Represents significance at .025 level based on a Dickey-Fuller test for unit roots.

[c]Represents significance at .05 level based on a Dickey-Fuller test for unit roots.

[d]Represents significance at .10 level based on a Dickey-Fuller test for unit roots.

Table 4.7. Convergence Results of AR(3) and AR(4) Models for WTP Series and WTA Series: Implications of Subsets of Outliers.

	---------- AR(3) Models ----------	
	Oct. = Sept. + Aug. + July	Nov. = Oct. + Sept. + Aug.
Full data set	R<1	R<1
w/o Noise	R>1	R>1
w/o Large	R<1	R<1
w/o Large + Noise	R<1	R>1
w/o Influence	R<1	R<1

	---------- AR(4) Models ----------
	Nov. = Oct. + Sept. + Aug. + July
Full data set	R<1
w/o Noise	R>1
w/o Large	R<1
w/o Large + Noise	R>1
w/o Influence	R>1

Note: R refers to the roots of the characteristic equation associated with an appropriate AR model. The relative magnitude of all roots are indicated relative to the length of the radius of the unit circle, e.g., R<1 indicates that all roots lie inside the unit circle.

(1980) influence procedure had mixed results, with the majority of cases showing divergence behavior.

4.4. Conclusions and Directions

Previous applications of the CV method have not considered the use of time series techniques to examine and test for evidence of convergence in CV bid behavior. Such an approach applied to panel data can assist in research in this area and is to be encouraged.

To summarize the findings of this investigation it is noted that on the basis of mean values as illustrated in the series' graphs provides evidence, that especially for the specie bluefish some level of convergence in WTP bids occurred. WTP bids declined from initial levels over time and this result was somewhat robust without the presence of $0 bids.

However, when time-series techniques are used to examine convergence of a series findings are somewhat mixed. For example, convergence of WTP bid series for striped bass had mixed results, and this result was found to be robust over various subsets of suspect observations. For all time periods and AR series representations based on the full data set results indicated convergence behavior, but for the AR series that involved October and September time periods results were opposite. There was no single data subset whereby its exclusion caused results that outperformed one another, i.e., the results were robust across subsets of suspect observations.

The results and findings based on convergence of WTP and WTA series and possible effects of $0 bids appear to indicate the WTP series exhibited more convergence behavior than did the WTA series, and that results were robust over subsets without $0 bids. A reexamination of the graphs of the series would tend to confirm this finding, that the WTP series appears to exhibit more convergence behavior than does the WTA series. It is suspected that the data associated with fluke suffers from too few observations since the number of observations that contain CV bids over the various time periods is small.

Perhaps this is a problem for all data sets. however. controlled experiments usually involve from 7 to 15 or more individuals per trial. The WTP series associated with bluefish trips appears to exhibit convergence behavior from the series graphs and is thought to be a good candidate for the empirical tests, however performance of AR processes beyond an AR(1) process degenerate.

It could be possible that convergence in a time series sense and based on time series techniques represents a strong test of convergence and could be overly sensitive to the variation present in these CV bids. Further research is the only means to address and resolve this concern.

Other possible explanations for the mixed findings of this investigation concerns the degree of substitutability of the goods in this experiment, and the fact that the experiment involved an applied field setting. To an active sport fisherman the decision and subsequent fishing tactics for each of the three species involve differences in effort, location, and in equipment. In this sense the three different species trips involves a fairly unique good with few substitutes. As Hanemann (1991) has argued and as Shogren *et al.* (1994) has demonstrated the degree of substitutability between goods matters and can become a critical factor concerning convergence. Unfortunately with sport fishing there may be few species and/or targets which can serve as substitutes for each other. Certain species are not abundant at the same time, and locations and fishing equipment are fairly specific to particular species, and it is quite possible that sport fishermen prefer specific species and tailor their fishing activity towards that specie. For these reasons it may not be possible to uniquely determine convergence or divergence behavior in sport fishing applications.

Overall the evidence found in this investigation based on quasi-experimental data (repeated games of sport fishermen over a season involving up to 5 repeated surveys/games each corresponding to a month, yielding 5 bidding periods) suggests that experienced, active recreational participants do revise their CV bids over multiple time periods. This could be due to their familiarity with the CV game over time, changes in their tastes and in repeated thinking about their true preferences, and in their success or failure in a particular time period.

Such findings suggest profound effects on future work in CV methods and in experimental methods that elicit preferences for nonmarket/public type goods. First off, these findings suggest that single point estimates based on one-time application of CV method may not be appropriate in field studies designed to assess values for NRDAs, no matter how good the application is. Because of tastes and experience, respondents can and will revise their CV bids in repeated game playing. Secondly, our findings suggest the need for continued experimental evidence in the CV method; to advance the technique, and to establish credibility and recommendations for its use.

References

Adamowicz, W.L., V. Bhardwaj and B. Macnab. 1993. "Experiments on the Difference Between Willingness to Pay and Willingness to Accept." *Land Economics*. 69(4): 416-427.

Banford, N.D., J.L. Knetsch and G.A. Mauser. 1977. *Compensating and Equivalent Variation Measures of Consumer's Surplus: Further Survey Results*. Department of Economics and Commerce, Simon Fraser University: Burnaby, British Columbia.

Belsley, D.A., E. Kuh and R.E. Welsch. 1980. *Regression Diagnostics*. John Wiley & Sons: New York, NY.

Bishop, R.C. and T.A. Heberlein. 1979. "Measuring Values of Extra–Market Goods: Are Indirect Measures Biased?" *American Journal of Agricultural Economics* 61: 926–30.

Bishop, R.C. and T.A. Heberlein. 1986. "Does Contingent Valuation Work?" In Cummings, R.G., B.S. Brookshire and W.D. Schulze. (eds.) 1986. *Valuing Public Goods: The Contingent Valuation Method.* Rowman and Allanheld: Totowa, NJ: 123-147.

Bishop, R.C., T.A. Heberlein and M.J. Kealy. 1983. "Contingent Valuation of Environmental Assets: Comparisons with a Simulated Market." *Natural Resources Journal* 23(July): 619-633.

Bishop, R.C., T.A. Heberlein, D. McCollum and M.P. Welsh. 1988. *A Validation Experiment for Valuation Techniques*. School of Natural Resources, University of Wisconsin: Madison, WI.

Bjornstad, D.J. and J.R. Kahn. (eds.) 1996. *The Contingent Valuation of Environmental Resources: Methodological Issues and Research Needs*. Edward Elger: Brookfield, VT.

Bohm, P. 1972. "Estimating the Demand for Public Goods: An Experiment." *European Economic Review* 3: 111-130.

Boyle, K.J., W.H. Desvousges, F.R. Johnson, R.W. Dunford and S.P. Hudson. 1994. "An Investigation of Part-Whole Biases in Contingent Valuation Studies." *Journal of Environmental Economics and Management* 27: 64-83.

Boyle, K.J., F.R. Johnson, D.W. McCollum, W.H. Desvousges, R.W. Dunford and S.P. Hudson. 1996. "Valuing Public Goods: Discrete versus Continuous Contingent-Valuation Responses." *Land Economics* 72(3): 381-396.

Brookshire, D.S., R.C. d'Arge, W.D. Schulze and M.A. Thayer. 1979. *Methods Development for Assessing Tradeoffs in Environmental Management, Vol. II, Experiments in Valuing Non–Market Goods: A Case Study of Alternative Benefit Measures of Air Pollution Control in the South Coast Air Basin of Southern California*. EPA–600/6–79–001b. Prepared for Office of Research Development, U.S. EPA, Washington, DC.

Brookshire, D.S., R.C. d'Arge, W.D. Schulze and M.A. Thayer. 1981. "Experiments in Valuing Public Goods." In Smith, V.K. (ed.) 1981. *Advances in Applied Microeconomics*. JAI Press: Greenwich, CT.

Brookshire, D.S., B.C. Ives and W.D. Schulze. 1976. "The Valuation of Aesthetic Preferences." *Journal of Environmental Economics and Management* 3(4): 325-346.

Brookshire, D.S., A. Randall and J.R Stoll. 1980. "Valuing Increments and Decrements in Natural Resource Service Flows." *American Journal of Agricultural Economics* 62(August): 478-488.

Brookshire, D.S., M.A. Thayer, W.D. Schulze and R.C. d'Arge. 1982. "Valuing Public Goods: A Comparison of Survey and Hedonic Approaches." *American Economic Review* 72(1): 165–177.

Cameron, T.A. 1988. "A New Paradigm for Valuing Non-Market Goods Using Referendum Data." *Journal of Environmental Economics and Management* 15: 355-379.

Cameron, T.A. 1992. "Combining Contingent Valuation and Travel Cost Data for the Valuation of Nonmarket Goods." *Land Economics* 68(3): 302-317.

Carson, R.T., W.M. Hanemann, R.J. Kopp, J.A. Krosnick, R.C. Mitchell, S. Presser, P.A. Rudd, and V.K. Smith. 1995. *Temporal Reliability of Estimates from Contingent Valuation*. Discussion Paper 95-37. Resources for the Future: Washington, DC.

Carson, R.T., N.E. Flores, K.M. Martin and J.L. Wright. 1996a. "Contingent Valuation and Revealed Preference Methodologies: Comparing the Estimates for Quasi-Public Goods." *Land Economics* 72(1): 80-99.

Carson, R.T., W.M. Hanemann, R.J. Kopp, J.A. Krosnick, R.C. Mitchell, S. Presser, P.A. Rudd, V.K. Smith, M. Conaway and K. Martin. 1996b. *Referendum Design and Contingent Valuation: The NOAA Panel's No-Vote Recommendation*. Discussion Paper 96-05. Resources for the Future: Washington, DC.

Carson, R.T., W.M. Hanemann, R.J. Kopp, J.A. Krosnick, R.C. Mitchell, S. Presser, P.A. Rudd, V.K. Smith, M. Conaway and K. Martin. 1996c. *Was the NOAA Panel Correct About Contingent Valuation?* Discussion Paper 96-20. Resources for the Future: Washington, DC.

Chow, G.C. 1969. "Tests of the Equality Between Subsets of Coefficients in Two Linear Regressions." *Econometrica* 28: 591-605.

Coursey, D.L., J. Hovis and W.D. Schulze. 1987. "The Disparity Between Willingness to Accept and Willingness to Pay Measures of Value." *Quarterly Journal of Economics* 102: 679-690.

Coursey, D.L. and W.D. Schulze. 1986. "The Application of Laboratory Experimental Economics to the Contingent Valuation of Public Goods." *Public Choice* 49(1): 47-68.

Cummings, R.G., L.A. Cox and A.M. Freeman. 1986a. "General Methods for Benefits Assessment." In Bentkover, J.D., V.T. Covello and J. Mumpower. (eds.) 1986. *Benefits Assessment: The State of the Art*. Boston, MA: D. Reidel Publishing Co.

Cummings, R.G., D.S. Brookshire and W.D. Schulze. 1986b. *Valuing Public Goods: The Contingent Valuation Method*. Rowman and Allanheld: Totowa.

Davis, D.D. and C.A. Holt. 1993. *Experimental Economics*. Princeton University Press: Princeton, NJ.

Davis, R.K. 1963. "The Value of Outdoor Recreation: An Economic Study of the Maine Woods." Ph.D. Dissertation. Harvard University: Cambridge, MA.

Davis, R.K. 1964. "The Value of Big Game Hunting in a Private Forest." In *Transactions of the 29th North American Wildlife and Natural Resources Conference*. Wildlife Management Institute: Washington, DC.

Desvousges, W.H., S.P. Hudson and M.C. Ruby. 1996. "Evaluating CV Performance: Separating the Light from the Heat." In Bjornstad, D.J. and J.R. Kahn. 1996. *The Contingent Valuation of Environmental Resources: Methodological Issues and Research Needs*. Edward Elger: Brookfield, VT: 117-144.

Desvousges, W.H., F.R. Johnson, R.W. Dunford, K.J. Boyle, S.P. Hudson and K. W. Wilson. 1993. "Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity." In Hausman, J.A. (ed.) 1993. *Contingent Valuation: A Critical Assessment*. North Holland: New York, NY: 91-164.

Desvousges, W.H., V.K. Smith and A. Fisher. 1987. "Option Price Estimates for Water Quality Improvements: A Contingent Valuation Study for the Monongahela River." *Journal of Environmental Economics and Management* 14: 248-267.

Desvousges, W.H., V.K. Smith and M.P. McGivney. 1983. *A Comparison of Alternative Approaches for Estimating Recreation and Related Benefits of Water Quality Improvements*. Prepared for Office of Policy Analysis, U.S. Environment Protection Agency, Washington, DC.

Diamond, P.A. and J.A. Hausman. 1993. "On Contingent Valuation of Nonuse Values." In Hausman, J.A. (ed.) 1993. *Contingent Valuation: A Critical Assessment*. North Holland: New York, NY: 3-38.

Diamond, P.A. and J.A. Hausman. 1994. "Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives* 8(4): 45-64.

Diamond, P.A. and J.A. Hausman, G.K. Leonard and M.A. Denning. 1993. "Does Contingent Valuation Measure Preferences? Experimental Evidence." In Hausman, J.A. (ed.) 1993. *Contingent Valuation: A Critical Assessment*. North Holland: New York, NY: 41-90.

Dillman, D.A. 1978. *Mail and Telephone Surveys: The Total Design Method*. John Wiley & Sons: New York, NY.

Enders, W. 1995. *Applied Econometric Time Series*. John Wiley & Sons: New York, NY.

Fisher, F.M. 1970. "Tests of the Equality Between Subsets of Coefficients in Two Linear Regressions: An Expository Note." *Econometrica* 38: 361-366.

Fomby, T.B., R.C. Hill and S.R. Johnson. 1988. *Advanced Econometric Methods*. (Corrected edition). Springer-Verlag: New York, NY.

Freeman. A.M. 1979. *The Benefits of Environmental Improvement: Theory and Practice*. Resources for the Future. Inc.: Washington. D.C.

Freeman. A.M. 1993. *The Measurement of Environmental and Resource Values: Theory and Method*. Resources for the Future. Inc.: Washington. D.C.

Gordon. I.M. and J.L. Knetsch. 1979. "Consumer's Surplus Measures and the Evaluation of Resources." *Land Economics* 55: 1–10.

Greene. W.H. 1993. *Econometric Analysis*. MacMillan: New York. NY.

Gregory. R. and L. Furby. 1987. "Auctions. Experiments and Contingent Valuation." *Public Choice* 55: 273-289.

Hamilton. J.D. 1994. *Time Series Analysis*. Princeton University Press: Princeton, NJ.

Hammack. J. and G.M. Brown. Jr. 1974. *Waterfowl and Wetlands: Toward Bioeconomic Analysis*. Johns Hopkins University Press: Baltimore, MD.

Hanemann. W.M. 1983. "Marginal Welfare Measures for Discrete Choice Models." *Economic Letters* 13: 129-136.

Hanemann W.M. 1984a. "Welfare Evaluation in Contingent Valuation Experiments with Discrete Responses." *American Journal of Agricultural Economics* 66(3): 332–341.

Hanemann, W.M. 1984b. "Discrete/Continuous Models of Consumer Demand." *Econometrica* 52(3): 541-561.

Hanemann. W.M. 1985. "Some Issues in Continuous- and Discrete-Response Contingent Valuation Studies." *Northeastern Journal of Agricultural and Resource Economics* 14(1): 5-13.

Hanemann. W.M. 1991. "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *American Economic Review* 81(3): 635-647.

Hausman, J.A. 1981. "Exact Consumer's Surplus and Deadweight Loss." *American Economic Review* 71: 662-676.

Hausman, J.A. (ed.) 1993. *Contingent Valuation: A Critical Assessment*. North Holland: New York, NY.

Johansson, P.O. 1987. *The Economic Theory and Measurement of Environmental Benefits*. Cambridge University Press: New York, NY.

Johansson, P.O. 1993. *Cost-Benefit Analysis of Environmental Change*. Cambridge University Press: New York, NY.

Jones-Lee, M.W., M. Hammerton and P.R. Philips. 1985. "The Value of Safety: Results of a National Sample Survey." *The Economic Journal* 95(March): 49-72.

Judge. G.G.. R.C. Hill. W.E. Griffiths. H. Lutkepohl and T.C. Lee. 1988. *Introduction to the Theory and Practice of Econometrics*. (Second edition.) John Wiley and Sons: New York, NY.

Just. R.E.. D.L. Hueth and, A. Schmitz. 1982. *Applied Welfare Economics and Public Policy*. Prentice–Hall, Inc.: Englewood Cliffs, NJ.

Kealy. M.J., J.F. Dovidio and M.L. Rockel. 1988. "Accuracy in Valuation is a Matter of Degree." *Land Economicsl* 64(2): 158-171.

Kealy, M.J.. M. Montgomery and J.F. Dovidio. 1990. " Reliability and Predictive Validity of Contingent Values: Does the Nature of the Good Matter?" *Journal of Environmental Economics and Management* 19: 244-263.

Knetsch, J.L. 1989. "The Endowment Effect and Evidence of Nonreversible Indifference Curves." *American Economic Review* 79(5): 1277-1284.

Knetsch, J.L. and R.K. Davis. 1966. "Comparisons of Methods for Recreation Evaluation." In Kneese, A.V. and S.C. Smith. (eds.) 1966. *Water Research*. Johns Hopkins University Press: Baltimore,MD.

Knetsch, J.L. and J.A. Sinden. 1984. "Willingness to Pay and Compensated Demand: Experimental Evidence of an Unexpected Disparity in Measures of Value." *Quarterly Journal of Economics* 99: 507-521.

Layard, P.R.G. and A.A. Walters. 1978. *Microeconomic Theory*. McGraw-Hill: New York. NY.

Loehman, E.T. 1984. *Willingness to Pay for Air Quality: A Comparison of Two Methods*. Staff Paper 84-18. Dept. of Agric. Econ., Purdue University: West Lafayette, IN.

Loehman, E.T., D.Boldt and K.C. Chaikin. 1981. *Measuring the Benefits of Air Quality Improvements in the San Francisco Bay Area*. SRI Report No. 8962. Stanford Research Institute International: Menlo Park, CA.

Loehman, E.T. and V.H. De. 1982. "Application of Stochastic Choice Modeling to Policy Analysis of Public Goods: A Case Study of Air Quality Improvements." *Review of Economics and Statistics* 64(3): 474-480.

Loehman, E.T., S. Park and D. Bolt. 1994. "Willingness to Pay for Gains and Losses in Visibility and Health." *Land Economics* 70(4): 478-498.

Loomis, J.B. 1989. "Test-Retest Reliability of the Contingent Valuation Method: A Comparison of General Population and Visitor Responses." *American Journal of Agricultural Economics* 71(February): 76-84.

Loomis, J.B. 1990. "Comparative Reliability of the Dichotomous Choice and Open-Ended Contingent Valuation Techniques." *Journal of Environmental Economics and Management* 18: 78-85.

McConnell, K.E. 1977. "Congestion and Willingness to Pay: A Study of Beach Use." *Land Economics* 53(2): 185-195.

Mitchell, R.C. and R.T. Carson. 1981. *An Experiment in Determining Willingness to Pay for National Water Quality Improvements.* Draft Report. Prepared for Office of Research and Development, U.S. Environment Protection Agency, Washington, DC.

Mitchell, R.C. and R.T. Carson. 1984. *A Contingent Valuation Estimate of National Freshwater Benefits*. Final Report. Prepared for Office of Research and Development. U.S. Environment Protection Agency. Washington, DC.

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

Mueller, D.C. 1979. *Public Choice*. Cambridge University Press: New York, NY.

Myles, G.D. 1995. *Public Economics.* Cambridge University Press: New York, NY.

Ofiara, D.D. and J.J. Seneca. 1998. *Valuing Economic Damages and Losses from Marine Pollution: A Handbook of Principles and Guidelines.* Island Press: Covelo, CA. Forthcoming.

Portney, P.R. 1994. "The Contingent Valuation Debate: Why Economists Should Care." *Journal of Economic Perspectives* 8(4): 3-18.

Randall, A. and J.P. Hoehn. 1996. "Embedding in Market Demand Systems." *Journal of Environmental Economics and Management* 30: 369-380.

Randall, A., J.P. Hoehn and D.S. Brookshire. 1983. "Contingent Valuation Surveys for Evaluating Environmental Assest." *Natural Resources Journal* (23): 635-648.

Randall, A.B.C. Ives and C. Eastman. 1974. "Bidding Games for Valuation of Aesthetic Environmental Improvements." *Journal of Environmental Economics and Management* 1: 132-149.

Randall, A., O. Grunewald, A. Pagoulatos, R. Ausness and S. Johnson. 1978. "Reclaiming Coal Surface Mines in Central Appalachia: A Case Study of the Benefits and Costs." *Land Economics* 54(4): 427-489.

Randall, A. and J.R. Stoll. 1980. "Consumer's Surplus in Commodity Space." *American Economic Review* 70(3): 449-455.

Ready, R.C., J.C. Buzby and D. Hu. 1996. "Differences between Contingent and Discrete Contingent Value Estimates." *Land Economics* 72(3): 397-411.

Reiling, S.D., K.J. Boyle, M.L. Phillips and M.W. Anderson. 1990. "Temporal Reliability of Contingent Values." *Land Economics* 66(2): 128-134.

Rowe, R.D., R.C. d'Arge and D.S. Brookshire. 1980. "An Experiment on the Economic Value of Visibility." *Journal of Environmental Economics and Management* 7(1): 1-19.

Rowe, R.D. and L.G. Chestnut. 1983. "Valuing Environmental Commodities: Revisited." *Land Economics* 59: 404–410.

Schulze. W.D. and D.S. Brookshire. 1983. "The Economic Benefits of Preserving Visibility in the National Parks of the Southwest." *Natural Resources Journal* 23: 149-173.

Schulze. W.D.. R.C. d'Arge and D.S. Brookshire. 1981. "Valuing Environmental Commodities: Some Recent Experiments. *Land Economics* 57:151-172.

Schulze, W.D.. G. McClelland, D. Waldman and J. Lazo. 1996. "Sources of Bias in Contingent Valuation." In Bjornstad, D.J. and J.R. Kahn. (eds.) 1996. *The Contingent Valuation of Environmental Resources: Methodological Issues and Research Needs.* Edward Elger: Brookfield, VT: 97-116.

Seller, C., J.R. Stoll and J.P. Chavas. 1985. "Validation of Empirical Measures of Welfare Change: A Comparison of Nonmarket Techniques." *Land Economics* 61(2): 156-175.

Shogren, J.F.. S.Y. Shin, D.J. Hayes and J.B. Kliebenstein. 1994. "Resolving Differences in Willingness to Pay and Willingness to Accept." *American Economic Review* 84(1): 255-270.

Shone. R. 1997. *Economic Dynamics.* Cambridge University Press: New York. NY.

Sinclair, W.F. 1976. *The Economic and Social Impact of the Kemano II Hydroelectric Project on British Columbia's Fisheries Resources.* Fisheries and Marine Service, Department of the Environment: Vancouver, BC.

Smith, V.K. and W.H. Desvousges. 1986. *Measuring Water Quality Benefits.* Kluwer. Nijhoff Publishing: Boston, MA.

Smith, V.K. and L.L. Osborne. 1996. "Do Contingent Valuation Estimates Pass a "Scope" Test? A Meta-Analysis." *Journal of Environmental Economics and Management* 31: 287-301.

Smith, V.L. 1979. "Incentive Compatible Experimental Processes for the Provision of Public Goods." In Smith. V.L. (ed.) 1979. *Research in Experimental Economics.* JAI Press: Greenwich, CT.

Smith, V.L. 1980. "Experiments with a Decentralized Mechanism for Public Goods Decisions." *American Economic Review* 70: 584-599.

Smith, V.L. 1982. "Microeconomic Systems as an Experimental Science." *American Economic Review* 72: 923-955.

Smith, V.L. 1986. "Comments." In Cummings, R.G., D.S. Brookshire, and W.D. Schulze. (eds.) 1986. *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method.* Rowman & Allanheld: Totowa, NJ: 197-204.

Stevens, T.H., T.A. More and R.J. Glass. 1994. "Interpretation and Temporal Stability of CV Bids for Wildlife Existence: A Panel Study." *Land Economics* 70(3): 355-363.

Thayer, M.A. 1981. "Contingent Valuation Techniques for Assessing Environmental Impacts: Further Evidence." *Journal of Environmental Economics and Management* 8: 27-44.

U.S. Department of Commerce, National Oceanic Atmospheric Administration. 1993. "Natural Resource Damage Assessments: Advance Notice of Proposed Rulemaking, Extension of Comment Period, and Release of Contingent Valuation Methodology Report." *Federal Register* 58(10), Friday, January 15: 4600-4614.

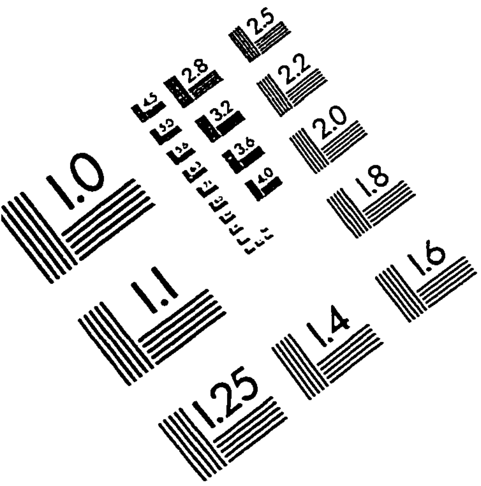
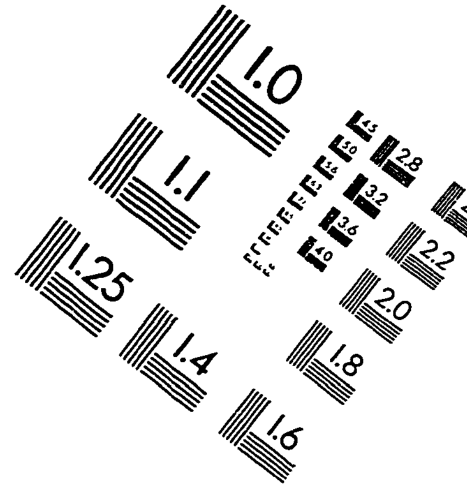
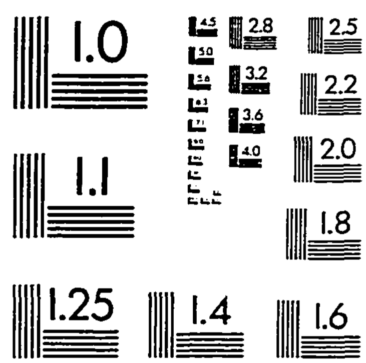
Whitehead, J.C., G.C. Blomquist, T.J. Hoban and W.B. Clifford. 1995. "Assessing the Validity and Reliability of Contingent Values: A Comparison of On-Site Users, Off-Site Users, and Non-users." *Journal of Environmental Economics and Management* 29: 238-251.

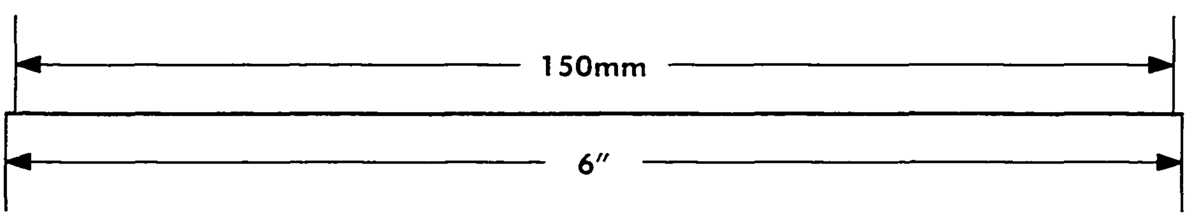
Willig, R.D. 1976. "Consumer's Surplus Without Apology." *American Economics Review* 66(4): 589–597.

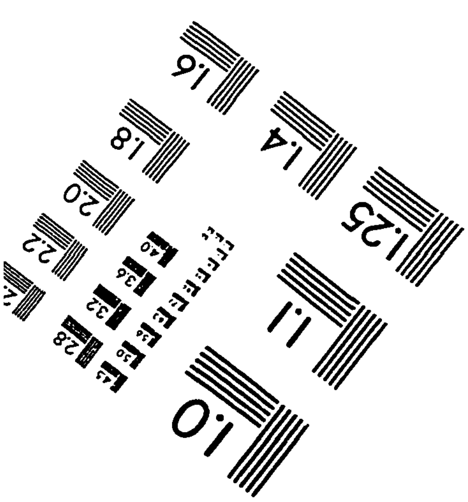
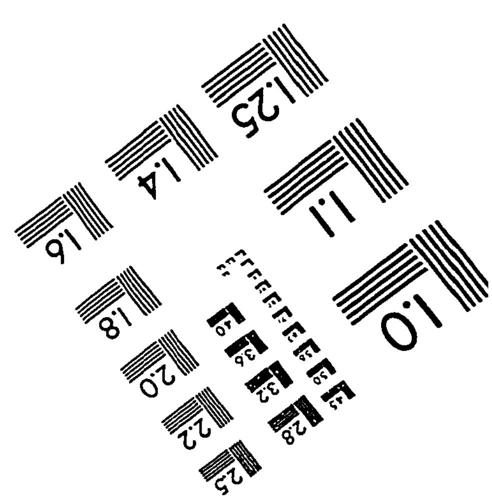
Willis, K.G. 1995. "Contingent Valuation in a Policy Context: The National Oceanic and Atmospheric Administration Report and Its Implications for the Use of Contingent Valuation Methods in Policy Analysis in Britain." In Willis, K.G. and J.T. Corkindale. (eds.) 1995. *Environmental Valuation: New Perspectives.* CAB International: Wallingford, U.K: 118-143.

Willis, K.G. and J.T. Corkindale. (eds.) 1995. *Environmental Valuation: New Perspectives.* CAB International: Wallingford, U.K.

IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989