# Economic value of instream flows in Montana: A travel cost model approach 

Christopher J. Neher
The University of Montana

Follow this and additional works at: https://scholarworks.umt.edu/etd Let us know how access to this document benefits you.

## Recommended Citation

Neher, Christopher J., "Economic value of instream flows in Montana: A travel cost model approach" (1989). Graduate Student Theses, Dissertations, \& Professional Papers. 8515.
https://scholarworks.umt.edu/etd/8515

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, \& Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

## COPYRIGHT ACT OF 1976

## THIS IS AN UNPUBLISHED MANUSCRIPT IN WHICH COPYRIGHT SUBSISTS. ANY FURTHER REPRINTING OF ITS CONTENTS MUST BE APPROVED BY THE AUTHOR.

Mansfield Library

University of Montana
DATE: 1989

## THE ECONOMIC VALUE OF INSTREAM FLOWS IN MONTANA:

## A TRAVEL COST MODEL APPROACH

By<br>\section*{Christopher J. Neher}

B.S. University of Idaho, 1979.

Presented in partial fulfillment of the requirements
for the degree of

Master of Arts
University of Montana

1989

Approved by


All rights reserved
INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.
In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.

UMI EP39316
Published by ProQuest LLC (2013). Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code


ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346

Ann Arbor, MI 48106-1346

## The Economic Value of Instream Flows in Montana: A Travel Cost Model Approach (143 pp.)

Director: John W. Duffield


The allocation of river flows between consumptive and instream uses is becoming a critical resource problem in Montana. Identifying an efficient allocation is difficult because markets do not exist to provide valuation of recreational uses. This paper derives dollar values for recreational uses of water which are comparable to the dollar prices associated with consumptive uses of water flows.

Nineteen blue-ribbon Montana trout streams were studied. A regional travel cost model was estimated to show the relation of travel distance and mean water flows to river visitation by fishermen. The pooled, crosssectional, time series data came from a fisherman pressure survey conducted by the Montana Department of Fish Wildlife and Parks between 1982-1985, and from U.S. Geological Survey data on water flows.

The major finding is that a strong correlation of mean water flows to fisherman site visitation exists. The model suggests that a hypothetical reduction of water flows by $25 \%$ will result in a reduction of fishing pressure of $9.92 \%$. This change translates into a marginal value of flows of $\$ 1.03$ acre/foot.

Although a significant relationship between flows and visitation was estimated, the limitations of the data and the formulation of the flow variable on an annual basis suggest that the estimated values for water flows may be serious underestimates of actual values. Further research employing more comprehensive pressure data and a longer time series would aid in a more precise estimation of values associated with instream water flows.

## Table of Contents

Abstract ..... ii
List of Figures ..... $v$
List of Tables ..... vi
Acknowledgments ..... vii

1. Statement of the problem ..... 1
1.1. Introduction ..... 1
1.2. Economic Basis of the Problem ..... 2
1.3. Proposed Research ..... 4
1.4. Expected Findings and Thesis Organization ..... 6
2. Review of Literature ..... 10
2.1. Introduction ..... 10
2.2. Assumptions of the Travel Cost Method ..... 10
2.3. Alternative Specifications of Travel Cost Models ..... 13
2.4. Choice of Benefit Estimation Techniques ..... 15
2.5. Summary of Instream Flow Studies ..... 17
3. Modeling Methods ..... 20
3.1. Model Estimation and Calculation of Benefits ..... 20
3.2. The Flow-Visitation Relationship ..... 22
4. Data Sources and Preliminary Analysis ..... 25
4.1. DFWP Fishing Pressure Data ..... 25
4.2. Problems with Pressure Data ..... 30
4.3. U.S. Geological Survey Flow Data ..... 32
4.4. Measures of Central Tendency and Magnitudes ..... 33
4.5. Preliminary Analysis of the Hypothesized Flow Level, Fishing Pressure ..... 41 Relationship
5. Model Estimation and Calculation of Benefits Associated with Alternative ..... 46 Flow Levels
5.1. Travel Cost Model Estimation ..... 46
5.2. Alternative Specification of the Model ..... 49
5.3. Model Prediction ..... 51
5.4. Site-Specific Regression Results and Predictions ..... 54
5.5. Calculation of Benefits Associated with Actual Flow Levels ..... 58
5.6. Hypothetical Flow Variation and its effect on Total Site Benefits ..... 59
5.7. Comparison of Regional and Single-site Models ..... 70
6. Conclusions and Future Research ..... 78
6.1. Overall Conclusions ..... 78
6.2. Suggested Directions for Future Research ..... 79
Appendix A. Data Manipulation and Transformation Processes ..... 81
A.1. Introduction ..... 81
A.2. Merging zip code files with pressure data ..... 81
A.3. State, County, and Unique Water Transformations ..... 83
A.4. Combining Flow Variables with the Pressure File ..... 89
A.5. Ordering and Clustering Visitations ..... 97
Appendix B. Montana DFWP Survey Questionnaire. ..... 132
Appendix C. Residual Analysis of Regional Model Regression Results ..... 134
C.1. Variable Plots ..... 134
C.2. Analysis of Residuals ..... 134

## List of Figures

Figure 3-1: Total Values Associated with Flow Levels ..... 24
Figure 3-2: Marginal and Average Value / Flow Relationship ..... 24
Figure 4-1: Sample Comparison: Big Hole Flows to Historical Averages ..... 43
Figure 5-1: Range of Model Validity in Flow-Value Relationship ..... 73
Figure 5-2: The Average Value - Marginal Value Relationship ..... 73
Figure 5-3: Hypothesized Flow Fishing Pressure Relationship ..... 74
Figure C-1: Plot of $\operatorname{In}(T R I P S)$ and $\ln (D I S T)$ ..... 135
Figure C-2: Plot of $\ln (T R I P S)$ and $\operatorname{In}(Y R F L O W)$ ..... 136
Figure C-3: Histogram of Standardized Residuals ..... 138
Figure C-4: Residual Scatterplots of Independent Variables ..... 139

## List of Tables

Table 1-1: River Codes and Sections for 19 Unique Waters ..... 7
Table 4-1: Wave to Month Correspondence for the Study Years ..... 27
Table 4-2: Total Pressure Survey Response Rates ..... 29
Table 4-3: Survey Responses for the 19 Unique Waters ..... 29
Table 4-4: Total Pressure on 19 Unique Waters ..... 34
Table 4-5: 1982 License Year Pressure Calculations ..... 35
Table 4-6: 1983 License Year Pressure Calculations ..... 36
Table 4-7: 1985 License Year Pressure Calculations ..... 37
Table 4-8: Average River Flow by Time Period ..... 39
Table 4-9: Historical Flows and Standard Deviations ..... 42
Table 5-1: Sample Predictions from Double Log Model ..... 53
Table 5-2: Trip Predictions from Full Sample Model ..... 55
Table 5-3: Regression Results for River Specific Equations ..... 56
Table 5-4: Site-specific predictions for the 4 River Subsample ..... 57
Table 5-5: Consumer Surplus per day: Regional Model ..... 60
Table 5-6: Consumer Surplus per day: Site-specific Models ..... 61
Table 5-7: Total Site Benefits: Regional Mode! ..... 62
Table 5-8: Total Site Benefits: Site-specific Models ..... 63
Table 5-9: Comparison of Hypothetical Flows and DFWP Minimum Flows ..... 65
Table 5-10: Effects of Hypothetical Flow Variation: Regional Model ..... 71
Table 5-11: Acre/Foot Values Associated with a $25 \%$ Decrease in Flows ..... 72
Table 5-12: Effects of 25\% Decrease in Flow Levels:Site-Level Models ..... 75
Table 5-13: Effects of 50\% Decrease in Flow Levels:Site-Level Models ..... 76

## Acknowledgments

Second to the task of writing a thesis is that of reading and critically analyzing one. For this service I am indebted to my committee members John Duffield, Michael Kupilik, and David Patterson. Rob Brooks and Bob MacFarland of DFWP provided invaluable assistance in the compilation of fishing pressure data for this study as did Mel White of USGS with the river flow data. An extra measure of thanks goes to John Duffield for arranging the Montana Water Resources Council grant which made work on this thesis possible. Thanks also to John for his suggestions and criticism throughout.

In a large part this thesis would not have been possible without the friendship and encouragement of Robert Sevy, who showed me at an early date the hidden value of a day spent on a river and has been generous enough to share many such days with me.

Finally, for urging me to return to school, for listening, for asking common sense questions when l could see only abstract details, and for her friendship through it all, I thank Lucy.

## Chapter 1

## Statement of the problem

### 1.1. Introduction

Water is rapidly becoming the pivotal natural resource issue in the Western United States. Competition among uses of this increasingly scarce commodity has forced many states to redefine traditional conceptions of what constitute productive uses of water. The past two decades have seen a dramatic increase in water based recreation and as a result, many states have come to regard, either formally or informally, that maintaining instream flows for recreation is a beneficial use of water. Indeed, this position was explicitly stated in the 1975 Montana Water Use Act. It is significant that recreational uses have gained this equal-footing among many decision makers but unfortunately, this change of status has not obviated all of the problems associated with cost benefit comparisons between recreational and consumptive uses of water. The values associated with consumptive uses of water have long been understood, as these uses are generally marketable and have a market determined price associated with them. The estimation of the values of nonconsumptive, or recreational, uses however is still an emerging area of research.

The objective of this thesis is to estimate the net willingness to pay for fishing on 19 trout streams in Montana as a function of stream flow levels.

Estimates of the values associated with instream flow reservation can then be derived, which are directly comparable to traditional consumptive values for the same waters. The determination of comparable consumptive and recreational values is essential if informed and optimal decisions about water usage are to be made. Current Montana water reservation rules require that estimations of the direct economic benefits of instream flows to protect fisheries be a part of all applications for instream flows. This study is designed to, at least in part, provide the tools necessary for that analysis.

### 1.2. Economic Basis of the Problem

Economics is fundamentally concerned with the issues of scarcity and efficiency. Scarcity, because resources are limited while human wants and the demands they put on those resources know few constraints. Efficiency, since the basic economic question entails determining a use of our resources which maximizes the satisfaction that society, both present and future, derives from that use.

Under the theoretical system of competitive, laissez faire capitalism the basic economic question is rhetorical. In such a system the "market", or the interplay between those supplying and those demanding goods and services, determines the optimal use of resources. It does this through the price mechanism which is ubiquitous to all economic decision making. In a less than perfectly competitive economy such as our own, however, there are many goods and services consumed daily whose worth are not measured adequately, or even at all, by the proxy
yardstick price. These goods and services fall broadly under the label of public goods, or goods which, to one degree or another, have the properties of nonexcludability, and zero marginal cost. it has long been recognized in economic literature that public goods present a special problem to the economic decision maker. Due to the lack of a pricing mechanism this type of good tends to be underproduced in a market economy, resulting in a sub-optimal usage of available resources. Many times government has taken over for this market failure by spending tax dollars on the provision of public goods at a level greater than would be achieved through reliance on the market alone. These efforts, however, have often been made with no true knowledge of what the proper level of public good expenditure should be. Rather they have been roughly guided by the loose and erratically applied political reigns of the republic.

This document addresses the problem of accurately determining the optimal level of resource commitment (in this case instream water flows) for the provision of a public good (fishable populations of trout) in several rivers and streams of Montana.

The estimates of net economic values in this research do not represent the total values associated with the fisheries studied. Several types of values remain unestimated in this study. Specifically, existence, bequest, and option values ${ }^{1}$ will not be examined. These three values best characterize the public good qualities of a fishery resource. Existence values are associated with the utility one derives

[^0]simply from knowing that a high quality fishery exists. Similarly, option value is the amount a person would be willing to pay to maintain the option of fishing a particular river in the future, and bequest value is the value they associate with knowing the fishery will exist for future generations. These three "non-use" values can account for a large portion of the total economic value of a fishery. For catastrophic changes such as the destruction of an entire fishery these values may represent up to $80 \%$ of the total economic value. ${ }^{2}$ This study, however, is primarily interested in changes in flow levels which are deleterious but probably not catastrophic to a fishery, therefore an estimation of recreational use values is appropriate for most policy and resource management decisions pertaining to these fisheries.

### 1.3. Proposed Research

This project was originally proposed and is largely possible because of the existence of an extensive multi-year data base of angler pressure collected by the Montana Department of Fish Wildlife and Parks (DFWP) in the license years 1982-83 through 1985-86. In addition to these monthly pressure surveys a detailed telephone survey was administered in September and October of 1985. This survey included questions on socio-economic status as well as travel time and travel costs related to fishing trips. Combined, these two sources of data provide a high quality picture of fishing behavior in Montana.

[^1]This thesis has as its underlying goal to aid in the identification of optimal levels of instream flow reservation on 19 rivers and streams in Montana (See Table 1-1). In order to arrive at this goal the relationship between observed net economic values of fishery related recreation and historical levels of instream flow must first be identified. This relationship will be explored through the use of a regional (multi-site) travel cost model. Once the flow coefficient is estimated, it can be used to calculate the site specific benefits of alternative hypothetical flow levels. This can be achieved by varying the flow variable while other variables remain constant, and calculating alternative consumer surplus values. This procedure will produce estimated values for varying levels of instream flow reservation which are directly comparable to the marketable consumptive use values for the same waters.

In this study economic benefits for the recreational user are measured in terms of net willingness to pay. This is the difference between the maximum an individual would be willing to pay before foregoing the use of a resource, and the amount they actually must pay. The travel cost method (TCM) is one of two widely accepted methods for calculating net willingness to pay. The other is the contingent valuation method (CVM). Several recent studies have shown the estimates for TCM and CVM to be in the same general range ${ }^{3}$, and both methods are recommended by the U.S. Water Resources Council $(1978,1983)$ for valuing recreational use of resources in federal cost benefit analysis. Just. Heuth and

[^2]Schmitz (1982) describe net willingness to pay as the correct method of measurement in applied welfare economics. Net willingness to pay can be used in the calculation of benefits associated with both marketable and non-marketed commodities, thus its use gives the researcher a theoretically consistent technique with which to measure the varying economic impacts of alternative policy choices.

In the travel cost method observed travel distance is used as a measure of price and the number of trips to a site is used as quantity to statistically construct a demand equation. This resulting "first stage" demand equation can be used to calculate the additional amount a fisheries user would be willing to pay, over their actual incurred travel costs, to have access to the fishery site in question. This additional amount, or net willingness to pay, can be calculated using a "second stage" demand curve that relates additional travel costs to visitation rates. ${ }^{4}$ The travel cost method is used in this study because of the available data bases suitability to this type of modeling.

### 1.4. Expected Findings and Thesis Organization

The hypothesis being examined in this study, is that a non-zero coefficient will be estimated for the flow variable in the travel cost modeling procedure. It seems intuitively obvious that the amount of water in a stream will directly impact fish populations and the quality of fishing experiences to be found at a site. What is not so obvious is how this impact will emerge in the coefficient of flow. While

[^3]| CODE | RIVER | DESCRIPTION |
| :--- | :--- | :--- |
| 80 | Beaverhead | Mainstem |
| $\mathbf{8 1}$ | Big Hole | Mainstem |
| 82 | Bitterroot | Mainstem to E. and W. Forks |
| 83 | Blackfoot | Mainstem |
| 84 | Boulder | Mainstem |
| 85 | Bighorn | Mainstem |
| 86 | Upper Clark Fork | Mainstem above Milltown |
| 87 | Middle Clark Fork | Mainstem below Milltown |
| 89 | Upper Flathead | Flathead Lake to of S. Fork |
| 90 | Gallatin | Mainstem |
| 91 | Kootenai | Mainstem |
| 92 | Madison | Mainstem |
| 93 | Missouri | Mainstem Holter to Cascade |
| 94 | Rock Creek | Mainstem |
| 95 | Smith | Mainstem |
| 96 | Stillwater | Mainstem |
| 97 | Swan | Mainstem |
| 98 | Up. Yellowstone | Springdale to Gardner |
| 99 | Mid. Yellowstone | Springdale to Bighorn R. |

Table 1-1: River Codes and Sections for 19 Unique Waters
higher than average flows in the late summer might increase the attractiveness of a fishery, lower than average flows in the late spring might foster the same attractiveness. Any ambiguities which the dynamic yearly flow cycle of the study rivers might present should be avoidable by analyzing the data as individual months or month clusters. Whatever the level of disaggregation necessary to estimate a good model, it is expected that positive values will be shown as associated with instream flow reservation in order to protect fisheries and the recreation they foster.

This thesis is organized in six chapters and three appendices. The scope of each of these sections is detailed below.

Chapter 2 - Review of Relevant Literature

This chapter examines the current body of literature on both travel cost methodology and instream flow issues. A discussion of the development and refinement of travel cost modeling procedures is included here along with an examination of several studies in the relatively new area of instream flow analysis.

## Chapter 3 - Modeling Methods

The basic form of the hypothesized flow-visitation relationship is presented here along with the assumptions relevant to that hypothesis.

Chapter 4 - Data Sources and Preliminary Analysis

In this section the sources of and problems relating to the data used in this study are examined. The preparation of DFWP data for the travel cost modeling procedure entailed a significant effort and Chapter 3 in conjunction with Appendix $A$ detail the procedures and assumptions relevant to this effort. Additionally, an initial analysis of the relationship between fishing pressure and stream flows is discussed in this chapter.

Chapter 5 - Travel Cost Model Assumptions and Application

This chapter begins with a discussion of the procedures followed in the application of the model to the data base as described in Chapter 3. The results of the model regressions are shown and the statistical significance of these results are discussed. In a second major portion of this chapter the calculation of dollar values relating to the model is shown and discussed. A final section looks at changes in net valuation associated with hypothetical fluctuations in flow levels.

Chapter 6 - Conclusions and Areas of Future Research

This final chapter draws the major conclusions possible from the preceding analysis. Suggestions are made for the appropriate direction of future research.

## Chapter 2

## Review of Literature

### 2.1. Introduction

Two major areas of inquiry are combined in the analysis contained in this study. Travel cost modeling is a mature area of research with a large and growing body of literature associated with it. Instream flow studies, however, are a much rarer breed. This chapter is intended to give an overview of the significant advances in, and applications of, the travel cost method since it was first suggested by Harold Hotelling 40 years ago. Additionally, the current body of work relating to instream flow issues will be discussed with emphasis on those studies examining the valuation of instream flows and optimization of those flows.

### 2.2. Assumptions of the Travel Cost Method

The underlying assumption of the travel cost method is that the travel costs incurred by the recreationist to get to a recreation site can be interpreted as the price they are willing to pay to visit that site. This assumption can be considered valid only if two factors are addressed in the definition of travel costs. First, trips with multiple destinations must be excluded from any analysis of the travel cost model. Only trips made for the sole purpose of accessing a recreation site must be included (Dwyer, Kelly and Bowes 1977). If the recreationist stopped at several destinations on one trip it would be impossible to divide travel costs among each
of the destinations appropriately. In the DFWP survey respondents were asked if the main purposes of their trip was to fish. Those who responded "no" were excluded from the sample. A "no" response indicated that the respondent would have likely made the trip even if a fishing opportunity had not been available. Consequently, inclusion of their responses in the data base would have biased benefit estimates upward.

The second qualifying factor necessary for the accurate interpretation of travel costs as site access prices relates to the issue of travel time. The old saw is quoted "time is money", and the application of the travel cost method is no exception. The formulation of the price variable in this study was made to include both the variable costs of travel associated with operating a vehicle and the opportunity cost of time spent traveling. Cesario's (1976) work in transportation planning indicates that time spent traveling is indeed viewed as costly. Dwyer, Kelly and Bowes (1977) urge the inclusion of the value of travel time in the TCM in order to avoid a significant underestimation of the benefits associated with recreational activities. The U.S. Water Resources Council $(1979,1983)$ drew heavily on Cesario's (1976) work in formulating their recommendations for the valuation of travel time. They suggest that one-half to one-fourth the wage rate be used as a measure of the opportunity cost of time. This study, which follows closely the methodologies of Duffield, Loomis and Brooks (1987) uses one-third the wage rate in accordance with U.S. WRC recommendations.

Dwyer, Kelly and Bowes (1977) point out two additional assumptions of the travel cost method: (1) All relevant variables are specified in the model and (2) there is no unobserved demand due to capacity restrictions.

In their study of cold water fishing in Montana Duffield, Loomis and Brooks (1987) found that two explanatory variables usually included in travel cost analysis, income and a measure of the attractiveness of substitute sites, were not significant in their final unique waters model. This study, drawing heavily on the results of Duffield et al. (1987), made no attempt to include income and a substitute variable in its analysis since both methodology and data sources used in this study correspond to those used in the 1987 study. While the exclusion of substitutes and demographic variables from this analysis deviates from procedures used in many TCM studies Dwyer, Kelly, and Bowes (1977) point out that "[T]he process of choosing significant variables and correctly specifying their interrelationship is partly an art" (p83).

The assumption relating to capacity restrictions provides no obstacles for this model. During the time period covered by the data base no restrictions other than self imposed ones caused by congestion existed on the study rivers. There is a large body of literature pertaining to the effects of congestion on recreational participation, however, to include the effects of congestion in this analysis would be beyond the scope of this study.

Another assumption of the ordinary least squares modeling procedure which will be explored in more depth in chapter 4 , is the assumption of a constant variance for the dependent variable. Bowes and Loomis (1980) found that the unequally weighted population zones of the zonal travel cost method led to a problem of heteroscedasticity. They suggested an application of weighted least squares to the demand equation in order to correct for this problem. Duffield,

Loomis and Brooks (1987) chose a double log specification of the TCM in order to minimize the effects of heteroscedasticity. Their relative success in this effort will be further explored in chapter 4.

### 2.3. Alternative Specifications of Travel Cost Models

Since the early work of Clawson (1959) and Knetsch (1963) there have been many variations of the travel cost model. Two major genres of TCM's are the zonal and the individual models.

In the zonal model (Cesario and Knetsch, 1976) visits to a site are aggregated by equidistant zones and visits are regressed on distance to trace out a demand relationship. This model makes two major assumptions: (1) it is assumed that tastes are homogeneous within origin zones, and (2) the assumption is made that travel costs are constant across zones.

Duffield, Loomis and Brooks (1987) utilized this zonal model and the current study follows their lead by regressing visits per capita on round trip distance. As in the 1987 study round trip distance is used as a composite variable including both variable travel costs and the value of travel time. This specification sidesteps problems of collinearity of the independent variables travel cost and travel time. This collinearity problem proved to be a disadvantage of early specifications of the zonal modet.

The individual model (Gum and Martin 1975; Brown and Nawas 1973) does not make the assumptions of homogeneous tastes and constant travel costs made by the zonal models. This disaggregated model sums visits per individual over a
specific time period and regresses them on costs to estimate a demand function. There are several advantages of the individual model. First, the individual specification eliminates problems of nonconstant variance of the error term associated with the dependent variable. Second, the use of individual costs induces more precision into the coefficient estimation. The major drawback of the individual model lies in its inability to correct for participation levels. This inability leads to significant overestimation of benefits.

One advantage of the zonal model is that estimates of consumer surplus may be expanded to represent the entire populations of the zones of origin. In general practice however the form and quality of the investigators data base often dictates the choice of an appropriate model. Sorg and Loomis (1985) ${ }^{5}$ point out that the U.S. Water Resources Council has recommended that regional, or large multi-site, models be used in place of single site models. They add that while regional, or large multi-site, models may not be as accurate as single site models in identifying simple benefits of an existing site, for many planning issues the regional approach is desirable.

[^4]
### 2.4. Choice of Benefit Estimation Techniques

The calculation of benefits (consumer surplus) in this study is a straightforward application of the technique of Duffield, Loomis and Brooks (1987). No attempt was made to modify or improve upon their estimates of variable travel costs or their formulation of consumer surplus calculations.

Duffield et al. (1987) calculated variable travel cost using 2 methods. Initially a Water Resources Council (1983) method was used. According to the Department of Transportation report (1985) 6 cents per mile was chosen as the representative variable cost associated with the typical "fishing vehicle". This amount was summed with 9.2 cents variable gas and tax costs per mile and divided by an average of 2.76 persons per vehicle to arrive at a variable travel cost of 5.6 cents per mile. This amount was added to the calculated opportunity cost of time to get the standard variable cost figure of 12.6 cents per mile. ${ }^{6}$

Duffield et al. (1987) felt that the costs derived from the Department of Transportation information made assumptions which were not necessarily applicable to Montana anglers. Specifically, the DOT information assumes the use of a new car while the average age of cars in the U.S. is 7.4 years. Also, the DOT does not consider the less than ideal driving conditions encountered by many Montana fishermen (i.e. winding mountain roads, dirt roads). Finally, the DOT does not consider the large number of low-gas-efficiency vehicles ( 4 wheel drives, pickup campers) used by fishermen in Montana.

[^5]In an effort to get around these shortcomings of the DOT method Duffield, Loomis and Brooks (1987) included in their 1985 phone survey questions regarding trip expenditures, travel costs and the value of travel time. They found that a variable vehicle cost of 22.6 cents per individual was reported by respondents and that this travel cost was homogeneous across origin zones in accordance with the assumptions of the zonal model. Their queries on valuation of travel time led them to a valuation equal to one fifth of the wage rate, or 4.6 cents per mile. The total travel cost reported by Montana fishermen was therefore 27 cents per mile. These actual reported costs are used rather than the less realistic DOT estimates in the analysis to follow.

When the per capita demand equation has been estimated through the use of ordinary least squares regression there are two major methods available for calculating benefits from the regression results. Integration of the first stage demand curve between the current observed distance and some measure of a maximum distance which would cause visitation to fall to less than one could be done in order to calculate net willingness to pay for each origin zone. Total site benefits could then be calculated as the population weighted sum of each zones net willingness to pay. This is the method that was used in the current study. Selection of an appropriate upper limit for the integration was drawn from Duffield's supplemental technical paper to the 1987 study, concerning sensitivity analysis relating to the model specification (Duffield 1987b). The integration limit chosen was the maximum observed travel distance to a specific site plus the current distance.

A second method of calculating consumer surplus associated with a recreation site is that detailed exhaustively by Dwyer, Kelly and Bowes (1977). This method entails the derivation of a second stage, or site, demand curve which relates visitation to travel costs over and above existing travel costs. The area under this second stage curve represents net willingness to pay. Burt and Brewer (1977) and Menz and Wilton (1983) have demonstrated that these two alternative methods of calculating consumer surplus are equivalent.

### 2.5. Summary of Instream Flow Studies

Daubert and Young (1981) studied the Poudre River in northern Colorado in an attempt to derive a marginal value of water flows to recreation. They selected a Contingent Valuation approach for their study which utilized an iterative bidding game. What they found was that total willingness to pay for flows by fishermen followed a non-linear relationship increasing over a certain range and then decreasing. The marginal willingness to pay for flows by fishermen was therefore a linear function decreasing with increased flows. Daubert and Young estimated the acre/foot value of water flows to recreationist at levels comparable to values for consumptive uses (irrigation) of the same water. In general they found that diminishing marginal utility of flows was found for fishermen but not for whitewater enthusiasts. ${ }^{7}$

[^6]Narayanan (1986) outlined three alternative methods for determining demand for instream flows. One method is that followed by Daubert and Young (1981) and Boyle, Walsh, Bishop, and Baumgartner (1987) of asking people (using various CVM techniques) what they would be willing to pay for an increase, or to avert a decrease, in stream flows. A second method Narayanan suggests is the use of cross-sectional data on household characteristics and visitation for multiple sites. He notes, however, that it is difficult to use this cross-sectional analysis to evaluate demand for flows. An extension of this approach would be to estimate site-specific demand equations including an interaction term in each equation consisting of price and flow variables. This model attributes differences in demand not explained by the model to varying flow levels.

Narayanan (1986) used a combination of a cross-sectional travel cost model and a CVM survey in his model. The travel cost model was used to estimate recreation demand on the Blacksmith Fork river in Utah. A Contingent Valuation survey was then used to determine the effect flow reductions would have on visitation rates. In general Narayanan found positive marginal values associated with averting reductions in average flows. The values were on the average an order of magnitude smaller than those found by Daubert and Young (1981).

A third method of determining demand for instream flows, suggested by Narayanan (1986), is the use of time series data for a site to examine changes in recreation demand as a function of stream flows. Narayanan points out a dilemma associated with the use of time series data. Often if a long time series is used there can be structural changes in the series which will cause invalid coefficients
to be estimated. On the other hand, the use of a short time series often will not capture the effects of variations in flows adequately.

This study relies upon a short time series to estimate demand for flows. The ability to use a short time series enabled the assumption to be reasonably made that estimated coefficients are stable across time periods (Judge et al. 1980).

## Chapter 3

## Modeling Methods

### 3.1. Model Estimation and Calculation of Benefits

The process of examining the hypothesized flow-visitation relationship can be broken down into four stages.

1. Estimation of the model through the application of ordinary least squares regression.
2. Testing the reliability of the model through examining its predictive powers.
3. Calculation of the benefits associated with flows in the estimated model.
4. Examination of the effects of hypothetical shifts in flow levels on the benefits calculated from the model.

The first three of these steps follow closely the methods of Duffield, Loomis and Brooks (1987). The functional form of the travel cost model which they found most significant (the double log specification) will be first examined. Additionally, the straightforward methods which they employed in the testing of model prediction and the calculation of benefits from the model will be used in this study as well. The additions and modifications to the work of Duffield. Loomis and Brooks (1987) which this study makes are three:

1. The method used in this study for defining individual trips deviated from that used in the 1987 study. Due to the nature of the DFWP data base the average number of days per trip calculated in this
study was two times that calculated by Duffield et al. This deviation from the previous study suggests that there is an underestimation of trips associated with the methods used in this study.
2. The use of pooled time series and cross sectional data is a deviation from the single year study of Duffield et al. The use of pooled data is an innovative use of the travel cost model.
3. The addition of a flow variable into the study is an extension of the 1987 study. The inclusion of a flow variable (which necessitated the use of time series data) and the measurement of its relationship to visitation is the primary objective of this thesis.

An additional area in which the current study deviates from Duffield, Loomis and Brooks' 1987 analysis is the exclusion of demographic and site attribute variables (other than flow) from the model. Duffield et al. found significance in the inclusion of several of these measures. The explanatory power of the present model could doubtless have been improved by the inclusion of some, or all, of these explanatory variables. To do so, however, would have been beyond the scope of this study. It must be noted that excluding possibly significant site and demographic variables from this model could introduce an omitted variable bias into the results. Koutsoyiannis (1977) describes the possible effect of this bias as the overestimation of the residual variance and overestimation of the standard errors of the estimates.

### 3.2. The Flow-Visitation Relationship

There are several a priori assertions which we are able to make regarding the relationship of flows to recreational values.

Assertion (1)

$$
\text { If flows }=0 \text { then Total Value }=0
$$

This assertion is intuitively obvious. Without water there can be no fish. Without fish, no fishing, thus no value associated with fishing.

## Assertion (2)

$$
\text { If flows }=\infty \text { then Total Value }=0
$$

In this assertion $\infty$ is meant not to represent an infinite amount of water per se, but some flow beyond the highest observed historical floods. Assertion (2), like (1), has a reasonable basis. When rivers are in a normal seasonal flood most carry a heavy silt load which reduces visibility in the water, fishing success, and therefore fishing pressure. To assert that flows could exist where this detrimental effect of flood waters reduces fishing pressure to zero is a reasonable extension of the normal effects of high flows.

Assertion (3)

$$
\begin{gathered}
\text { Where } 0<\text { flows }<\infty \\
\text { then } \\
\text { Total Value }>0
\end{gathered}
$$

Assertion (3) states that the value of recreational fishing is positive when flows are between zero and infinity. Logic would suggest that this is a natural extension following from assertions (1) and (2).

The exact shape of the flow-total value relationship is indeterminate. Drawing from the work of Daubert and Young (1982) and Narayanan (1986), however, we can make the assumption that the relationship depicted in figure 3-1 is not unreasonable. Specifically, we can assert that there is a global maximum associated with the relationship. Figure 3-2 shows the marginal and average value curves associated with figure 3-1. The calculation of marginal and average values associated with flows, and the relationship of these values to one another will, at the least, be indicative of what region of the total value curve this study is effectively modeling.


Figure 3-1: Total Values Associated with Fiow Levels


Figure 3-2: Marginal and Average Value / Flow Relationship

## Chapter 4

## Data Sources and Preliminary Analysis

### 4.1. DFWP Fishing Pressure Data

The data used for this study came from two sources. Data on fishing pressure was originally gathered by the Montana DFWP while streamflow data came from U.S. Geological Survey records.

The data on fishing pressure was collected by the DFWP in the years 1982-1985 in order to estimate yearly fishing pressure on rivers, streams and lakes throughout Montana. A questionnaire, designed by DFWP, (see Appendix B) was mailed to a sample of fishing license holders each month during the four study years. This survey collected information on where respondents fished, the number of fish they caught and kept, the main purpose of their trip, whether they stayed overnight, and the round trip distance they traveled. Not all of these categories were present in the first two study years, but the variables, where fished, number of days, and anglers location of residence (origin) were present in all survey responses. This lack of consistency across years limited the depth of analysis possible, but the data was sufficient for the construction of a regional (multi-site) travel cost model.

In order for the raw fishing pressure data to be usable for travel cost modeling, extensive transformations and manipulations of that data were
necessary. These transformations were primarily mechanical in nature and no significant assumptions relating to the modeling process were made in their course. A detailed exposition of the programming steps used in these transformations is included in appendix $A$ for the interested reader.

The DFWP pressure survey was a stratified survey which was performed on a monthly or semi-monthly basis. Each of these mailings was called a wave, and the prospective respondents were randomly sampled from the entire fishing license population as it was known at that time. The correspondence of waves to months for the four study years is shown in Table 4-1.

In 1984 the license year was changed from starting on May 1 to starting on March 1. The months of March and April 1984 were included in both the 1983 and 1984 pressure files to ensure that both years samples covered a full 12 month period. Also in 1984 the sampling procedure was done semi-monthly rather than monthly in an effort to minimize recall bias. This change when combined with another change in the sampling timing during the summer months of 1984 led to a serious downward bias in total pressure figures for those few months. This problem and how it was dealt with will be discussed more fully below.

In addition to the monthly, or semi-monthly, sampling of resident and nonresident season fishing license holders an end of the year sampling was done of all non-resident 2-day fishing license holders. These 2-day licenses make up more than two thirds of non-resident fishing pressure in a typical year.

In 1982 the DFWP sampled 1500 resident and 110 non-resident license holders per wave. Additionally, 1200 non-resident 2-day license holders were

| Wave | 1982 | 1983 | 1984 | 1985 |
| :---: | :---: | :---: | :---: | :---: |
| 1 | JAN 83 | JAN 84 | MAR1 84 | MAR 85 |
| 2 | FEB 83 | FEB 84 | MAR2 84 | APR 85 |
| 3 | MAR 83 |  | APR1 84 | MAY 85 |
| 4 | APR 83 |  | APR2 84 | JUN1 85 |
| 5 |  |  | MAY 84 | JUN2 85 |
| 6 | MAY/JUN 82 | MAY/JUN 83 | JUN1 84 | JUL1 85 |
| 7 | JUL 82 | JUL 83 | JUN2 84 | JUL2 85 |
| 8 | AUG 82 | AUG 83 | JUL1 84 | AUG1 85 |
| 9 | SEP 82 | SEP 83 | JUL2 84 | AUG2 85 |
| 10 | OCT 82 | OCT 83 | AUG1 84 | SEP1 85 |
| 11 | NOV 82 | NOV 83 | AUG2 84 | SEP2 85 |
| 12 | DEC 82 | DEC 83 | SEP1 84 | OCT1 85 |
| 13 |  |  | SEP2 84 | OCT2 85 |
| 14 |  | MAR1 84 | OCT1 84 | NOV 85 |
| 15 |  | MAR2 84 | OCT2 84 | DEC 85 |
| 16 |  | APR1 84 | NOV 84 | JAN 86 |
| 17 |  | APR2 84 | DEC 84 | FEB 86 |
| 18 |  |  | JAN 85 |  |
| 19 |  |  | FEB 85 |  |
| 99 | 2-DAY | 2-DAY | 2-DAY | 2-DAY |

Table 4-1: Wave to Month Correspondence for the Study Years
sampled at the end of the year. The ratio of these numbers was based upon past surveys indicating a $80 / 20$ split between resident and non-resident fishing pressure. In 1983 the sampling was done much the same as in 1982 with two exceptions. The non-resident 2-day license sample was doubled to 2400 , and March and April were sampled semi-monthly at half the previous rate. 1984 saw the sampling change to semi-monthly in the higher pressure months (June through October) in an attempt to reduce recall bias. The 1985 sampling stayed with the 1984 format but increased the sampling rate to 3000 residents and 220 non-residents per month. This doubling of the rate was adopted to accommodate the needs of the Montana Bioeconomic Study undertaken in that year. Table 4-2 shows the questionnaire totals and response rates for the 1982-1985 license years. These response rates included both those respondents who had fished in the time period covered by the questionnaire, as well as those who had not fished. The significant drop in response rates in 1985 can be attributed to the DFWP going to a bulk rate mailing system rather than the first class used previously.

Since the data collected by DFWP included information on fishing on all Montana waters it was necessary to select out only those responses pertaining to the 19 rivers to be dealt with in this study. The year specific sample sizes for this subsample of the total pressure file are shown in Table 4-3.

|  |  | MAILED | RETURNED | FISHED | \% RET. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1982 | RESIDENT | 16860 | 12436 | 3705 | 74\% |
|  | NON-RES | 2240 | 1435 | 830 | 64\% |
| 1983 | RESIDENT | 17438 | 10977 | 2584 | 63\% |
|  | NON-RES | 3531 | 2671 | 1697 | 76\% |
| 1984 | RESIDENT | 17873 | 11119 | 2303 | 62\% |
|  | NON-RES | 3847 | 2234 | 1127 | 58\% |
| 1985 | RESIDENT | 36969 | 15277 | 3698 | 41\% |
|  | NON-RES | 7914 | 3834 | 2202 | 48\% |

Table 4-2: Total Pressure Survey Response Rates

|  |  |  | FULL YEAR | SUM.MONTHS |
| :---: | :---: | :---: | :---: | :---: |
| 1982 | RESIDENT | ....... | 1288 | 738 |
|  | NON-RES | .......... | 422 | $90^{\prime}$ |
| 1983 | RESIDENT | .......... | 802 | 487 |
|  | NON-RES | .......... | 802 | 80 |
| 1984 | RESIDENT | ..... | 705 | 371 |
|  | NON-RES | .......... | 420 | 38 |
| 1985 | RESIDENT | $\ldots$ | 1269 | 468 |
|  | NON-RES | .......... | 1043 | 75 |

${ }^{1}$ Nonresident Summer figures do not include 2-day respondents

Table 4-3: Survey Responses for the 19 Unique Waters

### 4.2. Problems with Pressure Data

The data set received from DFWP, while extensive, was not without problems. The most compromising problem, which has been previously mentioned, concerns a sampling bias during the summer months of $1984 .^{8}$ At the end of each month fishing license vendors are required to send in that months licenses and remittances to Helena, Montana. From Hefena copies of the licenses are forwarded to DFWP headquarters in Bozeman. On the 20th of the month, after all of the previous months newly sold licenses are received by DFWP the sample for the previous months pressure is drawn and mailed. This method was used consistently through 1982, 1983, and early 1984. In the summer month's of 1984, however, the samples were drawn on the first of the month for the preceding months pressure estimate. This meant that the samples were drawn before the most recently sold licenses (whose holders are most likely to have fished during the month) had been entered into the sample population. The DFWP statistician in charge of administering the survey agrees that this change resulted in a serious downward bias in those months pressure estimates. Since there was no way to correct for this bias the 1984 data was dropped from the travel cost analysis.

The final problem with the pressure data arose in regard to differences in aggregation techniques used on the individual study years. The pressure files for 1982 and 1983 were, when they were received for this study, aggregated according

[^7]to the following scheme. Each record in these files detailed the person doing the fishing, the water that was fished, and the number of days that it was fished. The records in these years were aggregated such that records showing the same individual fishing the same water in one particular wave in subsequent records were combined into one record containing the sum of days fished in those multiple records. Since the goal of the travel cost model is to estimate trips as a function of round trip distance, each record of this aggregated file was taken to represent one multi-day trip. The effect of this aggregation scheme was likely to underestimate the number of actual trips made to specific rivers. It may be reasonable to assume that non-residents fishing a specific river in a particular month did so on one multi-day trip from their home, but this assumption is not nearly so valid for residents, who may live but a few miles away from their river destination. A comparison of the 1985 data aggregated according to the DFWP: method with that same data aggregated by Duffield, Loomis and Brooks in their 1987 study shows significantly different results. For the 19 unique waters Duffield et al. found a total of 2837 trips in their sample. The same data using the DFWP aggregation shows a total of 2312 trips or nearlv $20 \%$ fewer. The overall effect of this difference is one of conservatively estimating the number of trips taken and should lead to a defensible lower bound for any valuation estimates in this study.

For the sake of consistency the same aggregation scheme used by DFWP in the 1982 and 1983 data was used for the 1985 data. It was felt that in the absence of any way to disaggregate 82 and 83 the best method to follow was to at least ensure consistency across all the study years.

### 4.3. U.S. Geological Survey Flow Data

The U.S. Geological Survey publishes a yearly compendium of recorded daily water flows for each of their water gauging stations in Montana. It was from these records that the flow data for this study was collected. Gauging stations were chosen on each of the 19 study rivers for their proximity to the most heavily fished sections of those rivers. Initially, a file was created which contained a unique water code identifying each particular river, and an mean monthly flow for each month of the four study years. Additionally, a mean yearly flow variable was added to each record along with mean flows for two month clustering schemes. Finatly an identifying variable was attached to each record to indicate whether flows in that particular river were completely free of, mildly controlled by, or heavily controlled by releases from upstream reservoirs.

Once completed, the river flow file was merged with the DFWP fishing pressure file. Before this was possible, however, it was necessary to choose a standard for time measurement between the months used in the flow file, and the waves used in the pressure file. It was decided that since the waves most closely represented the fishing year, they would be used as a standard. References to specific years in this paper, (eg. 1984) therefore, should be understood to mean the license year 1984-85 as indicated in Table 4-1. The actual procedure for merging the flow file with the pressure file is described in detail in Appendix $A$.

### 4.4. Measures of Central Tendency and Magnitudes

The 19 river sections included in this analysis constitute approximately 25 percent of the total fishing pressure in Montana. That pressure, which totals over 2 million fishing days per year, is roughly divided between 80 percent resident and 20 percent non-resident fishing. Table 4-4 illustrates how the total number of fishing days on these 19 rivers is split between the resident, non-resident, and 2day sampling subgroups. In Table 4-4, average days per year is the sum of weighted means of the samples drawn from each wave in a given year. The calculation of these weighted means and the total use days for each of the three study years is shown in Tables 4-5 through 4-7.

Flow levels for the 19 rivers are measured in cubic feet per second (CFS). The monthly flow levels in the data set are the means of the daily flow levels. Table 4-8 shows average flow levels for the study rivers. These averages are calculated on a yearly level as well as for the period May-September. The partial year averages were calculated in order to more accurately identify the periods relating to the heaviest fishing pressure and the greatest fluctuation in flow levels. In addition to the flow averages Table 4-8 shows a control level assigned to each river. This level (which takes a value of 1,2 or 3 ) tricotomizes the rivers according to the degree to which their flows are controlled by reservoir releases. A control level of 1 designates a free flowing stream with no control by dam releases. Control levels of 2 and 3 designate rivers which have moderate or significant flow control respectively. Reservoir control has the effect of smoothing out high and low river levels by creating a more even intertemporal distribution of a river

## AVERAGE DAYS/YEAR TOTAL FISHING DAYS

| RESIDENTS |  |  |
| :--- | :--- | :--- |
|  |  |  |
| 1982 | 2.65 | 419750 |
| 1983 | 2.18 | 380227 |
| 1985 | 2.61 | 427881 |
|  |  |  |
| NON-RESIDENTS |  | 62247 |
|  |  | 43980 |
| 1982 | 5.62 | 38621 |
| 1983 | 4.05 |  |
| 1985 | 4.30 |  |
|  |  | 86588 |
| $2-D A Y$ |  | 122166 |
|  |  | 96999 |

## TOTAL DAYS FISHED

1982
.568585

1983 ............... 546372
1985 ............... 563501

Table 4-4: Total Pressure on 19 Unique Waters

1982
\(\left.\begin{array}{llll}RESIDENT \& \& <br>
WAVE \& ELIGIBLE <br>

\& \& POOL\end{array}\right]\) USE DAYS |  |
| :--- |
| 1 |

NON-RESIDENT

| 1 | 0.01 | 12727 | 141 |
| :--- | :--- | :--- | :--- |
| 2 | 0.01 | 12740 | 172 |
| 3 | 0.03 | 12753 | 440 |
| 4 | 0.27 | 12766 | 3408 |
| 6 | 0.84 | 8846 | 7425 |
| 7 | 1.52 | 10391 | 15821 |
| 8 | 2.02 | 11668 | 23601 |
| 9 | 0.68 | 12344 | 8376 |
| 10 | 0.08 | 12510 | 957 |
| 11 | 0.15 | 12663 | 1906 |
| 12 | 0.00 | 12689 | 0.00 |
|  | $\underline{5.62}$ |  | 62248 |
|  |  |  |  |
| $2-$ DAY |  |  | 86589 |
| 99 | 0.80 | 108225 |  |

Table 4-5: 1982 License Year Pressure Calculations

|  | 1983 |  |  |
| :--- | :---: | :---: | :---: |
| RESIDENT |  |  |  |
| WAVE | ELIGIBLE |  |  |
|  |  | POOL | USE DAYS |
| 1 | 0.09 | 187220 | 17311 |
| 2 | 0.09 | 187408 | 17069 |
| 6 | 0.51 | 152550 | 78118 |
| 7 | 0.41 | 171290 | 69723 |
| 8 | 0.35 | 181223 | 63739 |
| 9 | 0.18 | 183659 | 33888 |
| 10 | 0.13 | 185908 | 23289 |
| 11 | 0.05 | 186283 | 9844 |
| 12 | 0.04 | 186470 | 7248 |
| 14 | 0.10 | 186658 | 17981 |
| 15 | 0.06 | 186845 | 10982 |
| 16 | 0.11 |  | 21426 |
|  |  |  |  |

NON-RESIDENT

| 1 | 0.01 | 11746 | 130 |
| :--- | :--- | :--- | :--- |
| 2 | 0.04 | 11758 | 424 |
| 6 | 1.06 | 9571 | 10169 |
| 7 | 0.82 | 11370 | 8464 |
| 8 | 0.82 | 11370 | 9290 |
| 9 | 0.43 | 11522 | 4919 |
| 10 | 0.29 | 11663 | 3350 |
| 11 | 0.07 | 11687 | 770 |
| 12 | 0.01 | 11699 | 118 |
| 14 | 0.00 | 11710 | 0.00 |
| 15 | 0.16 | 11722 | 1823 |
| 16 | 0.16 | 11734 | 1853 |
| 17 | 0.23 | 11746 | 2667 |
|  |  |  | 43980 |
| $2-$ DAY | 4.05 |  |  |
| 99 |  |  | 111364 |

Table 4-6: 1983 License Year Pressure Calculations

1985
RESIDENT

| WAVE | AVE.DAYS | $\begin{gathered} \text { ELIGIBLE } \\ \text { POOL } \end{gathered}$ | USE DAYS |
| :---: | :---: | :---: | :---: |
| 1 | 0.35 | 44490 | 15389 |
| 2 | 0.35 | 82283 | 28872 |
| 3 | 0.28 | 147642 | 41659 |
| 4 | 0.17 | 168609 | 29366 |
| 5 | 0.16 | 189575 | 30135 |
| 6 | 0.24 | 199037 | 48363 |
| 7 | 0.19 | 208498 | 38593 |
| 8 | 0.18 | 213455 | 38593 |
| 9 | 0.14 | 218412 | 30881 |
| 10 | 0.11 | 221456 | 24271 |
| 11 | 0.08 | 224498 | 18184 |
| 12 | 0.09 | 227402 | 21235 |
| 13 | 0.05 | 230204 | 11634 |
| 14 | 0.05 | 232617 | 10580 |
| 15 | 0.06 | 233469 | 14654 |
| 16 | 0.03 | 234675 | 7252 |
| 17 | 0.07 | 235163 | 17042 |
|  | 2.61 |  | 427882 |

NON-RESIDENT

| 1 | 0.39 | 1006 | 397 |
| :--- | :--- | :--- | :--- |
| 2 | 0.70 | 2226 | 1550 |
| 3 | 0.46 | 4804 | 2212 |
| 4 | 0.22 | 6955 | 1545 |
| 5 | 0.30 | 9106 | 2742 |
| 6 | 0.57 | 10955 | 6208 |
| 7 | 0.59 | 12804 | 7606 |
| 8 | 0.22 | 13904 | 3082 |
| 9 | 0.33 | 15003 | 4985 |
| 10 | 0.11 | 15470 | 1751 |
| 11 | 0.11 | 16936 | 1783 |
| 12 | 0.17 | 16055 | 2702 |
| 13 | 0.03 | 16174 | 564 |
| 14 | 0.04 | 16336 | 648 |
| 15 | 0.01 | 16487 | 242 |
| 16 | 0.01 | 16567 | 193 |
| 17 | 0.02 | 16683 | 412 |
|  |  |  | 38622 |
| $2-$ OAY | 4.30 |  |  |
| 99 |  |  | 96999 |

Table 4-7: 1985 License Year Pressure Calculations
drainages water supply. It is anticipated that the introduction of a control level variable will allow the analysis of river subgroups in which the effects of drought and flood are most acutely felt.

Table 4-9 is presented in order to give a feel for how the 3 study years (1982, 1983, and 1985) relate to historical averages on the 19 rivers. Included in table 4-9 are the historical means of yearly flows for the rivers as well as the standard deviations around those means. Figure 4-1 shows an example of how one river (the Big Hole) compares to historical means and variations in flow.

A cursory comparison of the flows in table 4-8 with the historical averages shown in table 4-9 indicates that in general 1982 was a higher than average flow year, 1983 was an average flow year, and 1985 was a lower than average flow year. It is interesting to note that the means and standard deviations of flows for the three study years correspond very closely to the historical means and standard deviations on almost all rivers. This indicates that the study years are very representative of historical flows.

| RIVER | YEAR | $\begin{aligned} & \text { YEAR } \\ & \text { AVERAGE } \end{aligned}$ | MAY-SEPT. AVERAGE | CONTROL LEVEL ${ }^{1}$ |
| :---: | :---: | :---: | :---: | :---: |
| BEAVERHEAD | 1982 | 627 | 838 | 3 |
|  | 1983 | 735 | 854 |  |
|  | 1985 | 401 | 640 |  |
| BIG HOLE | 1982 | 1595 | 2982 | 1 |
|  | 1983 | 1301 | 2050 |  |
|  | 1985 | 795 | 1053 |  |
| BITTERROOT | 1982 | 1169 | 2325 | 2 |
|  | 1983 | 792 | 1243 |  |
|  | 1985 | 629 | 994 |  |
| BLACKFOOT | 1982 | 1965 | 3639 | 1 |
|  | 1983 | 1361 | 2095 |  |
|  | 1985 | 1261 | 1839 |  |
| BOULDER | 1982 | 646 | 1295 | 1 |
|  | 1983 | 641 | 1048 |  |
|  | 1985 | 371 | 663 |  |
| BIGHORN | 1982 | 4173 | 3921 | 3 |
|  | 1983 | 4839 | 5570 |  |
|  | 1985 | 2480 | 1953 |  |
| UPPER | 1982 | 3996 | 6850 | 1 |
| CLARK FORK | 1983 | 2878 | 4023 |  |
|  | 1985 | 2352 | 2950 |  |
| MIDDLE | 1982 | 9503 | 16846 | 1 |
| CLARK FORK | 1983 | 6744 | 9664 |  |
|  | 1985 | 6090 | 8247 |  |
| UPPER | 1982 | 10058 | 15688 | 3 |
| FLATHEAD | 1983 | 9857 | 13279 |  |
|  | 1985 | 10133 | 14682 |  |

Table 4-8: Average River Flow by Time Period

| GALLATIN | 1982 | 1556 | 2371 | 2 |
| :---: | :---: | :---: | :---: | :---: |
|  | 1983 | 1439 | 1899 |  |
|  | 1985 | 841 | 809 |  |
| KOOTENAI | 1982 | 12639 | 10298 | 3 |
|  | 1983 | 11797 | 10580 |  |
|  | 1985 | 9778 | 6760 |  |
| MADISON | 1982 | 2166 | 2697 | 3 |
|  | 1983 | 2200 | 2403 |  |
|  | 1985 | 1733 | 1705 |  |
| MISSOURI | 1982 | 7434 | 9323 | 3 |
|  | 1983 | 6881 | 7286 |  |
|  | 1985 | 4362 | 3440 |  |
| ROCK CREEK | 1982 | 722 | 1374 | 2 |
|  | 1983 | 529 | 814 |  |
|  | 1985 | 419 | 641 |  |
| SMITH | 1982 | 234 | 367 | 2 |
|  | 1983 | 161 | 199 |  |
|  | 1985 | 110 | 87 |  |
| STILLWATER | 1982 | 1035 | 1914 | 2 |
|  | 1983 | 1201 | 1896 |  |
|  | 1985 | 632 | 1093 |  |
| SWAN | 1982 | 1332 | 2225 | 1 |
|  | 1983 | 1179 | 1776 |  |
|  | 1985 | 1157 | 1679 |  |
| UPPER | 1982 | 4765 | 8902 | 1 |
| YELLOWSTONE | 1983 | 4217 | 6252 |  |
|  | 1985 | 2981 | 4946 |  |
| MIDDLE | 1982 | 9110 | 16604 | 1 |
| YELLOWSTONE | 1983 | 8125 | 11840 |  |
|  | 1985 | 5066 | 7608 |  |

[^8]
### 4.5. Preliminary Analysis of the Hypothesized Flow Level, Fishing Pressure Relationship

Before proceeding with the travel cost modeling procedure for the entire sample, a simpler regression analysis was undertaken to identify the flow levelfishing pressure relationship. In this preliminary analysis the average number of days per year a river was fished by respondents was regressed against the mean monthly May-September flow for that river. In this analysis the average number of days per respondent was used rather than total days fished in an effort to reduce the effect of changes in the structure of fishing licensing regulations, and license prices. These changes directly affect the total number of licenses sold in a year but should have minimal impact on the number of days a fisherman, once licensed. fishes. Regressions were run for the entire aggregated sample (19 rivers, resident and non-resident), for resident and non-resident subgroups, and for each of the three control levels within each residency class. The following equation shows the results obtained for the entire sample regression.
LNPRESSURE $=1.874+1.22$ LNFLOW
T-stats $(21.55) \quad(4.64)$
Std.Err. $(.087) \quad(.26)$

$N=114 \quad R^{2}=.35$

Where:
LNPRESSURE = In(average number days fished per year per fisherman)


The above equation shows a significant relationship between flow and pressure.

RIVER
FLOW ${ }^{1}$ (HISTORICAL) STD.DEV. (HISTORICAL)

| BEAVERHEAD | 585 | $(428)$ | 143 | $(156)$ |
| :--- | :--- | :--- | :--- | :--- |
| BIG HOLE | 1205 | $(1172)$ | 335 | $(360)$ |
| BITTERROOT | 870 | $(936)$ | 220 | $(254)$ |
| BLACKFOOT | 1447 | $(1658)$ | 281 | $(466)$ |
| BOULDER | 539 | $(614)$ | 133 | $(140)$ |
| BIGHORN | 3637 | $(3914)$ | 1037 | $(1023)$ |
| UP.CLARK FK. | 3020 | $(3055)$ | 686 | $(890)$ |
| MD.CLARK FK. | 7252 | $(7598)$ | 1470 | $(2159)$ |
| UP.FLATHEAD | $10036(9758)$ | 114 | $(2052)$ |  |
| GALLATIN | 1239 | $(1069)$ | 324 | $(289)$ |
| KOOTENAI | $11238(12160)$ | 1200 | $(2398)$ |  |
| MADISON | 2024 | $(1759)$ | 216 | $(301)$ |
| MISSOURI | 5807 | $(5648)$ | 1405 | $(1318)$ |
| ROCK CREEK | 525 | $(616)$ | 124 | $(195)$ |
| SMITH | 168 | $(207)$ | 47 | $(59)$ |
| STILLWATER | 928 | $(968)$ | 257 | $(219)$ |
| SWAN | 1203 | $(1171)$ | 70 | $(261)$ |
| UP.YELLOW | 4182 | $(3769)$ | 630 | $(720)$ |
| MD.YELLOW | 7189 | $(7072)$ | 1780 | $(1634)$ |

${ }^{1}$ Flow $=$ mean monthly flow for the three study years
Std.Dev. $=$ standard deviation of flows for the three years
(Historical) $=$ historical flows and standard deviations

Table 4-9: Historical Flows and Standard Deviations


Figure 4-1: Sample Comparison: Big Hole Flows to Historical Averages

This regression was run for the three license years 82,83 , and 85 , eliminating the problematic data of the 84 license vear. The coefficient on flow has the expected positive sign and a T statistic which shows significance to the $99 \%$ level. The low $R^{2}$ of .35 is not entirely unexpected as flow level is only one of a myriad of determinants of fishing pressure. Additionally, no effort is made in this model to explain variation across sites other than as a function of flows. Therefore this unexplained variation dominates the variance of the sample leaving little ultimately explained in the model.

When regressions were run on less aggregated samples the most interesting result was the difference between the resident and the non-resident subsamples. The resident subsample again showed a relationship between flow and pressure of the expected sign and significant at the $95 \%$ level. The nonresident subsample, however, showed no significant relationship between flow and pressure. Additionally, regressions run for the flow-control subgroupings within each of the two residency classes showed no relationship between flow and pressure for the non-residents.

The absence of a strong relationship between flows and pressure for the non-resident population in this initial analysis is not entirely counter intuitive. There is obviously an information gap between the resident's knowledge of existing flow levels and fishing conditions and the knowledge of these things by the nonresident. Additionally, nonresidents often plan fishing trips months ahead of time and are less likely than the resident to cancel or modify these plans due to unfavorable flow conditions.

When the resident subgroup was broken down by river control level it was found that fishing pressure was more strongly related to flows on the rivers with minimal, or no, dam control. Significance at the $80 \%$ level was found for control levels 1 and 2, while no significant relationship was observed for the control level 3 subgroup.

In addition to estimating the flow-pressure relationship with a double log specification several alternative functional forms were tried. The model was estimated as a straight linear relationship with the results showing no significance of coefficients. Also the flow variable was entered in both a linear and a quadratic form with the result also showing no significant relationship.

# Chapter 5 <br> Model Estimation and Calculation of Benefits Associated with Alternative Flow Levels 

### 5.1. Travel Cost Model Estimation

In preparation for the modeling procedure the three year data set was aggregated according to individual origin/destination pairings within each study year. This aggregated file then required two additional transformations before it was suitable for the estimation of a regional travel cost model.

1. Since the rivers in the regional model are of varying sizes and their average flows have a range of an order of magnitude, it was necessary to normalize the flows across rivers. To accomplish this the flow variables for each river were indexed to the three year mean for that variable. The flow values in the travel cost model were therefore expressed as a percentage of the mean flow. The same indexing procedure was done for the number of trips and number of days. This accomplished, any observed variation in flows or visitation was no longer due to river size or popularity of the stream but solely to variation across the three years in the time series.
2. The three years in the data set were unequally sampled. To avoid any estimation bias due to unequal sample sizes the 1983 and 1985 samples were scaled to equal the 1982 sample size.

The specification of the regional travel cost model chosen for this study was the double log specification shown in equation (1). This is the same specification which Duffield, Loomis and Brooks (1987) found most successful.

## General Form of the Regional Travel Cost Model

$$
\text { (1) } \ln \left(\text { VISITS }_{i j} / P_{i j}\right)=\alpha-B_{i} \ln \left(\text { DIST }_{i j}\right)+B_{2} \ln \left(\text { FLOW }_{j}\right)
$$

Where:

VISITS $_{i j}=$ Trips from origin (i) to river ( j )
POP $_{i}=$ Population of origin (i)
DIST $_{i \mathrm{ij}}=$ Round trip distance from origin (i) to river ( j )
FLOW $_{j}=$ Mean monthily flow (cts) of river (j)
When the double log specification of the regional travel cost model was estimated for the three year, 19 river sample it returned the following results.

Double Log Regional Model
(2) $\operatorname{In}\left(\right.$ VISITS $_{i j} /$ POP $\left._{i}\right)=-1.64-1.98 \ln \left(\right.$ DIST $\left._{i j}\right)+.352 \ln \left(\right.$ YRFLOW $\left._{j}\right)$
$T$ stats (-8.3) (-68) (1.86)
Std.Err. (.198) (.03) (.189)

$$
N=1040 \quad R^{2}=.82
$$

Where: YRFLOW $_{j}=$ average monthly flow (cfs) for river ( j )
This model shows the distance variable very precisely estimated at the $99 \%$ level and showing the expected sign. Specifically, the price variable (round trip distance) shows a negative sign indicating decreased site visitation associated with increased travel distance. The coefficient on flow has a positive sign indicating that higher flows on specific rivers are associated with increased visitation and is significantly different than zero at $\alpha=.1$. The adjusted $R^{2}$ value of .82 shows that a large proportion of the variation in the dependent variable is explained by the model.

It is often beneficial to analyze the residuals of a regression procedure in an effort to determine the degree to which the model conforms to the assumptions of the ordinary least squares model. Additionally, plots of dependent and independent variables can provide clues as to the correctness of the functional form chosen for the model. Appendix $C$ shows these residual and variable plots and discusses their meaning and importance. Briefly, the dependent vs. independent variable plots showed the double log specification to be a suitable functional form for the estimated model. An analysis of the residuals showed that the assumption of normality was satisfied while that of homoscedasticity was violated by the price variable. How this problem of non-constant variance was dealt with is detailed in the following sections.

In addition to estimating the model with all 19 rivers pooled, the data also was broken down into river control level subgroups. Of the three models estimated (control=1 free flowing; control=2 slight dam control; control=3 heavy dam control) only the controi level 1 model identified a significant relationship between flows and visitation. The results of this estimation are shown below.

## Double Log Regional Model; Control Level 1 Subgroup

(3) $\begin{aligned} \operatorname{In}\left(\text { VISITS }_{\mathrm{ij}} /\right. & \left.\text { POP }_{\mathrm{i}}\right)= \\ \mathrm{T} \text { stats } & -2.71 \\ (-9) & -1.84 \ln \left(\text { DIST }_{\mathrm{ij}}\right)+\underset{(2.49)}{+826 \ln \left(\text { YRFLOW }_{\mathrm{i}}\right)} \\ \text { Std.Err. } & (.3) \\ & (.045)\end{aligned}$

$$
N=405 \quad R^{2}=80
$$

It can be seen from equation (3) that the coefficient on flow is both larger in absolute terms and more significant than in the 19 river model. This makes a
certain amount of sense. Dam control allows the flows of a river to be spread out more evenly over the entire year than would occur with normal runoff patterns. The rivers with no dam control, therefore, will experience higher high flows and lower low flows than those with dam control. In a drought year like 1985 the inability to reserve flows for the dry late summer periods could result in flows so low as to seriously impact fishing opportunities.

### 5.2. Alternative Specification of the Model

As an alternative to entering the flow variable as strictly a "shift" variable another specification of the model was tried. An application of the "varying parameter approach" (Vaughan and Russell 1982) was attempted which entered a composite variable (flow * distance) into the model. The varying parameter approach is a single equation method which allows a quality variable to affect both the intercept and the slope of the demand equation. The basic form of the varying parameter equation is shown in equation (3):

$$
\begin{equation*}
\left(V_{i j} / P_{i}\right)=B_{0}+B_{1} A_{i}+B_{2} T C_{i j}+B_{3}\left(T C_{i j}^{*} A_{j}\right)+\ldots \tag{3}
\end{equation*}
$$

Where:
$V_{i j}=$ visits from origin $i$ to site $j$.
$P_{i}=$ population of origin $i$.
$A_{j}=$ measure of site $j$ 's attractiveness.
$T C_{i j}=$ travel cost from origin $i$ to site $j$.
The measure of attractiveness used in the current model was the flow variable. The results from this regression analysis were identical to those from the simple yearly aggregation model shown in equation (2). In the varying parameter model the composite variable in(AVFLOW1 " DIST) was not significant in the
estimated regression model. Therefore, since YRFLOW was the only measure of flow which entered the regression the model was the same as reported in equation (2).

The varying parameter approach was also modeled entering the interaction variable (AVFLOW1*DIST) in a linear rather than a log form. Additionally, flow variables in both linear and quadratic form were entered into the model. The result of a stepwise regression of these formulations of the explanatory variables showed no significant relationship between flow and visitation. Since this formulation showed no difference in coefficients from equation (2), and thus offered no improvement in explanatory power over equation (2), the double log model specified in equation (2) was chosen for the remainder of the analysis in this study. The use of the functional form of equation (2) offers consistency with the techniques of Duffield, Loomis and Brooks (1987) and therefore will serve as a check of validity throughout the remaining analysis. An attempt was also made to explore a transformation on the flow variable. As an alternative to entering logged measures of average flow into the model, the variable AVFLOW1 was entered into regression (2) in a squared as well as a linear form. The regression showed, however, that the original measure of flow $\operatorname{In}(A V F L O W 1)$ was the only statistically significant explanatory flow variable.

### 5.3. Model Prediction

Very important for the calculation of consumer surplus is how well the estimated model predicts total trips for origin-destination pairings. The calculation of predicted trips is straightforward. The regression coefficients are plugged into the double log specification resulting in predicted trips per capita for each origindestination pair. Multiplying this figure by the population of the origin results in total predicted trips for each pairing. These figures can then be aggregated on a river level to arrive at predictions on a site by site basis.

Overall the regional model overpredicts trips by nearly $160 \%$. Actual trips were 974 while predicted trips were 2529. As was found by Duffield, Loomis and Brooks (1987), however, the vast majority of this overprediction could be accounted for by a small number of origin-destination pairings where the round trip distance is less than 30 miles. Table $5-1$ shows a matrix of trip predictions when the double log model is evaluated at a series of fixed average flow levels and average round trip distances. This table gives an impression of what the surface of the TRIPS, DISTANCE, FLOW plot is shaped like. It can be seen that mean origin/destination predictions are very large at the level of a 20 mile distance but fall off rapidly when distance is increased to 250 miles. In subsequent increases in distance the number of predicted trips diminishes much more slowly. This pattern is consistent with the finding of Duffield et al. (1987) that the double log specification of the model is heteroscedastic and tends to overpredict at very close origin/destination distances.ln a technical paper supplemental to their 1987 study Duffield (1987a) addressed this problem of
overprediction and showed that the vast majority of prediction error could be eliminated by use of a shifted distance transformation on the distance variable. That is, by the simple addition of a certain mileage ( 90 miles in his study) to the distance variable predictions could be brought into line with actual observations.

There is an intuitive economic rational for the application of the shifted distance transformation. The additional, or fixed, distance which is added to each observation distance can be thought of as a fixed cost associated with undertaking a fishing trip (Duffield 1987a). While this additional cost may not weigh heavily on the decision of a person coming 1000 miles to fish, it will severely impact those deciding whether to travel 10 or 20 miles to fish.

Bowes and Loomis (1980) suggested that another solution to the heteroscedasticity arising from unequally weighted population zones in the zonal model was the application of weighted least squares to the demand estimation. The weight they suggested was the square root of each origin zones population. For this analysis the shifted distance transformation suggested by Duffield (1987a) has been chosen over the application of weighted least squares. The shifted distance transformation was chosen for three reasons: (1) The transformation has the desired effect of correcting for the majority of overprediction error, (2) the transformation is computationally easy to apply, and (3) the economic rational behind the application of the shifted distance transformation is intuitively appealing.

In applying a shifted distance transformation to equation (2) it was found that a shift of only 10 miles was necessary to bring predictions and actual observations

| DISTANCE | 1500 | 2500 | FLOW | 4000 | 5000 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| DISTANCE | 1500 | 2500 |  | 4000 | 5000 |
| 20 mi | $\begin{aligned} & 23757^{1} \\ & (1308)^{2} \end{aligned}$ | $\begin{aligned} & 28456 \\ & (1566) \end{aligned}$ | $\begin{aligned} & 31219 \\ & (1719) \end{aligned}$ | $\begin{aligned} & 33596 \\ & (1849) \end{aligned}$ | $\begin{aligned} & 36352 \\ & (2001) \end{aligned}$ |
| 250 mi | $\begin{aligned} & 331 \\ & (18.2) \end{aligned}$ | $\begin{aligned} & 396 \\ & (21.8) \end{aligned}$ | $\begin{aligned} & 435 \\ & (23.9) \end{aligned}$ | $\begin{aligned} & 468 \\ & (25.7) \end{aligned}$ | $\begin{aligned} & 506 \\ & (27.8) \end{aligned}$ |
| 500 mi | $\begin{aligned} & 87 \\ & (4.8) \end{aligned}$ | $\begin{aligned} & 104.5 \\ & (5.7) \end{aligned}$ | $\begin{aligned} & 114.6 \\ & (6.3) \end{aligned}$ | $\begin{aligned} & 123 \\ & (6.8) \end{aligned}$ | $\begin{aligned} & 133.5 \\ & (7.3) \end{aligned}$ |
| 1000 mi | $\begin{aligned} & 22.5 \\ & (1.24) \end{aligned}$ | $\begin{aligned} & 27 \\ & (1.48) \end{aligned}$ | $\begin{aligned} & 29.6 \\ & (1.6) \end{aligned}$ | $\begin{aligned} & 31.9 \\ & (1.75) \end{aligned}$ | $\begin{aligned} & 34.5 \\ & (1.9) \end{aligned}$ |
| 2000 mi | $\begin{aligned} & 5.8 \\ & (.3) \end{aligned}$ | $\begin{aligned} & 6.92 \\ & (.38) \end{aligned}$ | $\begin{aligned} & 7.59 \\ & (.42) \end{aligned}$ | $\begin{aligned} & 8.17 \\ & (.45) \end{aligned}$ | $\begin{aligned} & 8.84 \\ & (.49) \end{aligned}$ |
| 3000 mi | $\begin{aligned} & 2.6 \\ & (.14) \end{aligned}$ | $\begin{aligned} & 3.11 \\ & (.17) \end{aligned}$ | $\begin{aligned} & 3.42 \\ & (.19) \end{aligned}$ | $\begin{aligned} & 3.67 \\ & (.20) \end{aligned}$ | $\begin{aligned} & 3.98 \\ & (.22) \end{aligned}$ |

[^9]Table 5-1: Sample Predictions from Double Log Model
of trips into line. This is significantly less than the 90 mile optimal shift which Duffield found in his study. Table 5-2 shows actual trips by river, trips predicted by equation (2) and equation (2) predictions after the application of the 10 mile shifted distance transformation.

### 5.4. Site-Specific Regression Results and Predictions

When estimating benefits associated with existing sites it has been found that single-site models often provide more accurate results on a river specific basis than larger multi-site regional models (Sorg and Loomis 1985). In addition to the entire sample the double log, year level specification was run for each of the 19 study rivers on an individual site basis. Of these 19 rivers only 4 returned results which showed significance of the flow variable. The coefficients and $T$ statistics for these 4 regressions are reported in Table 5-3. All of the coefficients in these equations showed the expected signs for distance and flow. Additionally, the coefficients on distance and flow, with just one exception, were estimated at the $90 \%$ confidence level or higher.

The predictive power of the 4 single site models was far superior to that of the larger multi-site regional model. Table 5-4 shows actual and predicted trips for the site-specific travel cost models. For this 4 river subsample the overprediction error was reduced from $650 \%$ for the regional model to $63 \%$ for the single site models. No attempt was made to apply a shifted distance transformation to the site-specific TCM's.

| CODE | RIVER | ACTUAL <br> TRIPS | PREDICTED <br> TRIPS | PREDICTED <br> TRIPS(SHIFT) |
| :--- | :--- | :--- | :--- | :--- |
|  |  |  |  |  |
| 80 | BEAVERHEAD | 44 | 38 | 28 |
| 81 | BIG HOLE | 89 | 54 | 38 |
| 82 | BITTERROOT | 42 | 292 | 118 |
| 83 | BLACKFOOT | 59 | 56 | 45 |
| 84 | BOULDER | 48 | 77 | 28 |
| 85 | BIGHORN | 53 | 23 | 22 |
| 86 | UP.CLARK FK. | 44 | 105 | 65 |
| 87 | MD.CLARK FK. | 33 | 70 | 43 |
| 89 | UP.FLATHEAD | 35 | 290 | 93 |
| 90 | GALLATIN | 69 | 73 | 51 |
| 91 | KOOTENAI | 32 | 40 | 31 |
| 92 | MADISON | 88 | 38 | 35 |
| 93 | MISSOURI | 67 | 36 | 32 |
| 94 | ROCK CREEK | 40 | 42 | 35 |
| 95 | SMITH | 45 | 33 | 29 |
| 96 | STILLWATER | 54 | 37 | 26 |
| 97 | SWAN | 32 | 43 | 34 |
| 98 | UP.YELLOW | 61 | 665 | 96 |
| 99 | MD.YELLOW | 36 | 516 | 180 |

${ }^{1}$ PREDICTED TRIPS USING A 10 MILE SHIFT ON THE DISTANCE VARIABLE
(Full sample predictions from equation (2))

Table 5-2: Trip Predictions from Full Sample Model

| RIVER | INTERCEPT | DIST | FLOW | N |
| :---: | :---: | :---: | :---: | :---: |
| BITTERROOT | $\begin{aligned} & -3.91 \\ & (-5) \end{aligned}$ | $\begin{aligned} & -1.765 \\ & (-15.7) \end{aligned}$ | $\begin{aligned} & .886 \text { AVFLOW1 }{ }^{1} \\ & (1.48)^{85 \%} \end{aligned}$ | 37 |
| UP.CLARK FK. | $\begin{aligned} & -2.52 \\ & (-3.7) \end{aligned}$ | $\begin{aligned} & -1.905 \\ & (-17.4) \end{aligned}$ | $\begin{aligned} & .856 \text { AVFLOW2 }{ }^{2} \\ & (1.73)^{90 \%} \end{aligned}$ | 43 |
| UP.FLATHEAD | $\begin{aligned} & -3.874 \\ & (-3.95) \end{aligned}$ | $\begin{aligned} & -1.747 \\ & (-13.3) \end{aligned}$ | 8.14 AVFLOW1 <br> (2.5) | 37 |
| UP.YELLOWST. | $\begin{aligned} & -4.46 \\ & (-6) \end{aligned}$ | $\begin{aligned} & -1.60 \\ & (-14.9) \end{aligned}$ | 3.24 YRFLOW ${ }^{3}$ <br> (2.65) | 59 |

Note: All coefficients are significant at $95 \%$ level uniess otherwise indicated
(Equation (2) results for individual rivers)
${ }^{1}$ Average monthly flow for May-September
${ }^{2}$ Average monthly flow for July-September
${ }^{3}$ Average monthly flow for entire year

Table 5-3: Regression Results for River Specific Equations

| CODE | RIVER | ACTUAL <br> TRIPS | PREDICTED <br> TRIPS |
| :--- | :--- | :--- | :--- |
| 82 | BITTERROOT | 42 | 59 |
| 86 | UP.CLARK FK. | 43 | 59 |
| 89 | UP.FLATHEAD | 34 | 87 |
| 98 | UP.YELLOWST. | 61 | 89 |

(Predictions from equation (2) coefficients; individual rivers)

Table 5-4: Site-specific predictions for the 4 River Subsample

### 5.5. Calculation of Benefits Associated with Actual Flow Levels

Benefits from fishing on the rivers included in this study were calculated as the area under the site demand curves. This was done through direct integration of the first stage demand curve between the current origin to destination distance and the sum of the maximum observed distance to the destination and the current distance. The choice of this upper limit of integration was made based on Duffield's (1987b) technical report exploring the sensitivity of benefit estimates to alternative specifications of the upper limit of integration. Equation (5) shows the method used to calculate consumer surplus estimates:
(5) $\quad C S=\left(e^{b_{0} / b_{1}+1}\right) \cdot F_{i}^{b_{2}} \cdot\left(D_{u i}^{b_{1}+1}-D_{i j}^{b_{1}+1}\right)$

## Where:

```
\(F_{j}=\) Flow variable for river (j)
\(D_{\text {ui }}=\) Maximum observed travel distance to river ( j )
\(\mathrm{D}_{\mathrm{ij}}=\) Current distance from origin (i) to river ( j )
\(b_{0}=\) The estimated intercept
\(b_{1}=\) The estimated coefficient on distance
\(b_{2}=\) The estimated coefficient on flow
```

The consumer surplus variables shown in Table 5-5 are the regional model values from the results of regression (2) with the 10 mile shifted distance transformation. The model shows significant variation in values across sites and comparison with the results of Duffield, Loomis and Brooks (1987) shows slightly larger consumer surplus estimates. Table 5-5 shows a mean consumer surplus per day of $\$ 166.32$ for the 19 study rivers. An estimate of total benefits associated with fishing in each river is calculated for 1985 by multiplying angler pressure estimates for that year by the consumer surplus per day calculated for each river. These river-specific benefits are shown in Table 5-7.

Consumer surplus values were also calculated for the site-specific regression equations shown in Table 5-3. The mean consumer surplus per day for this 4 river subsample was $\$ 66.43$. The benefit estimates for these 13 rivers are shown in Table 5-6. As was done for the regional model, the consumer surplus estimates derived from the river-specific models were multiplied by the angler pressure estimate for 1985 to arrive at total site benefits for that year. These total benefits are shown in Table 5-8.

### 5.6. Hypothetical Flow Variation and its effect on Total Site Benefits

So far in this study we have: (1) Specified a travel cost model which identifies a robust relationship between average river flows and visitation to those rivers, (2) calculated, based on the specified model, per day net values for recreational fishing on these rivers, and (3) estimated total site benefits for 1985 based on the calculated per day values and DFWP estimates of fishing pressure for that year. What remains in this analysis is to observe how these total site benefits change with hypothetical changes in the magnitude of the flow variable.

This portion of the study, through invoking the hypothetical, treads gingerly into the world of speculative calculations. The first question which must be answered is what the realistic range of hypothetical flow variation should be. A cursory examination of the mean flow variables reported in Table 3-8 indicate that variations of + or $-50 \%$ are not out of the ordinary. This is particularly true when viewing the summer months of May through September, when the vast majority of fishing on Montana's rivers takes place. Therefore, in estimating the effects of

|  |  | CONS.SURPLUS <br> PER DAY | AVE.ZONES <br> PER YEAR |
| :--- | :--- | :--- | :--- |
|  | RIVER |  |  |
| 80 | BEAVERHEAD | 259.51 | 19 |
| 81 | BIG HOLE | 237.84 | 33 |
| 82 | BITTERROOT | 56.34 | 13 |
| 83 | BLACKFOOT | 161.85 | 20 |
| 84 | BOULDER | 235.30 | 17 |
| 85 | BIGHORN | 336.54 | 20 |
| 86 | UP.CLARK FK. | 86.02 | 14 |
| 87 | MD.CLARK FK. | 86.20 | 12 |
| 89 | UP.FLATHEAD | 51.90 | 12 |
| 90 | GALLATIN | 172.14 | 26 |
| 91 | KOOTENAI | 166.07 | 12 |
| 92 | MADISON | 267.63 | 40 |
| 93 | MISSOURI | 203.94 | 22 |
| 94 | ROCK CREEK | 190.02 | 17 |
| 95 | SMITH | 191.70 | 15 |
| 96 | STILLWATER | 206.18 | 17 |
| 97 | SWAN | 148.10 | 11 |
| 98 | UP.YELLOWSTONE | 80.67 | 20 |
| 99 | MD.YELLOWSTONE | 22.22 | 11 |

(Values from equation (2) coefficients)

Table 5-5: Consumer Surplus per day: Regional Mode!

CONS.SURPLUS AVE. ZONES

| CODE | RIVER | PER DAY | PER |
| :--- | :--- | :--- | :--- |
| 82 | BITTERROOT | 63.54 | 13 |
| 86 | UP.CLARK FK. | 67.03 | 14 |
| 89 | UP.FLATHEAD | 45.30 | 12 |
| 98 | UP.YELLOWSTONE | 89.88 | 20 |

SITE AVERAGE

\$66.43
(Values from table 4-2 coefficients)

Table 5-6: Consumer Surplus per day: Site-specific Models

'SITE VALUES ARE IN 1000 'S
(Values from equation (2) coefficients)

Table 5-7: Total Site Benefits: Regional Model

| CODE | RIVER | VALUE <br> PER DAY | PRESSURE <br> 1985 | TOTAL <br> VALUE 1985 ${ }^{1}$ |
| :--- | :--- | :--- | :--- | :--- |
| 82 | BITTERROOT | 63.54 | 56024 | 3560 |
| 86 | UP.CLARK FK. | 96.63 | 17,578 | 1697 |
| 89 | UP.FLATHEAD | 70.23 | 15,262 | 1072 |
| 98 | UP.YELLOW | 130.10 | 52,016 | 6765 |

Table 5-8: Total Site Benefits: Site-specific Models
hypothetical changes in the "shifter" variable flow four alternative changes were analyzed: $+50 \%,+25 \%,-25 \%$, and $-50 \%$. The historical flow means and standard deviations shown in table 4-9 give an indication of how flow variation of $+50 \%$ to $\mathbf{- 5 0 \%}$ relate to historical changes in flow. The mean historical flow for the 19 rivers is 3346 cfs and the mean standard deviation is 784 cfs or $24 \%$ of the mean flow. Roughly speaking, therefore, a flow change of + or $-25 \%$ is equivalent to a 70 year flood or drought and a change of + or $-50 \%$ is equivalent to a 100 year flood or drought.

The Montana Department of Fish Wildlife and Parks has made an effort to define minimum acceptable instream flows on many rivers in Montana. Table 5-9 shows a comparison of the hypothetical flow reductions used in this analysis and the minimum acceptable flows set by the DFWP. Of the 19 rivers, DFWP minimum flow data was available for 15 . In general, a hypothetical reduction of $25 \%$, with one exception, did not cause flows to fall below the DFWP minimums. A reduction of $50 \%$, however, with three exceptions, caused flows to violate this level. These comparisons suggest that the reductions of $25 \%$ should be acceptable for modeling visitation changes due to flow reductions which are not yet destructive of fish populations. It should be pointed out that all reported flows are year averages, and due to the large variability of monthiy flows, and to the fact that 1985 was a drought year a flow reduction of $25 \%$ on a specific river in a specific month might indeed have disastrous effects on fish populations. These effects can not be captured by this simple formulation of the model and, therefore, marginal values of water flows could be significantly understated in such cases.

| RIVER | STUDY <br> AVE. | STUDY <br> $-25 \%$ | STUDY <br> $-50 \%$ | DFWP <br> MIN. |
| :--- | :--- | :--- | :--- | :--- |
| BEAVERHEAD | 585 | 439 | 292 | 200 |
| BIG HOLE | 1205 | 904 | 603 | 800 |
| BITTERROOT | 876 | 657 | 438 | MISSING |
| BLACKFOOT | 1447 | 1085 | 724 | 939 |
| BOULDER | 539 | 404 | 270 | 269 |
| BIGHORN | 3637 | 2728 | 1818 | 3422 |
| UP.CLARK FK | 3020 | 2265 | 1510 | MISSING |
| MD.CLARK FK | 7252 | 5939 | 3626 | MISSING |
| UP.FLATHEAD | 10036 | 7527 | 5018 | 4674 |
| GALLATIN | 1239 | 929 | 619 | 879 |
| KOOTENAI | 11238 | 8429 | 5619 | 4000 |
| MADISON | 2024 | 1518 | 1012 | 1280 |
| MISSOURI | 5807 | 4355 | 2903 | 3000 |
| ROCK CK. | 525 | 394 | 263 | 343 |
| SMITH | 168 | 126 | 84 | 100 |
| STILLWATER | 928 | 696 | 464 | 524 |
| SWAN | 1203 | 902 | 602 | MISSING |
| UP.YELLOW | 4182 | 3137 | 2091 | 2596 |
| MD.YELLOW | 7189 |  |  |  |

Table 5-9: Comparison of Hypothetical Flows and DFWP Minimum Flows

The effect of hypothetical changes in flow values is reflected in changes in the number of fishing trips predicted by the model. As in the preceding analysis, these changes were introduced into two alternative specifications of the model. The effects of the hypothetical changes on trip predictions for the single equation regional model are shown in table 5-10. Additionally, the changes in trip predictions for the site-specific 4 river subsample models are detailed in tables 5-12 and 5-13.

Table 5-10 shows changes in total valuation resulting from each of the hypothetical shifts in flow, for the single equation regional model. The results for this model are presented in an aggregated rather than river-specific form since the model coefficients are the same for each river in the sample.

Previous attempts to measure the acre/foot value of water for recreational uses have shown widely diverging estimates. Daubert and Young (1981) used a contingent valuation approach to arrive at a marginal value of September flows on the Poudre river of $\$ 7.68$ acre/foot. Narayanan (1986), on the other hand, calculated a marginal value of flow for recreation of 42 cents acre/foot for the Blacksmith Fork river. The results of this study help to bridge the gap between these two estimates.

Table 5-11 shows the calculated acre/foot valuation associated with a $25 \%$ decrease in flow levels. The aggregated 19 river regional model, shown first, returned a marginal valuation associated with this shift of $\$ 1.03$ acre/foot. The 4 rivers in the site-specific models showed values ranging from a low of 22 cents an acre/foot on the Upper Flathead to $\$ 6.54$ acre/foot on the Bitterroot river. A few preliminary comments on these figures are in order.

A brief examination of the correspondence of total discharge to acre/foot value shown in table 5-11 shows a degree of correlation. There may be several reasons for this negative correlation. One explanation might be that smaller rivers in general provide higher quality fisheries than large volume rivers. This explanation, however, probably portrays far too simple of a relationship between river size and quality of fishing. In many cases larger rivers have lower quality fisheries than smaller rivers because their volume and length attracted the attention and deleterious meddling of man at an earlier date. This meddling may take the form of dams or pollution (thermal, and chemical) both of which have a significant effect on the numbers and species of fish which inhabit the waters. Another possible explanation for the negative correlation between river size and the marginal value of flows is computational in origin. It may be that the model is returning a relatively stable total value change per river associated with the hypothetical shifts but this value change is being swamped by the huge differences in volume change across rivers implied by a $25 \%$ decrease in flows. Whatever the reason for this negative correlation it seems to be the case that for the study rivers there is a diminishing marginal value associated with the larger discharge rivers. It must be pointed out that the values shown in Table 5-11 are average acre/foot values for a $25 \%$ reduction in flows. There are undoubtedly specific periods on many of the rivers (eg. the high use, low flow periods of late summer) when these values would represent a severe understatement of true marginal values. The interpretation of the values shown in Table 5-10 and Table 5-11 present certain challenges to the analyst. At first blush the estimate of a
$\mathbf{2 2 . 9} \%$ reduction in net economic value of fishing being associated with a $50 \%$ reduction in flows and marginal values of flows between 22 cents and $\$ 6.54$ per acre/foot seem minimal. Two possible explanations for these values follow.

Figures 3-1 and 3-2 showed a possible form for the flow level-fishing value relationship. Very important to any conclusions we might draw from the effects of hypothetical flow shifts is that we first understand which portion of the flow-value relationship we are modeling.

For the 19 river regional model a $25 \%$ decrease in flows showed a marginal value associated with this decrease of $\$ 1.03$ acre/foot. A calculation of the average value of flows for 1985 shows a value of $\$ 2.83$ acre/foot. Figure 5-2 shows that the relationship of marginal to average values suggests the region of the total value modeled by this data set corresponds to the relatively flat portion of the curve near its maximum. Specifically, since marginal value is less than average value and both are positive, we are modeling between the horizon point and the maximum point in figure 5-1. Since we have modeled a small portion of the total value curve near its maximum extrapolation outside of that portion is risky and may provide gross under or overstatements of the marginal values assigned to differing flows.

Another problem with the extrapolation done in this analysis lies in the inability of the researcher to identify other variables which begin to affect visitation at abnormally high or low flow levels. A possible interpretation of the results shown in table 5-10 is that over a certain range visitation to rivers is only minimally sensitive to changes in flows. This range would correspond to " $X$ " in
figure 5-3. There most probably are however flow ranges within which the marginal value of additional flow levels to fishermen is very high. These ranges could be quite narrow, or even jump discontinuities in the functional relationship, and it is at these levels where the value associated with fishing is most sensitive to changes in flows. Points " $A$ "," $B$ " and " $C$ " on figure 5-3 correspond to these jump discontinuities. Two possible examples of these discontinuities, or threshold levels, follow:

1. If a significant portion of the fishing pressure on a stream is made possible by access gained through floating rafts or driftboats down the river then limitations which flows might place on floaters become very important. There may, for example, be a flow level below which it is very difficult or impossible for float craft to maneuver the river. The value of additional flows at or just below this threshold level then becomes quite high since flows below the threshold exclude an entire class of fisherman from the resource.
2. Another threshold, the violation of which has more disturbing consequences, is that level below which fish kill occurs due to unsatisfactory habitat conditions. The results of flows falling below such a threshold would impact not only the present years recreational value but also that of the future years until the fish population had fully recovered. It is easy to see that the marginal value of additional water flows at such a "biological" threshold level could be very substantial.

The overall effect of the exclusion of such important explanatory factors as threshold levels from this analysis argues for caution in interpreting the values shown in tables 5-10 through 5-11. The positive coefficient estimated on the flow variable indicates that overall there are positive economic values associated with increased flow levels. The interpretation of the range over which the coefficient on flow is valid is a more difficult problem. To attempt to extend prediction of the model beyond these threshold levels would be to vastly over or understate the
marginal value of flows beyond those threshold flows. It is entirely possible that the observations in the data set used in this study did not, or only minimally, violated these types of levels. Therefore, the coefficient estimated on flow shows a minimally sensitive, although highly significant relationship of visitation to flows. Just as there are threshold levels associated with low flows, logic suggests there are also those associated with high flood levels. There probably is, therefore, a very limited range within which we can accurately predict trips using this model without the inclusion of additional explanatory variables.

Tables 5-12 and 5-13 show the changes in valuation associated with hypothetical flow shifts of $-25 \%$ and $-50 \%$ being applied to the site-specific models. For the 4 site specific models two rivers, the Bitterroot and Upper Clark Fork, show variations in pressure due to flow variations comparable to the 19 river average of the regional model. The other two rivers, the Upper Flathead and the Upper Yellowstone, show a much higher sensitivity to flow variation with a 100 year drought cycle being associated with an 88 to $99 \%$ decrease in fishing values. The same caveats associated with the regional model are true of this single site specification. Consequently, caution in the interpretation of these values is urged.

### 5.7. Comparison of Regional and Single-site Models

Sorg and Loomis (1985) urged the use of regional estimation techniques in the application of the travel cost model where possible. They note, however, that single site models may provide more accurate estimates of existing site benefits. For this reason two alternative specifications of the travel cost model were run on the three year DFWP data set.

## AGGREGATE EFFECTS OF FLOW VARIATION

|  | +50\% | +25\% | -25\% | -50\% |
| :---: | :---: | :---: | :---: | :---: |
| \% CHANGE IN |  |  |  |  |
| PRESSURE | +15.07\% | +7.87\% | -9.92\% | -22.98\% |
| 1985 BASE |  |  |  |  |
| PRESS | 714,906 | 714,906 | 714,906 | 714,906 |
| CHANGEIN |  |  |  |  |
| PRESSURE | +107,736 | +56,263 | -70,919 | -164,285 |
| VALUE (DAY) | 166.32 | 166.32 | 166.32 | 166.32 |
| CHANGE IN |  |  |  |  |
| TOTAL VALUE | +17,981,652 | +9,357,662 | -11.795.248 | -27,323,881 |

(Results calculated using coefficients from equation (2))
Table 5-10: Effects of Hypothetical Flow Variation: Regional Model

REGIONAL 19 RIVER MODEL

| RIVER | TOT. DISCHARGE | VALUE ACRE/FOOT |
| :--- | :--- | :--- |
| ENTIRE SAMPLE | $45,727,381$ ACRE/FEET | $\$ 1.03$ ACRE/FOOT |

4 RIVER SITE-SPECIFIC MODELS
RIVER TOT. DISCHARGE VALUE ACRE/FOOT

| BITTERROOT | 479,080 | $\$ 6.54$ (ACRE/FOOT) |
| :--- | :--- | :--- |
| UP. CLARK FK. | $1,700,970$ | 60.9 CENTS |
| UP.FLATHEAD | $7,251,400$ | 22 CENTS |
| UP.YELLOWSTONE | $2,163,910$ | $\$ 5.09$ |

Table 5-11: Acre/Foot Values Associated with a $25 \%$ Decrease in Flows


Figure 5-1: Range of Model Validity in Flow-Value Relationship


Figure 5-2: The Average Value - Marginal Value Relationship


Figure 5-3: Hypothesized Flow Fishing Pressure Relationship

HYPOTHETICAL 25\% DECREASE IN FLOWS

| RIVER | \% CHANGE | CHANGE <br> IN DAYS | VALUE PER DAY | TOT. \$ CHANGE |
| :---: | :---: | :---: | :---: | :---: |
| BITTERROOT | -22\% | 12325 | 63.54 | 723,130 |
| UP.CLARK FK. | -21\% | 3867 | 86.02 | 332,639 |
| UP.FLATHEAD | -58\% | 8852 | 51.90 | 459,419 |
| UP.YELLOW | -59\% | 30085 | 80.67 | 2,426,957 |
| TOT. CHANGE |  |  |  | \$-4,002,145 |

Table 5-12: Effects of 25\% Decrease in Flow Levels:Site-Level Models

## HYPOTHETICAL 50\% DECREASE IN FLOWS

| RIVER | \% CHANGE | CHANGE <br> IN DAYS | VALUE <br> PER DAY | TOT. $\$$ <br> CHANGE |
| :--- | :--- | :--- | :--- | :--- |
| BITTERROOT | $-38 \%$ | 21289 | 56.34 | $1,199,429$ |
| UP.CLARK FK. | $-44 \%$ | 7746 | 86.02 | 666,331 |
| UP.FLATHEAD | $-99 \%$ | 15109 | 51.90 | 781,472 |
| UP.YELLOW | $-88 \%$ | 45767 | 86.67 | $3,692,098$ |
|  |  | TOT. CHANGE | $-\$ 6,339,330$ |  |

Table 5-13: Effects of 50\% Decrease in Flow Levels:Site-Level Models

The regional model specification was run for the 19 river sample. To improve its predictive powers a shifted distance transformation of 10 miles was applied to the distance (price) variable, as suggested by Duffield (1987a). This model showed a $\$ 166.32$ value for the per day consumer surplus associated with fishing on these rivers.

The site-specific models were run for each of the 19 study rivers, and 4 of them returned significant relationships for the distance and flow variables. The predictive power of these individual models was significantly better than the regional model so no shifted distance transformation was applied in the calculation of benefits. The single-site river subsample returned a mean consumer surplus per day of $\$ 66.43$. This value is comparable to the consumer surplus in the Regional model of $\$ 68.73$ for this 4 river subsample.

In the analysis of changes in prediction associated with alternative flow levels the changes returned by the 4 site specific models fell to both sides of the results from the regional model indicating that the regional model mav be returning a valid average value the 19 river model.

# Chapter 6 <br> <br> Conclusions and Future Research 

 <br> <br> Conclusions and Future Research}

### 6.1. Overall Conclusions

The major goal of this study was the determination of whether a systematic relationship between instream flow levels and recreational fishing exists on trout streams in Montana. Two modeling techniques were used to explore this relationship, and both showed a significant positive relationship between average flow levels and fishing pressure. This finding was expected and makes intuitive sense. In addition to modeling fishing pressure as a continuous function of flow levels this study attempted to predict changes in net willingness to pay associated with various hypothetical flow levels. The results from this part of the analysis were somewhat surprising. The implicit values for instream flows derived from these hypothetical shifts were lower than expected.

The estimation of unexpectedly low values associated with instream flows is most likely more indicative of the shortcomings of the model used than actual tow values associated with the flows. It appears that this model only effectively modeled a small portion of the total value curve associated with alternative flow levels. The portion modeled was probably relatively flat with low value changes being associated with changes in flows. Attempts to predict value changes associated with flows outside of this range would underestimate the marginal
value of these changes. Another possible explanation for these low values is that the functional relationship between values and flows is not as expected. The models as specified show visitation as a continuous function of flows. More realistic than this specification, however, might be a relationship showing "jump" discontinuities. These discontinuities in the relationship would represent threshold flow levels below which entire classes of fishermen would be excluded (eg. flows too low to allow float boats, or flows too low to support fish populations). The models included in this study fail to incorporate the effects of these threshold levels, and therefore most likely underpredict the effects of flow levels which violate these thresholds.

The most important result of this modeling is that a highly significant relationship, of the expected sign, between flows and fishing pressure was found in the data base. This relationship was expected and should encourage additional research into this important area.

### 6.2. Suggested Directions for Future Research

There are a number of improvements which could be made upon the procedures undertaken in this study. In general, these could be accomplished by enlargement of the time series used and incorporating more complex specifications of the flow variable.

The DFWP data base used in this analysis dictated that only three years be used in the time series. Normaliy, a three period time series would provide substantial problems in estimating a significant relationship of fishing with flow. It
was fortunate that a large amount of flow fluctuation occurred in the three years included in this study. Any expansion of the number of periods in the time series would help in a more accurate specification of the flow variable.

As was mentioned above, the flow variable is most likely more complex in its relationship to fishing pressure than was specified in this study. Any attempts to try to expand specification of the flow variable to include these complexities would doubtless lead to a more valid specification. With the increasing importance of water to uses, both marketed and non-marketed, efforts in this area should be strongly encouraged.

# Appendix A <br> Data Manipulation and Transformation Processes 

## A.1. Introduction

Before the actual analysis of DFWP data was possible in the Net Economic Value of Instream Flows study, it was necessary to make numerous transformations of that data. This appendix is a documentation of those transformation processes. The analysis of the DFWP monthly pressure survey data was performed on the Digital Equipment Corporation VAX Mainframe computer at the University of Montana. The statistical package used was SPSSk.

## A.2. Merging zip code files with pressure data

The raw pressure survey data files were divided into three study years. These study years were the fishing license years $82-83,83-84$, and $85-86$. The data files varied between years in the information they contained, but all three had in common the following variables: ID\#, WAVE\#, RESIDENCY, REGION, DRAINAGE, WATER-CODE, WATER-TYPE, DAYS-FISHED. A second set of files was prepared by the DFWP which contained: ID\#, WAVE\#, RESIDENCY, ZIP-CODE. This data was coded off of the original pressure survey forms when it was found that zip codes were not originally entered in the data records.

The transformations of these raw files into SPSSx system files, which are efficiently read and easily usable, entailed adding a few SPSSx commands to each
data file and running each as an SPSSx program. Once translated into SPSSX System Files these data files could be easily accessed from any SPSSx program using a FILE HANDLE command designating the desired System File and the GET FILE command.

With all data files now usable as SPSSx System Files the next step was to merge the zip code files with the main pressure files. In order to merge two files using SPSSx two conditions must be met. First, there must be a key variable or set of variables present in both files to direct the merging process. In our analysis a combination of the variables ID\#, WAVE\#, and RESIDENCY was used as the key to uniquely identify a survey respondent. Additionally, the files must be sorted in ascending order on these key variables. If duplicate keys are found in one file during the merge process the merge will be performed upon the first occurrence of the key only. A warning will then be generated by SPSSx to alert the investigator. This was a common occurrence with the files which we used since many survey respondents (each identified with a unique key) made more than one fishing trip in a given month. Program segment 1 shows the commands necessary to sort and merge the yearly pressure files with the zip code files.

## Program Segment 1

```
file handle condenc/name="condensed.dta"
file handle zip/name="zip82-system.dta"
file handle match/name="merged82dta"
file handle pres82/name="pres82-systemdta"
file handle zip2/name="zip2-sort-sys.dta"
file handle pres822/name="sys82-sort.dta"
get file=zip
sort cases by id wave resident
xsave outfile=zip2
execute
get file=pres82
```

```
sort cases by id wave resident
xsave outfile=pres822
execute
get file=zip2
execute
get file=pres822
match files file=zip2/ file=pres822/ by id wave resident
xsave outfile=merge82
execute
```

The above segment only sorts and merges the files for the 82 study year. This process was repeated for each of the years needing the zip code merge. Once the merging was accomplished it was necessary to fill in zip codes for all cases with multiple trips within a unique key. Program segment 2 accomplished this.

## Program Segment 2

```
file handle fill82/name="merged-filled82.dta"
file handle merge82/name="merge82.spssxsav"
get file=merge82
if missing(zip) zip=0
do if (zip=0 and id=lag(id,1) and wave=lag(wave,1)
    and resident=lag(resident,1))
compute zip=lag(zip,1)
end if
execute
xsave outfile=fill82
execute
```

Again, this procedure was performed upon each of the merged years.

## A.3. State, County, and Unique Water Transformations

The variables which were necessary for TCM analysis, as applied in this study, included STATE, COUNTY (within Montana), and UNIT2 (unique water code). STATE and COUNTY variables were obtained by performing transformations on the zip code variable, while the unique water code (UNIT2) was transformed from region, drainage, and water code data. Program segment 3 performed these transformations along with attaching a YEAR variable to each record.

## Program Segment 3

file handle input name="merged-filled82.dta" file handle unique 82 name="unique-sort.dta" get file=input
recode zip (1000 thru 2799=20) (2800 thru 3899=28) ( 3900 thru $4999=18$ ) ( 5000 thru 5999=44) ( 6000 thru $6999=6$ ) ( 7000 thru $8999=29$ ) (10000 thru 14999=31) (15000 thru 19699=37) (19700 thru 19999=7) (20000 thru 21999=19) (22000 thru 24699=45) (24700 thru 26999=47) (27000 thru 28999=32) (29000 thru 29999=39) ( 30000 thru $31999=10$ ) ( 32000 thru $33999=9$ ) ( 35000 thru $36999=1$ ) ( 37000 thru $38599=41$ ) (28600 thru 39999=23) (40000 thru 42999=16) (43000 thru 45999=34) (46000 thru 47999=13) (48000 thru 49999=21) (50000 thru 52999=14) ( 53000 thru $54999=48$ ) ( 55000 thru $56999=22$ ) (57000 thru 57799=40) (58000 thru 58999=33) (59000 thru 59999=25) (60000 thru 62999=12) ( 63000 thru $65999=24$ ) ( 66000 thru 67999=15) ( 68000 thru $69999=26$ ) ( 70000 thru 71599=17) ( 71600 thru $72999=3$ ) ( 73000 thru $74999=35$ ) (75000 thru $79999=42$ ) ( 80000 thru $81999=5$ ) ( 82000 thru $83199=49$ ) ( 83200 thru $83999=11$ ) $(84000$ thru $84999=43)(85000$ thru $86999=2)$ ( 87000 thru $87299=30$ ) ( 87300 thru $87399=5$ ) ( 87400 thru $87599=30$ ) ( 87600 thru $87699=5$ ) ( 87700 thru $87899=30$ ) ( 87900 thru $87999=5$ ) ( 88000 thru $88199=30$ ) ( 88200 thru 88299=5) ( 88300 thru $88499=30$ ) ( 88500 thru $88599=5$ ) ( 88600 thru $88799=30$ ) ( 88800 thru 88899=5) ( 88900 thru $88999=30$ ) ( 89000 thru $89999=27$ ) ( 90000 thru $96699=4$ ) ( 97000 thru $97999=36$ ) ( 98000 thru $99499=46$ ) ( 99500 thru $99999=50$ ) into state
do if any (zip, 59003,59004,59005,59012,59039,59040,59043,59083, 59312.59323.59327.59328.59329.59333.59334.59346.593471 compute county $=87$
else if any (zip, 59001.59019.59028.59057,59061,59063.59067.59069) compute county $=95$
else if any Izip. $59002,59006,59015,59024,59037,59044,59045.59064$, 59079,59088,59101,59102,59103)
compute county $=111$
else if any $\{$ zip, $59007,59008,59013,59026,59029,59051,59068,59070$, 59071,59080.59041)
compute county=9
else if any (zip,59010,59060,59076,59038) compute county $=103$
else if any (zip,59011,59033,59052,59055,59056)
compute county $=97$
else if any (zip,59014.59023.59046.59074)

```
    compute county=37
else if any (zip,59016,59022,59025,59031,59034,59035,59050,59066,
        59075,59089)
    compute county=3
else if any (zip,59017.59087.59084)
    compute county=69
else if any (zip,59020,59021,59018,59027,59030,59047,59048,49049,
        59065.59081,59082,59086.59874)
    compute county=67
else if any (zip,59032,59418,58423,59424,59429,59430,59438,59441.
        59445,5945 1,59457,59458,59459,59464,59471,59481,59489)
    compute county=27
else if any (zip,59036,59078,59085,59453)
    compute county=107
else if any (zip,59053,59642,59645,59646)
    compute county=59
else if any (zip,59054,59059,59072,59073)
    compute county=65
else if any (zip,59058,59077,59318,59322,59337)
    compute county=33
else if any (zip,59062,59314,59317,59321,59325,59343,59345,59348,
        59338,59854)
    compute county=75
else if any (zip,59173,59411,59415,59417,59427,59428,59434)
    compute county=35
else if any {zip,59201,59212,59213,59218,59226,59227,59228,59229,
        59245.59246,59255)
    compute county=85
else if any {zip,59211.59216,59219.59242.59247,59252,59254,59256,
        59257,59258,59275)
    compute county=91
else if any (zip,59214,59215,59274)
    compute county=55
else if any {zip,59217,59220,59221,59243,59262,59270,59271)
    compute county=83
else if any (zip,59222,59224,59253,59263,59276)
    compute county=19
else if any {zip,59223,59225,59230,59231,59240,59241,59244,59248,
        59249,59250.59260,59272,59273)
    compute county=105
else if any (zip,59301,59336,59340,59351,59352)
    compute county=17
else if any (zip.59311,59316,59319,59324,59332.59342)
        compute county=11
else if any (zip.59313,59344,59354)
        compute county=25
eise if any (zip.59315.59330.59331,59335.59339.59259.59711,59712)
        compute county=21
else if any {zip,59901,59902,59911.
        59912,59913.59916,59919,59920,59921,59922.59925.59926,
        59927.59928.59932,59936,59937)
        compute county=29
```

```
else if any (zip,59320,59353)
    compute county=109
else if any {zip,59261,59524,59533,59537,59538,59539,59543,59544,
        59546)
    compute county=71
else if any (zip.59326,59341,59349,59350)
    compute county=79
else if any {zip,59401,59402,59403,59404,59405,59412,59413,59414,
        5942 1,59443,59463,59465,59472,59473,59477,59478,59480,
        59483,59485,59486,59487)
    compute county=13
else if any {zip,59410,59548,59601,59602,59633,59635,59636,59639.
        59640)
    compute county=49
else if any {zip.59416,59425,59426.59432,59448,59456)
    compute county=73
else if any {zip.59419.59422,59433,59436.59467)
    compute county=99
else if any (zip.59420.59440,59442,59446,59450,59476,59520.59460)
    compute county=15
else if any (zip,59431,59435,59437,59444,59454,59466,59474,59482.
        59484)
    compute county=101
else if any (zip.59447.59452.59455,59462.59469,59479)
    compute county=45
else if any (zip,59461,59522,59531,59545)
    compute county=51
else if any {zip.59468,59730,59741,59742,59752,59753,59758,59760,
        59714,59715)
    compute county=31
else if any (zip,59501,59521,59525,59528,59530,59532,59540.59541)
    compute county=41
else if any (zip,59523,59526,59527,59529,59535,59542,59547)
    compute county=5
else if any (zip,59631,59632,59634,59638,59721,59759,59845)
    compute county=43
else if anv (zip,59641,59643,59644,59647)
    compute county=7
else if anv (zip.59701,59727.59743,59748,59750,59702,59703,59704,
        59705,59706,59707,59708,59709)
        compute county=93
else if any {zip.59710.59720.59729.59735.59737.59738.59740.59745,
        59747,59749.59751,59754,59755.59757)
    compute county=57
else if anv {zip.59713.59722,59723.59728.59731.59733.59734.59843)
    compute county=77
else if any {zip.59724,59725,59726,59732,59736,59739.59746.59761.
        59762)
        compute county=1
else if any (zip.59756)
        compute county=23
else if any {zip.59801,59802,59803,59804,59805,59806,59807.59808.
```

```
        59809.59823,59825,59826,59834,59836,59846,59847,59851.
        59862,59868,59869)
    compute county=63
else if any (zip,59820,59830,59842,59866,59867,59872)
    compute county=61
else if any (zip,59821,59824,59855,59860,59861,59863,59864,59865.
        59910,59914,59915,59929,59931)
    compute county=47
else if any (zip.59822.59831,59844,59848.59849,59852,59853,59856,
        59857,59859,59873)
    compute county=89
else if any (zip,59827,59828,59829,59833,59835,59840,59841,59870,
        59871,59875)
    compute county=81
else if any (zip,59832,59837,59838,59839,59850,59858)
    compute county=39
else if any (zip,59917.59918.59923,59924,59930.59933.59934,59935)
    compute county=53
end if
do if (region=1)
    do if (drainage=7)
                    do if (code=1560)
                compute unit2=89
                else if any (code.4560,4580)
                compute unit2=97
                end if
    else if (drainage=11)
                do if (code=3500)
                compute unit2=91
                end if
    end if
else if (region=2)
    do if (drainage=3)
                do if (code=475)
                compute unit2=82
                end if
    else if (drainage=4)
                do if any {code,600,630,660)
                compute unit2=83
                end if
    else if (drainage=5)
                do if (code=1456)
                compute unit2=87
                end if
    else if (drainage=6)
                do if any (code.1121,1140)
                compute unit2=86
                else if any (code.5263.5282)
                compute unit2=94
                end if
    end if
else if (region=3)
```

```
    do if (drainage=1)
            do if (code=500)
            campute unit2=80
            end if
    else if (drainage=2)
            do if any (code.425,450.475)
            compute unit2=81
            end if
    else if (drainage=9)
            do if (code=1710)
            compute unit2=88
            else if any (code,2090,6878.6916)
            compute unit2=90
            end if
    else if (drainage=13)
            do if any (code,3400.3440.3480)
            compute unit2=92
            end if
    else if (drainage=22)
        do if any (code,7058,7070,7072,7084)
        compute unit2=98
        end if
    end if
else if (region=4)
    do if (drainage=17)
                do if (code=4896)
                compute unit2=93
                else if any (code,6816,6832)
                compute unit2=95
                end if
    end if
else if (region=5)
    do if (drainage=22)
                do if (code=495)
                compute unit2=85
                else if any (code,742.756,770)
                compute unit2=84
                else if any (code,6104,6118,6132)
                compute unit2=96
                else if any {code.7000,7001,7014.7015,7028,
                        7042,70431
                compute unit2=99
                else if any (code.7056,7057)
                compute unit 2=98
                end if
    end if
end if
execute
do if (id>0)
compute vear=82
end if
execute
```

xsave outfile=unique82
execute

The final step in this initial portion of the data transformation was concatenating the years together using the ADD FILES command and discarding unneeded cases. Cases were discarded which either lacked a zip code or did not correspond to one of the 19 unique water codes (80-99) under study. Program segment 4 illustrates this process.

Program Segment 4
add files file=unique82/file=unique83/file=unique84/file=unique85
/file=unique86
keep=year id wave resident zip state county unit2 days region
drainage code type/map
select if (zip $>0$ and unit $2>0$ )
xsave outfile=condenc
execute
finish

## A.4. Combining Flow Variables with the Pressure File

Water flow data for this study came from the U.S. Geological Survey publications on Montana water flows. This data listed the mean monthly flows for each of the 19 study rivers. A data file was created which included for each record; year, month, river, and monthly flow. A basic inconsistency existed between the flow data and the pressure data in that the basic unit of time for flows was a month, and that unit for pressure was a wave. It was decided to adopt the wave as the standard unit since it most closely corresponded to the fishing license year. A new variable, wave, was therefore added to the flow file and year designations were altered where necessary. These changes were made according to the wave, month relationships shown in Table 4-1.

In order to allow a wider range of model specification several measures of flow, in addition to monthly, were calculated and added to the flow file. These were average yearly flow and the month averages for two separate three way clustering schemes. The following program segment shows how the above transformations were performed, and how the resulting flow file was merged with the pressure file.

## Program Segment 5

```
COMMENT INPUT DATA FILES
COMMENT ******************
```

FILE HANOLE FLOW/NAME="FLOW.DTA"
FILE HANDLE CONDENSE/NAME="PRESSURE-ALL.DTA"
COMMENT IN PROCESS SCRATCH FILES
COMMENT ****************************
FILE HANDLE MOFLO1/NAME="SCRATCH-CLUS1.DTA"
FILE HANDLE MOFLO2/NAME="SCRATCH-CLUS2.DTA"
FILE HANDLE FLOW2/NAME="SCRATCH-FLOW2.DTA"
FILE HANDLE FLOW3/NAME="SCRATCH-FLOW3.DTA"
FILE HANDLE DAYS/NAME="SCRATCH-DAYS.DTA"
FILE HANDLE DAYS2/NAME="SCRATCH-DAYS2.DTA"
COMMENT INTER-PROGRAM DATA FILES
COMMENT $* * * * * * * * * * * * * * * * * * * * * * * * * * * *$

```
FILE HANDLE CONSORT/NAME="PRESSURE-MODIFY.DTA"
FILE HANDLE CONFLO/NAME="FLOW-ALL.OTA"
FILE HANDLE PRESFLOW/NAME="PRESSURE-FLOW.DTA"
GET FILE=FLOW
DO IF {YEAR=82}
    DO !F ANY (MONTH,1,2,3,4)
                COMPUTE YEAR=81
    ELSE IF ANY (MONTH,5,6)
                COMPUTE WAVE=6
    ELSE
        COMPUTE WAVE=MONTH
    END IF
ELSE IF (YEAR=83)
    DO IF ANY (MONTH.5.6)
        COMPUTE WAVE=6
    ELSE
        COMPUTE WAVE=MONTH
```

END IF
ELSE IF (YEAR=84)
DO IF (MONTH=1)
COMPUTE WAVE=1
ELSE IF (MONTH=2)
COMPUTE WAVE=2
ELSE IF (MONTH=3) COMPUTE WAVE=1
ELSE IF (MONTH=4) COMPUTE WAVE=3
ELSE IF (MONTH=5) COMPUTE WAVE=5
ELSE IF (MONTH=6) COMPUTE WAVE=6
ELSE IF (MONTH=7) COMPUTE WAVE=8
ELSE IF (MONTH=8) COMPUTE WAVE=10
ELSE IF (MONTH=9) COMPUTE WAVE=12
ELSE IF (MONTH=10) COMPUTE WAVE=14
ELSE IF (MONTH=11) COMPUTE WAVE=16
ELSE IF (MONTH=12) COMPUTE WAVE=17
END IF
ELSE IF (YEAR=85)
DO IF (MONTH=1) COMPUTE WAVE=18
ELSE IF (MONTH=2) COMPUTE WAVE=19
ELSE IF (MONTH=3) COMPUTE WAVE=1
ELSE IF (MONTH=4) COMPUTE WAVE=2
ELSE IF (MONTH=5)
COMPUTE WAVE=3
ELSE IF (MONTH=6)
COMPUTE WAVE=4
ELSE IF (MONTH*7) COMPUTE WAVE=6
ELSE IF (MONTH=8) COMPUTE WAVE=8
ELSE IF (MONTH=9)
COMPUTE WAVE $=10$
ELSE IF (MONTH=10)
COMPUTE WAVE $=12$
ELSE IF (MONTH=11)
COMPUTE WAVE=14
ELSE IF (MONTH=12)
COMPUTE WAVE=15

```
    END IF
ELSE IF (YEAR=86)
    DO IF (MONTH=1)
            COMPUTE WAVE=16
    ELSE IF (MONTH=2)
            COMPUTE WAVE=17
    END IF
ENO IF
EXECUTE
DO IF {YEAR=83)
    DO IF ANY (MONTH,1.2.3.4)
                    COMPUTE YEAR=82
    END IF
ELSE IF (YEAR=84)
    OO IF ANY (MONTH,1,2)
            COMPUTE YEAR=83
    END IF
ELSE IF (YEAR=85)
    DO IF ANY (WAVE,18,19)
            COMPUTE YEAR=84
    END IF
ELSE IF (YEAR=86)
    DO IF ANY (WAVE,16,17)
            COMPUTE YEAR=85
    END IF
END IF
DO IF ANY (MONTH.5,6)
    COMPUTE MOCLUS2=1
ELSE IF ANY (MONTH,7,8,9,10)
    COMPUTE MOCLUS2=2
ELSE
    COMPUTE MOCLUS2=3
END IF
DO IF ANY (MONTH,5.6.7)
    COMPUTE MOCLUST=1
ELSE IF ANY (MONTH,8,9,10)
    COMPUTE MOCLUS1=2
ELSE
    COMPUTE MOCLUST=3
END IF
EXECUTE
SORT CASES BY UNIT2 YEAR WAVE
EXECUTE
XSAVE OUTFILE=FLOW2
EXECUTE
AGGREGATE OUTFILE=:
            /PRESORTED
            /BREAK=UNIT2 YEAR
            /YRFLOW=MEAN(FLOW)
EXECUTE
XSAVE OUTFILE=FLOW3
EXECUTE
```

```
GET FILE=FLOW2
AGGREGATE OUTFILE=*
    /BREAK=UNIT2 YEAR MOCLUS1
    /AVFLOW1=MEAN(FLOW)
EXECUTE
XSAVE OUTFILE=MOFLO1
EXECUTE
GET FILE=FLOW2
AGGREGATE OUTFILE=*
    /BREAK=UNIT2 YEAR MOCLUS2
    /AVFLOW2=MEAN(FLOW)
EXECUTE
XSAVE OUTFILE=MOFLO2
EXECUTE
GET FILE=CONDENSE
DO IF (YEAR=84)
    DO IF (WAVE=2)
        COMPUTE WAVE=1
    ELSE IF (WAVE=4)
        COMPUTE WAVE=3
    ELSE IF (WAVE=7)
        COMPUTE WAVE=6
    ELSE IF (WAVE=9)
        COMPUTE WAVE=8
    ELSE IF (WAVE=11)
            COMPUTE WAVE=10
    ELSE IF (WAVE=13)
            COMPUTE WAVE=12
    ELSE IF (WAVE=15)
        COMPUTE WAVE=14
    END IF
ELSE IF (YEAR=85)
    DO IF (WAVE=5)
        COMPUTE WAVE=4
    ELSE IF (WAVE=7)
        COMPUTE WAVE=6
    ELSE IF (WAVE=9)
            COMPUTE WAVE=8
    ELSE IF (WAVE=11)
            COMPUTE WAVE=10
    ELSE IF (WAVE=13)
        COMPUTE WAVE=12
    END IF
END IF
EXECUTE
SORT CASES BY UNIT2 YEAR WAVE
EXECUTE
XSAVE OUTFILE=CONSORT
EXECUTE
GET FILE=FLOW2
MATCH FILES FILE = "/FILE=FLOW3/BY UNIT2 YEAR
EXECUTE
```

```
SORT CASES BY UNIT2 YEAR MOCLUS1
EXECUTE
XSAVE OUTFILE=CONFLO
EXECUTE
MATCH FILES FILE= */FILE=MOFLO1/BY UNIT2 YEAR MOCLUS1
EXECUTE
SORT CASES BY UNIT2 YEAR MOCLUS2
EXECUTE
XSAVE OUTFILE=CONFLO
EXECUTE
MATCH FILES FILE=CONFLO/FILE=MOFLO2/BY UNIT2 YEAR MOCLUS2
EXECUTE
XSAVE OUTFILE=CONFLO
EXECUTE
GET FILE=CONFLO
SORT CASES BY UNIT2 YEAR month
EXECUTE
IF MISSING(YRFLOW) YRFLOW=0
DO IF (YRFLOW=0 AND YEAR=LAG(YEAR,1) AND UNIT2=LAG(UNIT2,1))
COMPUTE YRFLOW=LAGIYRFLOW,1)
END IF
EXECUTE
SORT CASES BY UNIT2 YEAR WAVE
DO IF (YRFLOW=0 AND YEAR=&AG(YEAR,1) AND UNIT2=LAG(UNIT2,1))
COMPUTE YRFLOW=LAGIYRFLOW,1)
END IF
EXECUTE
XSAVE OUTFILE=CONFLO
EXECUTE
GET FILE=CONFLO
SORT CASES BY UNIT2 YEAR MOCLUST
EXECUTE
IF MISSING(AVFLOW1) AVFLOW1=0
DO IF (AVFLOW1=0 AND YEAR=LAG(YEAR,1) AND UNIT2=LAG(UNIT2,1))
COMPUTE AVFLOW1=LAG(AVFLOW1,1)
END IF
EXECUTE
SORT CASES BY UNIT2 YEAR MOCLUS2
EXECUTE
IF MISSING(AVFLOW2) AVFLOW2=0
DO IF (AVFLOW2=0 AND YEAR=LAG(YEAR,1) AND UNIT2=LAGIUNIT2.1))
COMPUTE AVFLOW2=LAG(AVFLOW2,1)
END IF
EXECUTE
XSAVE OUTFILE=CONFLO
EXECUTE
GET FILE=CONFLO
SORT CASES BY UNIT2 YEAR WAVE
EXECUTE
MATCH FILES FILE= %/FILE=CONSORT/BY UNIT2 YEAR WAVE
EXECUTE
XSAVE OUTFILE=PRESFLOW
```

/KEEP=UNIT2 YEAR MONTH WAVE ID RESIDENT STATE COUNTY FLOW YRFLOW AVFLOW1 AVFLOW2 DAYS OVERNITE PURPOSE HRSFISH
EXECUTE
GET FILE=PRESFLOW
IF MISSING (MONTH) MONTH=0
DO IF (MONTH=0 AND YEAR=LAG(YEAR,1) AND UNIT2=LAG(UNIT2,1)
COMPUTE MONTH=LAG(MONTH,1)
END IF
IF MISSING (FLOW) FLOW=0
DO IF (FLOW=0 AND WAVE=LAG(WAVE,1) AND YEAR=LAG(YEAR,1)
AND UNIT2=LAG(UNIT2.1))
COMPUTE FLOW=LAG(FLOW,1)
END IF
IF MISSING (YRFLOW) YRFLOW=0
DO IF (YRFLOW=0 AND YEAR=LAG(YEAR,1) AND UNIT2=LAG(UNIT2,1\})
COMPUTE YRFLOW=LAGIYRFLOW,1)
END IF
IF MISSING (AVFLOW1) AVFLOW1=0
DO IF (AVFLOWI=0 AND WAVE=LAG(WAVE,1) AND YEAR=LAG(YEAR,1)
AND UNIT2=LAGIUNIT2,1 II
COMPUTE AVFLOW1=LAG(AVFLOW1.1)
END IF
IF MISSING (AVFLOW2) AVFLOW2=0
DO IF (AVFLOW2=0 AND WAVE=LAG(WAVE,1) AND YEAR=LAG(YEAR,1)
AND UNIT2=LAG\{UNIT2.11)
COMPUTE AVFLOW2=LAG(AVFLOW2,1)
END IF
EXECUTE
SELECT IF RANGE(YEAR,82,86)
SELECT IF (NOT MISSING(ID))
SELECT IF (NOT UNIT2=88)
SELECT IF (MISSING(PURPOSE) OR PURPOSE=1)
EXECUTE
XSAVE OUTFILE=PRESFLOW
EXECUTE
GET FILE=PRESFLOW
SELECT IF (YEAR=84)
AGGREGATE OUTFILE=*
/BREAK=YEAR UNIT2 ID RESIDENT WAVE
/STATE=MAX(STATE)
/COUNTY=MAX(COUNTY)
/FLOW=MAX(FLOW)
MYRFLOW=MAX(YRFLOW)
/AVFLOW1 = MAX(AVFLOW1)
/AVFLOW2=MAX(AVFLOW2)
/SUMDAYS=SUMIDAYS:
EXECUTE
XSAVE OUTFILE=DAYS
EXECUTE
GET FILE=PRESFLOW
SELECT IF (YEAR=82)
AGGREGATE OUTFILE=*

```
    /BREAK=YEAR UNIT2 ID RESIDENT WAVE
    /STATE=MAX(STATE)
    /COUNTY=MAX(COUNTY)
    /FLOW=MAX(FLOW)
    MRFLOW=MAX(YRFLOW)
    /AVF&OW1=MAX(AVFLOW1)
    /AVFLOW2=MAX(AVFLOW2)
    /SUMDAYS=SUM(DAYS)
EXECUTE
XSAVE OUTFILE=DAYS2
EXECUTE
GET FILE=PRESFLOW
SELECT IF (YEAR=83)
AGGREGATE OUTFILE=*
    /BREAK=YEAR UNIT2 ID RESIDENT WAVE
    /STATE=MAX(STATE)
    /COUNTY=MAX(COUNTY)
    /FLOW=MAX(FLOW)
    MYRFLOW=MAX(YRFLOW)
    /AVFLOW1=MAX(AVFLOW1)
    /AVFLOW2=MAX(AVFLOW2)
    /SUMDAYS=SUM(DAYS)
EXECUTE
XSAVE OUTFILE=PRESFLOW
EXECUTE
GET FILE=PRESFLOW
ADD FILES FILE=DAYS/FILE=DAYS2/FILE=*
EXECUTE
XSAVE OUTFILE=PRESFLOW
EXECUTE
GET FILE=PRESFLOW
COMPUTE DAYS=SUMDAYS
EXECUTE
SORT CASES BY YEAR UNIT2 ID RESIDENT WAVE
EXECUTE
XSAVE OUTFILE=PRESFLOW/DROP=SUMDAYS
EXECUTE
GET FILE=PRESFLOW
PRINT FORMATS UNIT2 (F2) YEAR (F2) WAVE (F2) ID (F4)
        RESIDENT (F1) STATE (F4.1) COUNTY (F3) FLOW (F5)
        YRFLOW (F5) AVFLOW1 (F5) AVFLOW2 (F5) DAYS (F2)
FINISH
```


## A.5. Ordering and Clustering Visitations

The three key variables in any Travel Cost Model are trips, populations, and travel distances. In its simplest form, the regional (multi-site) Travel Cost Model is:

$$
\text { TRIPS }_{i j} / \text { POP }_{i}=B_{1}-B_{2} \text { DIST }_{i j}
$$

Where:
TRIPS $_{i j}=$ The number of visits from origin $i$ to site $j$.
POP ${ }_{i}$ if The population of origin $i$.
DIST $_{i j}=$ Round trip distance from origin $i$ to site $j$.

In order to arrive at the variables needed for this type of modeling it was necessary to first decide upon what would constitute a zone of origin. It was decided that within Montana origin zones would be counties or groups of counties. Outside of Montana, origins would be states or groups of states. Within Montana the same county clusterings were used as Duffield, Loomis, and Brooks used in their 1987 study of the same 19 unique waters. Similarly, the same population weighted round trip distances between county origin zones and each unique water were employed as used by Duffield et. al. The adopting of clusters and distances from the 1987 study greatly simplified the in-state origin question. The only programming required was to (1) cluster the county origin zones for each unique water, (2) assign a round trip distance for each origin/destination pairing, and (3) assign a total population for each county cluster. Program segment 6 shows this process.

## Program Segment 6

FILE HANDLE JAS1/NAME='ALL.DTA, 201'
FILE HANDLE RON1/NAME='STREAM85.DTA; $I^{\prime}$
FILE HANDLE AGGIT/NAME='CLUSTER.DTA. $1^{\prime}$
file handle penult/name='penultimate.dtaif ${ }^{\circ}$

```
GET FILE=JAS1
    /KEEP=STATE COUNTY UNIT2 TRIPS
SORT CASES BY STATE COUNTY UNIT2
DOCUMENT
* This section of programming breaks the counties in Montana *
* into clusters (ORDER) for each of the 20 unique waters *
******************************************************************
OO IF (UNIT2>79)
DO IF (STATE=25)
DO IF (UNIT2=80)
    DO IF ANY(COUNTY,53,29.89,47,61,63.81,39,23)
    COMPUTE ORDER=80.1
    ELSE IF ANYICOUNTY,59,7,43)
    COMPUTE ORDER=80.3
    ELSE IF ANYICOUNTY,1)
    COMPUTE ORDER=80.4
    ELSE IF ANY(COUNTY,93)
    COMPUTE ORDER=80.5
    ELSE IF ANY(COUNTY.77.49.13)
    COMPUTE ORDER=80.6
    ELSE
    COMPUTE ORDER=80.2
    END IF
ELSE IF (UNIT2=81)
    DO IF ANYICOUNTY,53,29,89,47,61,63)
    COMPUTE ORDER=81.01
    ELSE IF ANYICOUNTY,35,101,51,73,99)
    COMPUTE ORDER=81.2
    ELSE IF ANYICOUNTY.41.5)
    COMPUTE ORDER=81.3
    ELSE IF ANYICOUNTY.71.69.105,33.19.91,85,83.55,21,79.109)
    COMPUTE ORDER=81.4
    ELSE IF ANY(COUNTY,87.17.75,25,11)
    COMPUTE ORDER=81.5
    ELSE IF ANYICOUNTY,107,37,65,103,111,95,97,9,3)
    COMPUTE ORDER=81.6
    ELSE IF ANY(COUNTY,31,67)
    COMPUTE ORDER=81.7
    ELSE IF ANY(COUNTY.57)
    COMPUTE ORDER=81.8
    ELSE IF ANY(COUNTY,I)
    COMPUTE ORDER=81.9
    ELSE IF ANYICOUNTY,81,39,23)
    COMPUTE ORDER=81.10
    ELSE IF ANYICOUNTY.93)
    COMPUTE ORDER=81.11
    ELSE IF ANYICOUNTY.43.7.59)
    COMPUTE ORDER=81.12
    ELSE IF ANY(COUNTY,77.49)
```

```
    COMPUTE ORDER=81.13
    ELSE IF ANY(COUNTY,13,15,45,27)
    COMPUTE ORDER=81.14
    END IF
ELSE IF(UNIT2=82)
    DO IF ANY(COUNTY.63)
    COMPUTE OROER=82.2
    ELSE IF ANY(COUNTY,39)
    COMPUTE ORDER=82.3
    ELSE IF ANY(COUNTY.23,93)
    COMPUTE ORDER=82.4
    ELSE IF ANY(COUNTY,81)
    COMPUTE ORDER=82.5
    ELSE
    COMPUTE ORDER=82.1
    END IF
ELSE IFIUNIT2=83)
    OO IF ANY(COUNTY.53,29,89,47,61,49}
    COMPUTE ORDER=83.01
    ELSE IF ANY(COUNTY,35,101,51,41,73,99,15,13,45,27)
    COMPUTE ORDER=83.2
    ELSE IF ANYICOUNTY,5,71,105,19,91,85,83,55,21,79,109)
    COMPUTE ORDER=83.3
    ELSE IF ANYICOUNTY,69,33,103,87)
    COMPUTE ORDER=83.4
    ELSE IF ANYICOUNTY,17,75,25,11)
    COMPUTE ORDER=83.5
    ELSE IF ANYICOUNTY,65,111,3)
    COMPUTE ORDER=83.6
    ELSE IF ANY(COUNTY,107,37,95,97,9)
    COMPUTE OROER=83.7
    ELSE IF ANYICOUNTY,1,57,31,67,43,7.59)
    COMPUTE ORDER=83.8
    ELSE IF ANY(COUNTY.23.93)
    COMPUTE ORDER=83.9
    ELSE FF ANY(COUNTY,77)
    COMPUTE ORDER=83.10
    ELSE IF ANY(COUNTY.81.39)
    COMPUTE ORDER=83.11
    ELSE
    COMPUTE ORDER=83.12
    END IF
ELSE IFIUNIT2=84)
    DO IF ANYICOUNTY 53,29,89.47)
    COMPUTE ORDER=84.1
    ELSE IF ANY(COUNTY,103,87.3.17,75,25.11)
    COMPUTE ORDER=84.3
    ELSE IF ANY{COUNTY,111.95,9)
    COMPUTE ORDER=84.4
    ELSE IF ANY(COUNTY.97)
    COMPUTE ORDER=84.5
    ELSE IF ANY(COUNTY,1,57,31,67)
```

```
    COMPUTE ORDER=84.6
    ELSE IF ANYICOUNTY,61,63,77,49,81,39,23,93,43,7,59)
    COMPUTE ORDER=84.7
    ELSE
    COMPUTE ORDER=84.2
    END IF
ELSE IF (UNIT2=85)
    DO IF ANY(COUNTY.53.29,89.47)
    COMPUTE ORDER=85.01
    ELSE IF ANYICOUNTY.17.75,25.11)
    COMPUTE ORDER=85.3
    ELSE IF ANY(COUNTY.87)
    COMPUTE ORDER=85.4
    ELSE IF ANY(COUNTY,103)
    COMPUTE ORDER=85.5
    ELSE IF ANY(COUNTY,3)
    COMPUTE ORDER=85.6
    ELSE IF ANY(COUNTY,111)
    COMPUTE ORDER=85.7
    ELSE IF ANYICOUNTY,9)
    COMPUTE ORDER=85.8
    ELSE IF ANY(COUNTY,95)
    COMPUTE ORDER=85.9
    ELSE IF ANY(COUNTY,1,57,31,67)
    COMPUTE ORDER=85.10
    ELSE IF ANYICOUNTY,61,63,81,39,23,93.77.49,13,43,7.59)
    COMPUTE ORDER=85.11
    ELSE
    COMPUTE ORDER=85.2
    END IF
ELSE IF (UNIT2=86)
    DO IF ANY(COUNTY,53,89,61)
    COMPUTE ORDER=86.01
    ELSE IF ANY(COUNTY,29,47)
    COMPUTE ORDER=86.2
    ELSE IF ANY( COUNTY,1)
    COMPUTE ORDER=86.4
    ELSE IF ANYICOUNTY.93)
    COMPUTE OROER=86.5
    ELSE IF ANYICOUNTY,23)
    COMPUTE ORDER=86.6
    ELSE IF ANYICOUNTY.81,39)
    COMPUTE ORDER=86.7
    ELSE IF ANYICOUNTY.77I
    COMPUTE ORDER=86.8
    ELSE IF ANYICOUNTY,49;
    COMPUTE OROER=86 9
    ELSE IF ANYICOUNTY,63)
    COMPUTE OROER=86.10
    ELSE
    COMPUTE ORDER=86.3
    END IF
```

```
ELSE IF (UNIT2=87)
    DO IF ANY(COUNTY,53)
    COMPUTE ORDER=87.01
    ELSE IF ANYICOUNTY,29)
    COMPUTE OROER=87.2
    ELSE IF ANYICOUNTY,107,97,95,9,37,65,111,3,103,87,17,75,25,11)
    COMPUTE OROER=87.4
    ELSE IF ANYICOUNTY,77.49.23.93,1,57.43,7.59,31,67)
    COMPUTE ORDER=87.5
    ELSE IF ANYICOUNTY,81,39)
    COMPUTE ORDER=87.6
    ELSE IF ANYICOUNTY.63)
    COMPUTE ORDER=87.7
    ELSE IF ANYICOUNTY,61)
    COMPUTE ORDER=87.8
    ELSE IF ANY(COUNTY.89)
    COMPUTE ORDER=87.9
    ELSE IF ANYICOUNTY.47I
    COMPUTE ORDER=87.10
    ELSE
    COMPUTE ORDER=87.3
    END IF
ELSE IF (UNIT2=88)
    DO IF ANY(COUNTY,53,29,89,47.61,63.77.49,13,15.81,39,23,93,43.7,59)
    COMPUTE ORDER=88.1
    ELSE IF ANYICOUNTY,1.57.31)
    COMPUTE ORDER=88.3
    ELSE
    COMPUTE ORDER=88.2
    END IF
ELSE IF (UNIT2=89)
    DO IF AP:Y(COUNTY.53.89.61,63.77.81.39)
    COMPU` ミORDER=89.1
    ELSE IF ANY(COUNTY,29)
    COMPUTE ORDER=89.2
    ELSE IF ANYICOUNTY.47)
    COMPUTE ORDER=89.3
    ELSE IF ANYICOUNTY,35,101,51,73,41,5,99,15,13,45,27)
    COMPUTE ORDER=89.4
    ELSE IF ANYICOUNTY,23,93)
    COMPUTE ORDER=89.6
    ELSE IF ANYICOUNTY.49)
    COMPUTE ORDER=89.7
    ELSE
    COMPUTE ORDER=89.5
    END IF
ELSE IF {UNIT2=90}
    DO IF ANYICOUNTY,53.29,89.47.61.63.77,49,13,39,81,1,23,93.43,7,59)
    COMPUTE OROER=90.01
    ELSE IF ANYICOUNTY,35,73,101.51,41}
    COMPUTE ORDER=90.2
    ELSE IF ANYICOUNTY,5,71)
```


## COMPUTE ORDER=90.3

ELSE IF ANYICOUNTY.105,79,85,91,83)
COMPUTE ORDER=90.4
ELSE IF ANY(COUNTY,33,55,21,79.109)
COMPUTE ORDER=90.5
ELSE IF ANY(COUNTY,87,17,75,25,11)
COMPUTE ORDER=90.6

- ELSE IF ANY(COUNTY,111,103,3)
. COMPUTE ORDER=90.7
- ELSE IF ANYICOUNTY,107,37,97,95,9)
- COMPUTE ORDER=90.8
. ELSE IF ANYICOUNTY,99,15,45,27,69)
. COMPUTE OROER=90.9
. ELSE IF ANYICOUNTY,57,31,67)
COMPUTE ORDER=90.10
END IF
ELSE IF (UNIT2=91)
DO IF ANY(COUNTY,53)
COMPUTE ORDER=91.1
ELSE IF ANYICOUNTY,29)
COMPUTE ORDER=91.2
ELSE
COMPUTE ORDER=91.3
END IF
ELSE IF (UNIT2=92)
DO IF ANY(COUNTY,53.89,47)
COMPUTE ORDER=92.01
ELSE IF ANY(COUNTY,29)
COMPUTE ORDER=92.2
ELSE IF ANYICOUNTY,35,73,101,51,41,5,71)
COMPUTE ORDER=92.3
ELSE IF ANYICOUNTY,105,19,91,85,83,55,33,79,21,109)
COMPUTE ORDER=92.4
ELSE IF ANYICOUNTY,25,11,75)
COMPUTE ORDER=92.5
ELSE IF ANY(COUNTY, 103,87,17)
COMPUTE ORDER=92.6
ELSE IF ANY(COUNTY,107,37,65,97,95,111,3,9)
COMPUTE ORDER=92.7
ELSE IF ANY(COUNTY,99,15,45,27,69)
COMPUTE ORDER=92.8
ELSE IF ANY(COUNTY,7.59,67)
COMPUTE ORDER=92.9
ELSE IF ANY(COUNTY.31)
COMPUTE ORDER=92.10
ELSE (F ANY(COUNTY, 1,57)
COMPUTE ORDER=92.11
ELSE IF ANYICOUNTY,43)
COMPUTE ORDER=92.12
ELSE IF ANYICOUNTY,23.93)
COMPUTE ORDER=92.13
ELSE IF ANY(COUNTY.49)

```
    COMPUTE ORDER=92.14
    ELSE IF ANY{COUNTY,77I
    COMPUTE ORDER=92.15
    ELSE IF ANY(COUNTY,13)
    COMPUTE ORDER=92.16
    ELSE IF ANY(COUNTY,61,63,81,39)
    COMPUTE ORDER=92.17
    END IF
ELSE IF (UNIT2=93)
    DO IF ANYICOUNTY,53,29,89,47,61,63,77,81,39,23)
    COMPUTE ORDER=93.01
    ELSE IF ANY(COUNTY,35,101,51)
    COMPUTE ORDER=93.2
    ELSE IF ANY(COUNTY,41,5)
    COMPUTE ORDER=93.3
    ELSE IF ANY(COUNTY,67)
    COMPUTE ORDER=93.5
    ELSE IF ANY(COUNTY,1,57,31)
    COMPUTE ORDER=93.6
    ELSE IF ANY(COUNTY,93,43,7,59)
    COMPUTE ORDER=93.7
    ELSE IF ANY(COUNTY.49)
    COMPUTE ORDER=93.8
    ELSE IF ANY(COUNTY,13)
    COMPUTE ORDER=93.9
    ELSE IF ANY(COUNTY,15)
    COMPUTE ORDER=93.10
    ELSE IF ANY(COUNTY.99)
    COMPUTE ORDER=93.11
    ELSE IF ANY(COUNTY,73)
    COMPUTE ORDER=93.12
    ELSE
    COMPUTE ORDER=93.4
    END IF
ELSE IF (UNIT2=94)
    DO IF ANY(COUNTY,53.29)
    COMPUTE ORDER=94.1
    ELSE {F ANY(COUNTY,103,87,17,75,25,11)
    COMPUTE ORDER=94.3
    ELSE IF ANYICOUNTY,81,23,93,1,57.31,67,49,43,7,59,107,97,95)
    COMPUTE ORDER=94.4
    ELSE IF ANY(COUNTY.9.37,65.111,3)
    COMPUTE ORDER=94.4
    ELSE IF ANYICOUNTY,77)
    COMPUTE ORDER=94.5
    ELSE IF ANYICOUNTY,39)
    COMPUTE ORDER=94.6
    ELSE IF ANY(COUNTY,63)
    COMPUTE ORDER=94.7
    ELSE IF ANY(COUNTY,89,61,47)
    COMPUTE ORDER=94.8
    ELSE
```

```
    COMPUTE ORDER=94.2
    END IF
ELSE IF (UNIT2=95)
    DO IF ANYICOUNTY,53,89,47,29)
    COMPUTE ORDER=95.1
    ELSE IF ANY(COUNTY,35,73,101,51)
    COMPUTE ORDER=95.2
    ELSE IF ANY\COUNTY,41,15,99,45,27,69,5,71,33,105,19,91,85,83)
    COMPUTE ORDER=95.3
    ELSE IF ANY{COUNTY,21,79,109)
    COMPUTE ORDER=95.3
    ELSE IF ANY{COUNTY,107,97,95,37,65,111,9,3,103,87,17,75,25,11)
    COMPUTE ORDER=95.4
    ELSE IF ANY(COUNTY.59,67)
    COMPUTE ORDER=95.5
    ELSE IF ANY(COUNTY,1,57,31,93,43,7)
    COMPUTE ORDER=95.6
    ELSE IF ANYICOUNTY,61,63,81,39,23)
    COMPUTE ORDER=95.7
    ELSE IF ANY(COUNTY,77,49)
    COMPUTE ORDERI995.8
    ELSE IF ANY(COUNTY,13)
    COMPUTE ORDER=95.9
    END IF
ELSE IF (UNIT2=96)
    DO IF ANY(COUNTY,53,29.89.47,61,63.77.49,81.39,23,93.1,57,43)
    COMPUTE ORDER=96.1
    ELSE IF ANYICOUNTY,7,59,31,67,97,107.37)
    COMPUTE ORDER=96.1
    ELSE IF ANYICOUNTY,3.103,87.17,75.25.11)
    COMPUTE ORDER=96.3
    ELSE IF ANYICOUNTY.65)
    COMPUTE ORDER=96.4
    ELSE (F ANYICOUNTY,111)
    COMPUTE ORDER=96.5
    ELSE IF ANY{COUNTY,9)
    COMPUTE ORDER=96.6
    ELSE IF ANYICOUNTY,95)
    COMPUTE ORDER=96.7
    ELSE
    COMPUTE ORDER=96.2
    END if
ELSE IF (UNIT2=97)
    DO IF ANYICOUNTY,53.89,61,63,47)
    COMPUTE ORDER=97.1
    ELSE IF ANY(COUNTY,29)
    COMPUTE ORDER=97.2
    ELSE IF ANYICOUNTY,1,23.93,57.43,7.59,31.67)
    COMPUTE ORDER=97.4
    ELSE IF ANYICOUNTY81,39)
    COMPUTE ORDER=97.5
    ELSE IF ANYICOUNTY.77.49)
```

```
    COMPUTE ORDER=97.6
    ELSE
    COMPUTE ORDER=97.3
    END IF
ELSE IF (UNIT2=98)
    DO IF ANYICOUNTY,53,29,89,47,61,63,77,49,81,39)
    COMPUTE ORDER=98.1
    ELSE IF ANY(COUNTY.65,111,3.103,87,17.75,25,11)
    COMPUTE ORDER=98.3
    ELSE IF ANYICOUNTY,107,97,95,9,37)
    COMPUTE ORDER=98.4
    ELSE IF ANY(COUNTY.67)
    COMPUTE ORDER=98.5
    ELSE IF ANY(COUNTY,31)
    COMPUTE ORDER=98.6
    ELSE IF ANY(COUNTY,1,57,23,93,43,7.59)
    COMPUTE ORDER=98.7
    ELSE
    COMPUTE ORDER=98.2
    END IF
ELSE IF {UNIT2=99}
    DO IF ANYICOUNTY,53.29,89,47,61,63,77.49,81,39,23,93,43,7,59)
    COMPUTE ORDER=99.1
    ELSE IF ANY(COUNTY.103.87)
    COMPUTE ORDER=99.3
    ELSE IF ANYICOUNTY,3,9,95)
    COMPUTE ORDER=99.4
    ELSE IF ANY{COUNTY.111)
    COMPUTE ORDER=99.5
    ELSE IF ANYICOUNTY,65,37.107.97,67,31,57.1)
    COMPUTE ORDER=99.6
    ELSE
    COMPUTE ORDER=99.2
    END IF
END IF
END :F
DOCUMENT
```



```
* This section assigns a population weighted round trip distance *
* to each county cluster / unique water pairing
```



```
DO IF (STATE=25)
OO IF {UNIT2=80)
    DO IF ANYICOUNTY,53,29,89.47,61,63,81,39.23)
    COMPUTE DIST=490
    ELSE IF ANY(COUNTY,59,7,43)
    COMPUTE DIST=140
    ELSE IF ANYICOUNTY,1)
    COMPUTE DIST=18
    ELSE IF ANY{COUNTY,93)
```

```
. COMPUTE DIST=141
. ELSE IF ANY(COUNTY,77.49,13)
. COMPUTE DIST=380
. ELSE
. COMPUTE DIST=720
    END IF
ELSE IF (UNIT2=81)
    DO IF ANYICOUNTY,53,29,89,47,61,63)
    COMPUTE DIST=420
    ELSE IF ANY(COUNTY,35,101.51.73.99)
    COMPUTE DIST=550
    ELSE IF ANY(COUNTY.41.5)
    COMPUTE DIST=600
    ELSE IF ANY(COUNTY.71,69,105,33,19,91,85,83,55,21,79,109)
    COMPUTE DIST=965
    ELSE IF ANY(COUNTY,87,17.75.25.11)
    COMPUTE DIST=250
    ELSE IF ANYICOUNTY,107,37.65.103.111,95,97.9,3)
    COMPUTE DIST=520
    ELSE IF ANY(COUNTY,31,67)
    COMPUTE DIST=160
    ELSE IF ANY(COUNTY,57)
    COMPUTE DIST=13
    ELSE IF ANY(COUNTY,1)
    COMPUTE DIST=49
    ELSE IF ANY(COUNTY.81,39.23)
    COMPUTE DIST=171
    ELSE IF ANY(COUNTY.93)
    COMPUTE DIST=73
    ELSE IF ANY(COUNTY,43.7.59)
    COMPUTE DIST=130
    ELSE IF ANYICOUNTY,77.49)
    COMPUTE DIST=213
    ELSE IF ANYICOUNTY,13,15,45,27)
    COMPUTE DIST=380
    END IF
ELSE IFIUNIT2=82)
    DO IF ANYICOUNTY,63)
    COMPUTE DIST=16
    ELSE IF ANY(COUNTY,39)
    COMPUTE DIST=120
    ELSE IF ANYICOUNTY.23.93)
    COMPUTE DIST=296
    ELSE IF ANYICOUNTY,81)
    COMPUTE DIST=13
    ELSE
    COMPUTE DIST=600
    END IF
    ELSE IFIUNIT2=831
    DO IF ANYICOUNTY,53,29,89,47,61,49)
    COMPUTE DIST=200
    ELSE IF ANY(COUNTY,35,101,51,41,73,99,15,13,45,27)
```

```
    COMPUTE DIST=320
    ELSE IF ANY{COUNTY,5,71,105,19,91,85,83,55,21,79.109)
    COMPUTE DIST=490
    ELSE IF ANYICOUNTY,69,33,103,871
    COMPUTE DIST=1000
    ELSE IF ANY(COUNTY,17,75,25,11)
    COMPUTE DIST=1000
    ELSE IF ANY(COUNTY,65,111,3)
    COMPUTE DIST=700
    ELSE IF ANY(COUNTY,107,37,95,97,9)
    COMPUTE DIST=700
    ELSE IF ANY(COUNTY,1,57,31,67,43,7,59)
    COMPUTE DIST=280
    ELSE IF ANY(COUNTY,23,93)
    COMPUTE DIST=230
    ELSE IF ANY(COUNTY,77)
    COMPUTE DIST=90
    ELSE IF ANY(COUNTY,81,39)
    COMPUTE DIST=80
    ELSE
    COMPUTE DIST=42
    END IF
ELSE IF(UNIT2=84}
    DO IF ANYICOUNTY,53,29,89,47)
    COMPUTE DIST=320
    ELSE IF ANY(COUNTY,103,87,3,17,75,25,11)
    COMPUTE DIST=470
    ELSE IF ANY(COUNTY,111,95,9)
    COMPUTE OIST=182
    ELSE IF ANY(COUNTY,97)
    COMPUTE DIST=6
    ELSE IF ANY(COUNTY,1,57,31,67)
    COMPUTE DIST=128
    ELSE IF ANYICOUNTY,61,63,77,49,81,39,23,93,43,7,59)
    COMPUTE DIST=320
    ELSE
    COMPUTE DIST=480
    END IF
    ELSE IF (UNIT2=85)
    DO IF ANY(COUNTY,53,29,89,47)
    COMPUTE DIST=1040
    ELSE IF ANYICOUNTY,17,75,25,11)
    COMPUTE OIST=346
    ELSE IF ANY(COUNTY,87)
    COMPUTE DIST=244
    ELSE IF ANYICOUNTY,103)
    COMPUTE OIST= 126
    ELSE IF ANYICOUNTY,3)
    COMPUTE DIST=46
    ELSE IF ANYICOUNTY,1111)
    COMPUTE DIST=167
    ELSE IF ANYICOUNTY,9)
```

```
    COMPUTE DIST=235
    ELSE IF ANYICOUNTY,95)
    COMPUTE DIST=200
    ELSE IF ANYICOUNTY,1,57,31,67)
    COMPUTE DIST=450
    ELSE IF ANY(COUNTY,61,63,81,39,23,93,77.49.13,43,7,59)
    COMPUTE DIST=700
    ELSE
    COMPUTE DIST=670
    END IF
ELSE IF (UNIT2=86)
    DO IF ANY(COUNTY,53,89,61)
    COMPUTE DIST=240
    ELSE IF ANY(COUNTY,29,47)
    COMPUTE OIST=240
    ELSE IF ANY(COUNTY,1)
    COMPUTE DIST=160
    ELSE IF ANYICOUNTY.93)
    COMPUTE DIST=65
    ELSE IF ANYICOUNTY,23)
    COMPUTE DIST=25
    ELSE IF ANYICOUNTY,81,39)
    COMPUTE DIST=100
    ELSE IF ANY(COUNTY,77)
    COMPUTE DIST=18
    ELSE IF ANY(COUNTY.49)
    COMPUTE DIST=100
    ELSE IF ANYICOUNTY,63)
    COMPUTE DIST=27
    ELSE
    COMPUTE DIST=550
    END IF
ELSE IF (UNIT2=87)
    DO IF ANYICOUNTY.53)
    COMPUTE DIST=200
    ELSE IF ANY(COUNTY,29)
    COMPUTE DIST=200
    ELSE IF ANY(COUNTY,107,97,95,9,37,65,111,3,103,87,17,75,25,11)
    COMPUTE DIST=700
    ELSE IF ANY(COUNTY,77,49,23,93,1,57,43,7,59,31,67)
    COMPUTE DIST=300
    ELSE IF ANYICOUNTY,81,39)
    COMPUTE DIST=95
    ELSE IF ANY(COUNTY.63)
    COMPUTE DIST=35
    ELSE IF ANYICOUNTY.61)
    COMPUTE DIST=12
    ELSE IF ANYICOUNTY,89)
    COMPUTE DIST=80
    ELSE IF ANY(COUNTY,47)
    COMPUTE DIST=130
    ELSE
```

```
    COMPUTE DIST=550
    END IF
ELSE IF (UNIT2=88)
    DO IF ANY(COUNTY,53,29,89,47.61,63,77,49,13,15,81,39,23,93,43,7,59)
    COMPUTE DIST=600
    ELSE IF ANY(COUNTY,1,57.31)
    COMPUTE DIST=13
    ELSE
    COMPUTE DIST=430
    END IF
ELSE IF (UNIT2=89)
    DO IF ANY(COUNTY,53,89,61,63,77,81,39)
    COMPUTE DIST=240
    ELSE IF ANY(COUNTY,29)
    COMPUTE DIST=11
    ELSE IF ANY(COUNTY,47)
    COMPUTE DIST=100
    ELSE IF ANY(COUNTY,35,101,51,73,41,5,99,15,13,45,27)
    COMPUTE DIST=642
    ELSE IF ANY(COUNTY,23,93)
    COMPUTE DIST=470
    ELSE IF ANYICOUNTY,49}
    COMPUTE DIST=460
    ELSE
    COMPUTE DIST=800
    END IF
ELSE IF (UNIT2=90)
    DO IF ANYICOUNTY,53,29,89,47,61,63,77,49,13,39,81,1,23,93,43,7,59)
    COMPUTE DIST=628
    ELSE IF ANY(COUNTY,35,73,101,51,41)
    COMPUTE DIST=600
    ELSE IF ANY(COUNTY,5,71)
    COMPUTE DIST=630
    ELSE IF ANY{COUNTY,105,19.85,91,83)
    COMPUTE OIST=800
    ELSE IF ANY(COUNTY,33,55,21.79,109)
    COMPUTE OIST=840
    ELSE IF ANYICOUNTY,87,17,75,25,111
    COMPUTE DIST=600
    ELSE IF ANY(COUNTY,111,103,3)
    COMPUTE DIST=310
    ELSE IF ANY(COUNTY,107,37,97,95,9)
    COMPUTE DIST=190
    ELSE IF ANYICOUNTY,99,15,45,27,69)
    COMPUTE DIST=370
    ELSE IF ANY(COUNTY,57,31,67)
    COMPUTE DIST=29
    END IF
ELSE IF (UNIT2=91)
    DO IF ANYICOUNTY.53)
    COMPUTE DIST=25
    ELSE IF ANYICOUNTY,29)
```

```
    COMPUTE DIST=196
    ELSE
    COMPUTE DIST=795
    END IF
ELSE IF (UNIT2=92)
    DO IF ANY(COUNTY,53,89,47)
    COMPUTE DIST=780
    ELSE IF ANYICOUNTY,29)
    COMPUTE DIST=640
    ELSE IF ANYICOUNTY,35,73,101,51,41,5,71)
    COMPUTE DIST=600
    ELSE IF ANYICOUNTY,105,19.91,85.83,55,33,79,21,109)
    COMPUTE DIST=775
    ELSE IF ANYICOUNTY,25.11.75)
    COMPUTE DIST=800
    ELSE IF ANYICOUNTY,103,87.17)
    COMPUTE DIST=637
    ELSE IF ANYICOUNTY,107.37.65.97.95,111,3.9)
    COMPUTE DIST=347
    ELSE IF ANY{COUNTY,99,15,45,27,69)
    COMPUTE DIST=440
    ELSE IF ANY(COUNTY.7.59.67)
    COMPUTE DIST=200
    ELSE IF ANY(COUNTY,31)
    COMPUTE DIST=64
    ELSE IF ANYICOUNTY.1.57)
    COMPUTE DIST=55
    ELSE IF ANY(COUNTY.43)
    COMPUTE DIST=100
    ELSE IF ANY(COUNTY,23,93)
    COMPUTE DIST=148
    ELSE IF ANY(COUNTY.49)
    COMPUTE DIST=140
    ELSE IF ANY(COUNTY,77)
    COMPUTE DIST=220
    ELSE IF ANYICOUNTY.13)
    COMPUTE DIST=360
    ELSE IF ANY(COUNTY,61,63,81,39)
    COMPUTE OIST=400
    END IF
ELSE IF (UNIT2=93)
    OO IF ANY(COUNTY,53,29,89,47,61,63.77.81,39.23)
    COMPUTE DIST=440
    ELSE IF ANYICOUNTY,35,101,51)
    COMPUTE DIST=300
    ELSE IF ANYICOUNTY,41.5)
    COMPUTE DIST=290
    ELSE IF ANY(COUNTY.67)
    COMPUTE DIST=310
    ELSE IF ANYICOUNTY,1,57,31)
    COMPUTE DIST=263
    ELSE IF ANYICOUNTY,93.43.7.59)
```

```
    COMPUTE DIST=190
    ELSE IF ANY(COUNTY,49)
    COMPUTE DIST=73
    ELSE IF ANY(COUNTY,13)
    COMPUTE DIST=70
    ELSE IF ANY(COUNTY,15)
    COMPUTE DIST=90
    ELSE IF ANY(COUNTY,99)
    COMPUTE DIST=180
    ELSE IF ANY(COUNTY.73)
    COMPUTE DIST=172
    ELSE
    COMPUTE DIST=634
    END IF
ELSE IF (UNIT2=94)
    DO IF ANY(COUNTY.53.29)
    COMPUTE DIST=300
    ELSE IF ANY(COUNTY,103,87,17,75,25,11)
    COMPUTE DIST=930
    ELSE IF ANY(COUNTY,81,23,93,1,57,31,67.49,43,7,59,107,97,95)
    COMPUTE DIST=930
    ELSE IF ANY(COUNTY,9,37,65,111,3)
    COMPUTE DIST=530
    ELSE IF ANY(COUNTY,77)
    COMPUTE DIST=110
    ELSE IF ANY(COUNTY,39)
    COMPUTE DIST=20
    ELSE IF ANY(COUNTY.63)
    COMPUTE DIST=56
    ELSE IF ANYICOUNTY,89.61.47)
    COMPUTE DIST=200
    ELSE
    COMPUTE DIST=550
    END IF
ELSE IF (UNIT2=95)
    DO IF ANY(COUNTY,53,89,47,29)
    COMPUTE DIST=550
    ELSE IF ANYICOUNTY,35.73.101.51)
    COMPUTE DIST=240
    ELSE IF ANYICOUNTY,41,15,99,45,27,69.5,71,33,105,19,91,85,83)
    COMPUTE DIST=155
    ELSE IF ANY(COUNTY,21,79,109)
    COMPUTE DIST=155
    ELSE IF ANYICOUNTY,107,97.95.37,65,111,9,3.103.87,17.75,25,11%
    COMPUTE DIST=430
    ELSE IF ANY(COUNTY,59.67)
    COMPUTE DIST=60
    ELSE IF ANY(COUNTY,1,57,31,93,43.7)
    COMPUTE DIST=200
    ELSE IF ANYICOUNTY,61,63,81.39,23)
    COMPUTE DIST=400
    ELSE IF ANYICOUNTY.77.49)
```

```
    COMPUTE DIST=120
    ELSE IF ANY(COUNTY.13)
    COMPUTE DIST=65
    END IF
ELSE IF (UNIT2=96)
    DO IF ANYICOUNTY,53,29,89,47,61,63,77,49,81,39,23,93,1,57,43)
    COMPUTE DIST=770
    ELSE IF ANYICOUNTY,7,59,31,67,97,107,37)
    COMPUTE DIST=770
    ELSE IF ANYICOUNTY,3,103,87,17,75,25,11}
    COMPUTE DIST=380
    ELSE IF ANYICOUNTY,65)
    COMPUTE DIST=200
    ELSE IF ANY(COUNTY,111)
    COMPUTE DIST=96
    ELSE IF ANY(COUNTY,9)
    COMPUTE DIST=90
    ELSE IF ANY(COUNTY,95)
    COMPUTE DIST=16
    ELSE
    COMPUTE DIST=574
    END IF
ELSE IF (UNIT2=97)
    DO IF ANY(COUNTY,53,89,61,63,47)
    COMPUTE DIST=120
    ELSE IF ANY(COUNTY,29)
    COMPUTE DIST=40
    ELSE IF ANY(COUNTY,1,23.93,57.43.7.59,31.67)
    COMPUTE DIST=550
    ELSE IF ANYICOUNTY,81.39)
    COMPUTE DIST=425
    ELSE IF ANYICOUNTY,77.49)
    COMPUTE DIST=366
    ELSE
    COMPUTE DIST=725
    END IF
ELSE IF (UNIT2=98)
    DO IF ANYICOUNTY,53,29,89,47,61,63,77,49,81,39)
    COMPUTE DIST=740
    ELSE IF ANYICOUNTY,65,111,3,103,87,17,75,25,11)
    COMPUTE DIST=295
    ELSE IF ANYICOUNTY,107.97,95,9.37)
    COMPUTE DIST=9
    ELSE IF ANY(COUNTY.67)
    COMPUTE DIST=4
    ELSE IF ANYICOUNTY,31)
    COMPUTE DIST=123
    ELSE IF ANY(COUNTY,1,57,23.93,43.7.59)
    COMPUTE DIST=280
    ELSE
    COMPUTE DIST=510
    END IF
```

```
ELSE IF (UNIT2=99)
    DO IF ANYICOUNTY,53,29,89,47.61,63,77,49,81,39,23,93,43,7,59)
    COMPUTE DIST=670
    ELSE IF ANY(COUNTY,103,87)
    COMPUTE DIST=90
    ELSE IF ANY(COUNTY,3,9,95)
    COMPUTE DIST=10
    ELSE IF ANY(COUNTY,111)
    COMPUTE DIST=36
    ELSE IF ANYICOUNTY,65,37,107,97,67,31,57,1)
    COMPUTE DIST=13
    ELSE
    COMPUTE DIST=480
    END IF
END IF
END IF
DOCUMENT
***********************************************************
* These cornmands assign a total population for each
* of the river specific county clusters
```



```
IF (ORDER=80.1) CLUSPOP=221200
IF (ORDER=80.2) CLUSPOP =410800
IF (ORDER=80.3) CLUSPOP=13600
IF (ORDER=80.4) CLUSPOP=8700
IF (ORDER=80.5) CLUSPOP=35200
IF (ORDER=80.6) CLUSPOP=134500
IF (ORDER=81.01) CLUSPOP=182400
IF (OROER=81.2) CLUSPOP=33000
IF (ORDER=81.3) CLUSPOP=25900
IF (ORDER=81.4) CLUSPOP=71600
IF (ORDER=81.5) CLUSPOP=34800
IF (ORDER=81.6) CLUSPOP=157200
IF (ORDER=81.7) CLUSPOP=60900
IF (ORDER=81.8) CLUSPOP=5800
IF (ORDER=81.9)}\mathrm{ CLUSPOP=8700
IF (ORDER=81.10) CLUSPOP=38800
IF (ORDER=81.11) CLUSPOP=35200
IF (ORDER=81.12) CLUSPOP=13600
IF (ORDER=81.13) CLUSPOP=52700
IF (ORDER=81.14) CLUSPOP=103600
IF (ORDER=82.1) CLUSPOP=673500
IF (ORDER=82.2) CLUSPOP=76500
IF (ORDER=82.3) CLUSPOP=2800
IF (ORDER=82.4) CLUSPOP=46400
IF (ORDER=82.5) CLUSPOP=24800
```

[^10]IF (ORDER=87.9) CLUSPOP=9200 IF (ORDER=87.10) CLUSPOP=20400

## IF (ORDER=88.1) CLUSPOP=410700 IF (ORDER=88.2) CLUSPOP $=351200$ IF (ORDER=88.3) CLUSPOP=62100

IF (ORDEA=89.1) CLUSPOP=142600 IF (ORDER=89.2) CLUSPOP=53900 IF (ORDER=89.3) CLUSPOP=20400 IF (ORDER=89.4) CLUSPOP=162500 IF (ORDER=89.5) CLUSPOP=352400 IF (ORDER=89.6) CLUSPOP=46400 IF (ORDER=89.7) CLUSPOP=45800

IF (ORDER=90.01) $C L U S P O P=413200$
IF (ORDER=90.2) CLUSPOP=45100
IF (ORDER=90.3) $C L U S P O P=13100$ IF (ORDER=90.4) CLUSPOP=44700 IF (ORDER=90.5) CLUSPOP=20500 IF (ORDER=90.6) CLUSPOP=33800 IF (ORDER=90.7) CLUSPOP=135700 IF (ORDER=90.8) CLUSPOP=21300 IF (ORDER=90.9) CLUSPOP=28900 IF (ORDER=90.10) $\mathrm{CLUSPOP}=66700$

IF (ORDER=91.1) $C L U S P O P=18700$ IF (ORDER=91.2) CLUSPOP=53900 IF (ORDER=91.3) CLUSPOP=751400

IF (ORDER=92.01) CLUSPOP=48300 IF (ORDER=92.2) CLUSPOP $=53900$ IF (ORDER=92.3) $\mathrm{CLUSPOP}=58200$ IF (ORDER=92.4) CLUSPOP=65200 IF (ORDER $=92.5$ ) CLUSPOP $=8100$ IF (ORDER=92.6) CLUSPOP=27700 IF (ORDER=92.7) CLUSPOP=156200 IF (ORDER=92.8) CLUSPOP=28900 IF (ORDER=92.9) CLUSPOP $=18900$ IF (ORDER=92.10) CLUSPOP=47600 IF (ORDER=92.11) CLUSPOP=14500 If (ORDER=92.12) CLUSPOP=8000 IF (ORDER $=92.13$ ) $\mathrm{CLUSPOP}=46400$ IF (ORDER=92.14) $C L U S P O P=45800$ IF (ORDER=92.15) $\mathrm{CLUSPOP}=6900$ IF (ORDER=92.16) CLUSPOP $=81800$ IF (ORDER=92.17) CLUSPOP=107800

IF (ORDER=93.01) CLUSPOP=228100
IF (ORDER=93.2) CLUSPOP=19500
IF (ORDER=93.3) CLUSPOP=25900
IF (ORDER=93.4) $\mathrm{CLUSPOP}=279000$

IF (ORDER=93.5) CLUSPOP=13300 IF (ORDER=93.6) $C L U S P O P=62100$ IF (ORDER=93.7) CLUSPOP=48800 IF (ORDER=93.8) $\mathrm{CLUSPOP}=45800$ IF (ORDER=93.9) CLUSPOP=81800 IF (ORDER=93.10) $C L U S P O P=6200$ IF (ORDER=93.11) CLUSPOP=6400 IF (ORDER=93.12) $\mathrm{CLUSPOP}=7100$

```
IF (OROER=94.1) CLUSPOP=72600
IF (ORDER=94.2) CLUSPOP=234100
IF (ORDER=94.3) CLUSPOP=35800
IF (ORDER=94.4) CLUSPOP=362000
IF (ORDER=94.5) CLUSPOP=6900
IF (ORDER=94.6) CLUSPOP=2400
IF (ORDER=94.7) CLUSPOP=76500
IF (ORDER=94.8) CLUSPOP=33300
```

IF (ORDER=95.1) CLUSPOP $=102200$
IF (ORDER=95.2) CLUSPOP=26600
IF (ORDER=95.3) $\mathrm{CLUSPOP}=125500$
IF (ORDER=95.4) CLUSPOP=192000
IF (ORDER=95.5) $C L U S P O P=15500$
IF (ORDER=95.6) CLUSPOP=108700
IF (OROER=95.7) CLUSPOP=119000
IF (ORDER=95.8) $C L U S P O P=52700$
IF (ORDER=95.9) $C L U S P O P=81800$
IF (ORDER=96.1) CLUSPOP=404600
IF (ORDER=96.2) CLUSPOP=234100
IF (ORDER=96.3) CLUSPOP=47300
IF (ORDER=96.4) CLUSPOP=4700
IF (ORDER=96.5) CLUSPOP=118700
IF (ORDER=96.6) $\mathrm{CLUSPOP}=8600$
IF (ORDER $=96.7$ ) CLUSPOP $=6000$
(F (ORDER=97.1) CLUSPOP $=128500$
IF (ORDER=97.2) CL (OSPOP $=53900$
IF (ORDER $=97.3$ ) CLUSPOP $=426100$
IF (ORDER=97.4) CLUSPOP=135400
IF (ORDER=97.5) CLUSPOP $=27600$
IF IORDER=97.6) $\mathrm{CLUSPOP}=52700$
IF (ORDER $=98.1$ ) $\mathrm{CLUSPOP}=263500$
IF (ORDER $=98.2$ ) $\mathrm{CLUSPOP}=234100$
If (ORDER $=98.3$ ) $\mathrm{CLUSPOP=169700}$
IF (ORDER=98.4) $\mathrm{CLUSPOP}=21300$
IF (ORDER=98.5) CLUSPOP=13300
IF (ORDER $=98.6$ ) $\mathrm{CLUSPOP}=47600$
IF IORDER=98.7) $\mathrm{CLUSPOP}=74600$
IF (ORDER=99.1) CLUSPOP=327800

IF (ORDER=99.2) CLUSPOP $=255700$
IF (ORDER=99.3) $C L U S P O P=14200$
IF (ORDER=99.4) CLUSPOP=21600
IF (ORDER=99.5) CLUSPOP=118700
(F (ORDER=99.6) $C L U S P O P=86800$

A somewhat different approach was used in clustering visitations from outside Montana. In order to greatly reduce the amount of programming necessary the assumption was made that the 19 unique waters could be grouped into 3 clustered destinations. This assumption seemed reasonable due to the large distances often traveled by out of state visitors, and the often small distances between the rivers within the clusters.

Having clustered the 19 waters, 6 logical state paths emanating from Montana were defined and states or groups of states within those paths were ordered as to their proximity to Montana. Next, population weighted round trip distances between each state or state cluster / river cluster pairing were assigned (program segment 7).

## Program Segment 7

DOCUMENT

```
* This section clusters the unique waters into three general *
* groups for out-of-state analysis
```

DO IF ANY(UNIT2,84,96,85,98,99)
COMPUTE RIVCLUST=3
ELSE IF ANYIUNIT2,81,80,92,88,90,95,93)
COMPUTE RIVCLUST=2
ELSE
COMPUTE RIVCLUST=1
END IF

DOCUMENT


```
* This section one of six path numbers to each state *
```


DO IF ANY(STATE,11.1,46.36)
COMPUTE PATH=1
ELSE IF ANY(STATE,11.2,27.4.1,4.2)
COMPUTE PATH=2
ELSE IF ANY(STATE,49,43.5.30.2.42)
COMPUTE PATH=3
ELSE IF ANY(STATE,33,22,48,14,12,13,21,34,47,37,31,29,19,7,18,28,44,20,38,6)
COMPUTE PATH=4
ELSE IF ANY(STATE,40.26.15,24,35,3,17,23,16,41,1,10,39,32.45,9)
COMPUTE PATH=5
ELSE IF ANY(STATE,50)
COMPUTE PATH=6
END IF
DOCUMENT


* In this section the states within each path are ordered as to
* their distance from Montana. Those states within a path which *
* are roughly equidistant from Montana are given the same ordering *
* number and will be considered unique populations in the analysis *
* which is to follow

DO IF (PATH=1)
DO IF (STATE=11.1)
COMPUTE ORDER=1.1
ELSE IF (STATE=46)
COMPUTE ORDER=1.2
ELSE IF (STATE=36)
COMPUTE ORDER=1.3
END $1 F$
ELSE IF (PATH=2)
DO IF (STATE=11.2)
COMPUTE ORDER=2.1
ELSE IF ISTATE=27)
COMPUTE ORDER=2.2
ELSE IF (STATE=4.1)
COMPUT'E ORDER=2.3
ELSE IF (STATE=42)
COMPUTE ORDER=2.4
END IF
ELSE IF (PATH=3)
DO IF (STATE=49)
COMPUTE ORDER=3.1
ELSE IF ANY(STATE,43.5)
COMPUTE ORDER=3.2

```
    ELSE IF ANY(STATE,30,2)
    COMPUTE ORDER=3.3
    ELSE IF ISTATE=42)
    COMPUTE ORDER=3.4
    END IF
ELSE IF (PATH=4)
    DO IF (STATE=33)
    COMPUTE ORDER=4.1
    ELSE IF (STATE=22)
    COMPUTE ORDER=4.2
    ELSE IF ANY(STATE,48,14)
    COMPUTE ORDER=4.3
    ELSE IF (STATE=12)
    COMPUTE ORDER=4.4
    ELSE IF ANYISTATE,13,21)
    COMPUTE ORDER=4.5
    ELSE IF (STATE=34)
    COMPUTE ORDER=4.6
    ELSE IF ANYISTATE,47,37,31)
    COMPUTE ORDER=4.7
    ELSE IF ANYISTATE,29,19,7,18,28,14,20,38,6)
    COMPUTE ORDER=4.8
    END IF
ELSE IF (PATH=5)
    DO IF (STATE=40)
    COMPUTE ORDER=5.1
    ELSE IF (STATE=26)
    COMPUTE ORDER=5.2
    ELSE IF ANYISTATE,15,24)
    COMPUTE ORDER=5.3
    ELSE IF ANY(STATE,35,3)
    COMPUTE ORDER=5.4
    ELSE IF ANY(STATE,17,23,16.41)
    COMPUTE ORDER=5.5
    ELSE IF ANY(STATE,1,10,39,32,45)
    COMPUTE ORDER=5.6
    ELSE IF (STATE=9)
    COMPUTE ORDER=5.7
    END IF
ELSE IF (PATH=6)
    COMPUTE ORDER=6.1
END IF
```

DOCUMENT


* This section assigns a population weighted round trip distance
* to each state / river cluster pairing


```
DO IF (RIVCLUST=1)
    COMPUTE DIST=4280
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=4042
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=3598
    END IF
ELSE IF ANY (STATE,2,30)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=3086
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=2848
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=2404
    END IF
ELSE IF ANY {STATE,3,35}
    DO IF (RIVCLUST=1)
    COMPUTE DIST=3530
    ELSE :F (RIVCLUST=2)
    COMPUTE DIST=3292
    ELSE (F (RIVCLUST=3)
    COMPUTE DIST=2848
    END IF
ELSE IF (STATE=4.1)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=1688
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=1926
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=2370
    END IF
ELSE IF (STATE=4.2)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=1794
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=2032
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=02476
    END IF
ELSE IF ANY (STATE.5.43)
    DO IF (RIVCLUST=1)
    COMPUTE DIST }=65
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=896
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=1340
    END IF
ELSE IF ANY (STATE,6,38,20,44,28,18,7,19,29)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=5076
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST }=483
```

```
- ELSE IF (RIVCLUST*3)
        COMPUTE DIST=4394
        END IF
ELSE IF (STATE=9)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=5138
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=4900
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=4456
    END IF
ELSE IF (STATE=11.1)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=242
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=480
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=924
    END IF
ELSE IF (STATE=11.2)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=530
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=768
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=1212
    END IF
ELSE IF (STATE=12)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=3144
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=2906
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=2462
    END IF
ELSE IF ANY (STATE,13,21)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=3506
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=3268
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=2824
    END IF
ElSE IF ANY (STATE,14,48)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=2998
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=2760
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=2316
    END IF
```

```
ELSE IF ANY (STATE,15,24)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=2762
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=2524
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=2080
    END IF
ELSE IF ANY (STATE,16,17,23,41)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=3888
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=3650
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=3206
    END IF
ELSE IF (STATE=22)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=2336
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=2098
    ELSE IF {RIVCLUST=3)
    COMPUTE DIST=1654
    END IF
ELSE IF (STATE=26)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=2362
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=2124
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=1680
    END IF
ELSE IF (STATE=27)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=1236
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=1474
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=1918
    END IF
ELSE IF ANY (STATE,31,37,47)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=4760
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=4522
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=4078
    ENO IF
ELSE IF (STATE=33)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=1518
    ELSE IF (RIVCLUST=2)
```

```
. COMPUTE DIST=1280
- ELSE IF (RIVCLUST=3)
. COMPUTE DIST=836
    END IF
ELSE IF (STATE=34)
. DO IF (RIVCLUST=1)
. COMPUTE DIST=3898
. ELSE IF (RIVCLUST=2)
. COMPUTE DIST=3660
. ELSE IF (RIVCLUST=3)
    COMPUTE DIST=3216
    END IF
ELSE IF (STATE=36)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=1129
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=1358
    ELSE IF (RIVCLUST=3)
    COMPUTE DIST=1802
    END IF
ELSE IF (STATE=40)
    DO IF (RIVCLUST=1)
    COMPUTE DIST=1316
    ELSE IF (RIVCLUST=2)
    COMPUTE DIST=1074
    ELSE IF (RIVCLUST=3)
        COMPUTE DIST=634
        END IF
    ELSE IF (STATE=42)
        DO IF (RIVCLUST=1)
        COMPUTE DIST=3524
        ELSE IF (RIVCLUST=2)
        COMPUTE OIST=3286
        ELSE IF (RIVCLUST=3)
        COMPUTE OIST=2842
        END IF
    ELSE IF (STATE=46)
        DO IF (RIVCLUST=1)
        COMPUTE DIST=970
        ELSE IF (RIVCLUST=2)
        COMPUTE DIST=1208
        ELSE IF (RIVCLUST=3)
        COMPUTE DIST=1652
        END IF
    ELSE IF (STATE=49)
        DO IF (RIVCLUST=1)
        COMPUTE DIST=1592
        ELSE IF (RIVCLUST=2)
        COMPUTE DIST=1354
        ELSE IF (RIVCLUST=3)
        COMPUTE OIST=910
        END IF
```

```
ELSE IF (STATE=50)
    DO IF (RIVCLUST=1)
. COMPUTE DIST=4392
. ELSE IF (RIVCLUST=2)
. COMPUTE DIST=4392
. ELSE IF (RIVCLUST=3)
. COMPUTE DIST=4392
. END IF
END IF
END IF
```

The final programming challenge which remained was assigning the appropriate population clusters for the out of state zones of origin. As stated previously, for out of state analysis the U.S. was divided into 6 visitation paths. An algorithm was written which sampled all visitation from within a path and clustered the states and their populations to eliminate any interior blank origin zones from the paths. This is illustrated by the following example.

```
                                    Path out from Montana
MONTANA --m--->> STATE 1 ------> STATE 2 ---m->>> STATE3
If a visit from STATE 3 is observed.
```

Were there visits from STATE 2?


This method has the advantage of not arbitrarily declaring the whole U.S. as the relevant population base, but only populations extending out to the furthest visitation distances. While there is still a degree of arbitrariness in the decision to use states as origin zones, this method is closer to the theoretical ideal of origin zones as a continuous set of concentric circles emanating from each destination. Program segment 8 shows the algorithm used.

## Program Segment 8

## DOCUMENT

```
**************************************************************************
* Below the records are sorted by path and within that by order
* the file is then aggregated by order and a new variable (CLUSTRP)
* is created
SORT CASES BY UNIT2 ORDER
AGGREGATE OUTFILE= * /PRESORTED
    /BREAK=UNIT2 ORDER CLUSPOP DIST
    /CLUSTRP=SUMITRIPS)
XSAVE OUTFILE=AGGIT
    /KEEP=ALL
EXECUTE
GET FILE=AGGIT
DOCUMENT
```

*************************************************************************

* The following program creates a variable (CLUSPOP) for each
* record which originates outside of Montana. The variable *
* CLUSPOP is a path cumulative population. More precisely, *
* it is the relevant clustered population base of a state
* or group of states, considering all visitations from within *
* a particular path to a particular river
********************************************************************
COMPUTE POP11.1=218988
COMPUTE POP46=4315192
COMPUTE POP36 $=6984314$
COMPUTE POP11.2=722172
COMPUTE POP27=1522665
COMPUTE POP4. $1=10468481$
COMPUTE POP4.2=25088445
COMPUTE POP49 $=469557$

```
COMPUTE POP43=4820329
COMPUTE POP30=8842199
COMPUTE POP42=23069773
COMPUTE POP33=652717
COMPUTE POP22=4728687
COMPUTE POP48=12348137
COMPUTE POP12=23775551
COMPUTE POP13=38527881
COMPUTE POP34=49325505
COMPUTE POP4=80698586
COMPUTE POP29=105223760
COMPUTE POP40=690678
COMPUTE POP26=2260503
COMPUTE POP15=9514498
COMPUTE POP35=14853412
COMPUTE POP17=29831518
COMPUTE POP1=53539579
COMPUTE POPG=63286000
COMPUTE POP50=401851
COMPUTE MARK1=0
COMPUTE MARK2=0
COMPUTE MARK3=0
COMPUTE MARK4=0
COMPUTE MARK5=0
DO IF {ORDER=1.1 AND CLUSTRP>0)
    COMPUTE CLUSPOP=POP11.1
    COMPUTE MARK1=1
END IF
DO IF (ORDER=1.2 AND CLUSTRP>0 AND LAG(MARK1,1)=1)
    COMPUTE CLUSPOP=POP46-POP11.1
    COMPUTE MARK1=2
END IF
DO IF {ORDER=1.2 AND CLUSTRP>0 AND LAG(MARK1,1}=0)
    COMPUTE CLUSPOP=POP46
. COMPUTE MARKI=2
END IF
DO IF (ORDER=1.3 AND CLUSTRP>0 AND LAG(MARK1,1)>0)
    IF (LAG(MARK1.1)=21 CLUSPOP=POP36-POP46
    IF (LAG(MARK1,1)=1) CLUSPOP=POP36-POP11.1
END IF
DO IF (ORDER=1.3 AND CLUSTRP>0 AND LAG(MARK1.1)=0)
    COMPUTE CLUSPOP=POP36
ENO IF
DO IF IORDER=2.1 AND CLUSTRP>0}
    COMPUTE CLUSPOP=POP11.2
    COMPUTE MARK2=1
END IF
DO IF (ORDER=2.2 AND CLUSTRP>0 AND LAG(MARK2.1)=1)
    COMPUTE CLUSPOP=POP27-POP11.2
    COMPUTE MARK2=2
END IF
```

```
DO IF (ORDER=2.2 AND CLUSTRP>0 AND LAG(MARK2.1)=0)
    COMPUTE CLUSPOP=POP27
    COMPUTE MARK2=2
END IF
DO IF (ORDER=2.3 AND CLUSTRP>0 AND LAGIMARK2.1)>0)
    IF (LAG(MARK2,1)=2) CLUSPOP=POP4.1-POP27
    IF (LAG(MARK2,1)=1) CLUSPOP=POP4.1-POP11.2
    COMPUTE MARK2=3
END IF
DO IF (ORDER=2.3 AND CLUSTRP>0 AND LAG(MARK2.1)=0)
    COMPUTE CLUSPOP=POP4.1
    COMPUTE MARK2=3
END IF
OO IF (ORDER=2.4 AND CLUSTRP>0 AND LAG(MARK2,1)>0)
    (F (LAG(MARK2,1)=3) CLUSPOP=POP4.2-POP4.1
    IF (LAG(MARK2,1)=2) CLUSPOP=POP4.2-POP27
    IF (LAG(MARK2.1)=1) CLUSPOP=POP4.2-POP11.2
END IF
DO IF (ORDER=2.4 AND CLUSTRP>0 AND LAG(MARK2,1)=0)
    COMPUTE CLUSPOP=POP4.2
END IF
DO IF (ORDER=3.1 AND CLUSTRP>0)
    COMPUTE CLUSPOP=POP49
    COMPUTE MARK3=1
END IF
DO IF IORDER=3.2 AND CLUSTRP>0 AND LAG(MARK3.1)=1)
- COMPUTE CLUSPOP=POP43-POP49
    COMPUTE MARK3=2
END IF
DO IF IORDER=3.2 AND CLUSTRP>0 AND LAG(MARK3.1)=0)
    COMPUTE CLUSPOP=POP43
    COMPUTE MARK3=2
END IF
DO IF (ORDER=3.3 AND CLUSTRP>0 and LAG(MARK3.1)>0)
    IF (LAG(MARK3,1)=2) CLUSPOP=POP30-POP43
    IF (LAG(MARK3,1)=1) CLUSPOP=POP30-POP49
    COMPUTE MARK3=3
END IF
DO IF (ORDER=3.3 AND CLUSTRP>0 AND LAG(MARK3.1)=0)
    COMPUTE CLUSPOP=POP30
    COMPUTE MARK3=3
END IF
DO IF (ORDER=3.4 AND CLUSTRP>0 AND LAG(MARK 3, 1)>0)
    IF (LAG(MARK3.1)=3) CLUSPOP=POP42-POP30
    IF (LAG(MARK3.1)=2) CLUSPOP=POP42-POP43
    IF (LAG(MARK3.1)=1) CLUSPOP=POP42-POP49
END IF
DO IF IORDER=3.4 ANO CLUSTRP>0 AND LAG(MARK3.1)=0)
    COMPUTE CLUSPOP=POP42
END IF
```

DO IF IORDER=4.1 AND CLUSTRP>0)
COMPUTE CLUSPOP=POP33
. COMPUTE MARK4=1
END IF
DO IF (ORDER=4.2 AND CLUSTRP $>0$ AND LAG(MARK4.1)=1)
COMPUTE CLUSPOP=POP22-POP33
. COMPUTE MARK4=2
END IF
DO IF (ORDER=4.2 AND CLUSTRP>0 AND LAG $($ MARK4, 1$)=0$ )
COMPUTE CLUSPOP=POP22
. COMPUTE MARK4=2
END IF
DO IF (ORDER=4.3 AND CLUSTRP >0 AND LAG(MARK4,1)>0)
IF (LAG(MARK4,1)=2) CLUSPOP=POP48-POP22
IF (LAG(MARK4.1)=1) CLUSPOP=POP48-POP33
COMPUTE MARK4=3
END IF DO IF (ORDER=4.3 AND CLUSTRP>0 AND LAG(MARK4,1)=0)

COMPUTE CLUSPOP=POP48
COMPUTE MARK4=3
END IF DO IF $($ ORDER $=4.4$ AND CLUSTRP $>0$ AND LAG $($ MARK4, 1$)>0)$

IF (LAG(MARK4,1)=3) CLUSPOP=POP12-POP48
IF (LAG(MARK4,1)=2) CLUSPOP=POP12-POP22
IF (LAG(MARK4.1)=1) CLUSPOP=POP12-POP33
COMPUTE MARK4=4
END IF
DO IF (ORDER=4.4 AND CLUSTRP>0 AND LAG(MARK4,1)=0)
COMPUTE CLUSPOP=POP12
COMPUTE MARK4=4
END iF
DO IF (ORDER=4.5 AND CLUSTRP>0 AND LAG(MARK4, 1)>0)
IF (LAG(MARK4,1)=4) CLUSPOP=POP13-POP12
IF (LAG(MARK4.1)=3) CLUSPOP=POP13-POP48
IF (LAGIMARK4,1)=2) CLUSPOP=POP13-POP22
IF (LAG(MARK4.1)=1) CLUSPOP=POP13-POP33
COMPUTE MARK4=5
END IF
DO IF (ORDER=4.5 AND CLUSTRP>0 AND LAG(MARK4,1)=0)
COMPUTE CLUSPOP=POP13
COMPUTE MARK4=5
END IF OO IF IORDER=4.6 AND CLUSTRP >0 AND LAG(MARK4.1)>0)

IF (LAG(MARK4,1)=5) CLUSPOP=POP34-POP13
IF (LAG(MARK 4,1$)=4)$ CLUSPOP $=$ POP $34-$ POP 12
IF (LAG(MARK4.1)=3) CLUSPOP $=P O P 34-P O P 48$
IF (LAG(MARK4,1)=2) CLUSPOP=POP34-POP22
IF (LAGIMARK4.1)=1) CLUSPOP=POP34-POP33
COMPUTE MARK $4=6$
END : $F$
DO IF (ORDER=4.6 AND CLUSTRP $>0$ AND LAG(MARK4. 1$)=0$ ) COMPUTE CLUSPOP=POP34

```
    COMPUTE MARK4=6
END IF
DO IF (ORDER=4.7 AND CLUSTRP>0 AND LAG(MARK4,1)>0)
    IF (LAG(MARK4,1)=6) CLUSPOP=POP47-POP34
    IF (LAG(MARK4,1)=5) CLUSPOP=POP47-POP13
    IF (LAG(MARK4,1)=4) CLUSPOP=POP47-POP12
    IF (LAG(MARK4,1)=3) CLUSPOP=POP47-POP48
    IF (LAG(MARK4,1)=2) CLUSPOP=POP47-POP22
    IF (LAG(MARK4,1)=1) CLUSPOP=POP47-POP33
    COMPUTE MARK4=7
END IF
DO IF {ORDER=4.7 AND CLUSTRP>0 AND LAG(MARK4,1)=0)
    COMPUTE CLUSPOP=POP4
    COMPUTE MARK4=7
END IF
DO IF {ORDER=4.8 AND CLUSTRP >0 AND LAG(MARK4,1)>0)
    IF (LAG(MARK4,1)=7) CLUSPOP=POP29-POP47
    IF (LAG(MARK4.1)=6) CLUSPOP=POP29-POP34
    IF (LAG{MARK4,1)=5) CLUSPOP=POP29-POP13
    IF (LAG(MARK4,1)=4) CLUSPOP=POP29-POP12
    IF (LAG(MARK4,1)=3) CLUSPOP=POP29-POP48
    IF (LAG(MARK4,1)=2) CLUSPOP=POP29-POP22
    IF (LAG(MARK4,1)=1) CLUSPOP=POP29-POP33
END IF
DO IF (ORDER=4.8 AND CLUSTRP>0 AND LAG(MARK4,1)=0)
    COMPUTE CLUSPOP=POP29
END IF
DO IF (ORDER=5.1 AND CLUSTRP>0)
    COMPUTE CLUSPOP=POP40
    COMPUTE MARK5=1
END IF
DO IF (ORDER=5.2 AND CLUSTRP>0 AND LAG(MARK5,1)=1)
    COMPUTE CLUSPOP=POP26-POP40
    COMPUTE MARK5=2
END IF
DO IF (ORDER=5.2 AND CLUSTRP>0 AND LAG(MARK5.1)=0)
    COMPUTE CLUSPOP=POP26
    COMPUTE MARK5=2
END IF
DO IF (ORDER=5.3 AND CLUSTRP>0 AND LAG{MARK5.1)>0)
    IF (LAG(MARK5,1)=2) CLUSPOP=POP15-POP26
    IF (LAGIMARK5.1)=1) CLUSPOP=POP15-POP40
    COMPUTE MARK5=3
END IF
DO IF (ORDER=5.3 AND CLUSTRP>0 ANO LAG(MARK5.1)=0K
    COMPUTE CLUSPOP=POP15
    COMPUTE MARK5=3
END IF
OO IF (ORDER=5.4 AND CLUSTRP>0 AND LAG(MARK5.1)>0)
    IF (LAG(MARK5,1)=3) CLUSPOP=POP35-POP15
    IF (LAG(MARK5.1)=2) CLUSPOP=POP35-POP26
```

```
. IF (LAG(MARK5,1)=1) CLUSPOP=POP35-POP40
    COMPUTE MARK5=4
END IF
DO IF (ORDER=5.4 AND CLUSTRP >0 AND LAG(MARK5,1)=0)
    COMPUTE CLUSPOP=POP35
    COMPUTE MARK5=4
END IF
DO IF (ORDER=5.5 AND CLUSTRP>0 AND LAG(MARK5,1)>0)
    IF (LAG(MARK5,1)=4) CLUSPOP=POP17-POP35
    IF (LAG(MARK5,1)=3) CLUSPOP=POP17-POP15
    IF (LAG(MARK5,1)=2) CLUSPOP=POP17-POP26
    IF (LAG(MARK5,1)=1) CLUSPOP=POP17-POP40
    COMPUTE MARK5=5
END IF
DO IF (ORDER=5.5 AND CLUSTRP>0 AND LAG(MARK5,1)=0)
    COMPUTE CLUSPOP=POP17
    COMPUTE MARK5=5
END IF
DO IF (ORDER=5.6 AND CLUSTRP>0 AND LAG(MARK5,1)>0)
    IF (LAG(MARK5,1)=5) CLUSPOP=POP1-POP17
    IF (LAG(MARK5.1)=4) CLUSPOP=POP1-POP35
    IF (LAG(MARK5,1)=3) CLUSPOP=POP1-POP15
    IF (LAG(MARK5,1)=2) CLUSPOP=POP1-POP26
    IF (LAG(MARK5,1)=1) CLUSPOP=POP1-POP40
    COMPUTE MARK5=6
END IF
DO IF (ORDER=5.6 AND CLUSTRP >0 AND LAG(MARK5,1)=0)
    COMPUTE CLUSPOP=POP1
    COMPUTE MARK5=6
END IF
DO IF (ORDER=5.7 AND CLUSTRP>0 AND LAG(MARK5,1)>0)
    IF (LAG(MARK5,1)=6) CLUSPOP=POPG-POP1
    IF (LAG(MARK5,1)=5) CLUSPOP=POP9-POP17
    IF (LAG(MARK5.1)=4) CLUSPOP=POPG-POP35
    IF (LAG(MARK5.1)=3) CLUSPOP=POPG-POP15
    IF (LAG(MARK5,1)=2) CLUSPOP=POPG-POP26
    IF (LAG(MARK5,1)=1) CLUSPOP=POPG-POP40
END IF
DO IF (ORDER=5.7 AND CLUSTRP >0 AND LAG(MARK5,1)=0)
    COMPUTE CLUSPOP=POPG
END IF
DO IF {ORDER=6.1 AND CLUSTRP>0}
    COMPUTE CLUSPOP=POP50
END IF
XSAVE OUTFILE= PENULT
    /KEEP=UNIT2 ORDER CLUSTRP CLUSPOP DIST
EXECUTE
GET FILE=PENULT
PRINT / UNIT2 ORDER DIST CLUSTRP * CLUSPOPIF11I
EXECUTE
FINISH
```

The result from the aforementioned procedures was a data set which contained the following variables for each pairing of an in-state or out-of-state origin zone, with each of the 19 unique waters.

UNIT2 ORDER CLUSTRP CLUSPOP DIST
In addition to these variables, which are necessary for the specification of a regional travel cost model, the various variables relating to flow were also included in this data file.

## Appendix B <br> Montana DFWP Survey Questionnaire.

FISHERIES SURVEY
PLEASE CNECK THE TVPE OF LICENSE YOU PURCHASED
$\square$ FISHMO OR FISHINGICONSERVATIOM COMBINATIOM FOR DESIDENT OR MONAESIDENT
$\square$ SPOATSMAMdisabled resident comseayation
aestoent conservation only
(F YOU ARE A PIONEEN (E2 ON OLOER) OR A YOUTH (12 TO 14), DO YOU PLAM ON USMG YOUR CONSERVATIOM LICENSE TO FISMT
$\square$ res $\square$ no

PLEASE REFER TO THE MAPS TO HELP US IOENTIFY THE WATERS YOU FISHED

| DATE FISMED | NAHE OF LAKE Of STAEAM FISHEO | SECTION MUMEER TF indicated ON MAP | MEAREST TOWH AND/OR POANT OF ACCES\$ OR LANDMARK | total <br> HOURS <br> FISHED <br> PER OAY | TOTAL NUMBER OF FISH CAUGHT |  | TOTAL MAMBER OF FISH KEPT |  | $\left\{\begin{array}{l} \text { WAS THE MAIM } \\ \text { PURPOSE OF } \\ \text { YOUR TRIF TO } \\ \text { FISH? (Y OA N) } \end{array}\right.$ | DID YOU STAY <br> Overnight (YOM M) | ROUNO - RIP DISTANCE travelec |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| auaust $18-31$ |  |  |  |  | TAOUT <br> AND <br> SALMON | OTHEA SPORT Fish | TROUT <br> ANO <br> SALMON | OTHER SPORT F'GH |  |  |  |

ENTER EACH DAY AND EACH WATER FISHED ON A SEPARATE LINE LIST ALL FISHING IN MONTANA, NOT JUST WATERS INDICATED OR THE MAP


- SUCH AS: THE MMABEROF WHITEFISH, WALLEYE, PETCH, BASS, ETC.
- FOU STAYEC OVERNIGHT, PLEASE IFAKE A SEPAGATE ENTAY FOA EACH FISHING TRLP.

THIS tNFORMATION WRL BE HELD IN STRICT CONFDENCE AND WILZ QE USED FOA MAMAGEMENT PUAPOSES ONLY

# Appendix C <br> Residual Analysis of Regional Model Regression Results 

## C.1. Variable Plots

Plots of dependent and independent variables provide a good preliminary test of the appropriateness of the functional form of a regression model. In the case of the regional travel cost model specified in equation 2, chapter 5 the appropriate variable plots are $\ln (T R I P S)$ against $\ln (D I S T)$ and $\ln (T R I P S)$ against $\ln (Y R F L O W)$. These plots are shown in figure $\mathrm{C}-1$ and figure $\mathrm{C}-2$. The plot of $\operatorname{In}($ TRIPS $)$ against $\ln (D I S T)$ shows that the double log specification is a good specification of the tripdistance relationship. The $\operatorname{In}(T R I P S)$ against $\operatorname{In}$ (YRFLOW) plot, whife not nearly as good of a fit, is nevertheless not suggestive of any alternative specification which would improve the fit.

## C.2. Analysis of Residuals

A good test of the degree to which an estimated model obeys the assumptions of the ordinary least squares regression model is the analysis of the residuals from that modeling procedure. The plots shown in figure $\mathrm{C}-3$ and figure C-4 are illustrative of the validity of the assumptions of normality and constant variance of the regression residuals.

Figure $\mathrm{C}-3$ shows that the assumption of normality of the regression residuals is well satisfied by this specification of the model. Figure $\mathbf{C - 4}$, however,


Figure $C-1$ : Plot of $\ln$ (TRIPS) and $\ln (D I S T)$


Figure $\mathrm{C}-2$ : Plot of $\operatorname{In}($ TRIPS $)$ and $\operatorname{In}($ YRFLOW $)$
shows that there is a violation of the constant variance assumption on the part of the price variable $\ln (D I S T)$. The shifter variable $\operatorname{In}$ (YRFLOW) seems to satisfy the constant variance assumption.

Histogram *. Standarjized Residual


Figure C-3: Histogram of Standardized Residuals


Figure C-4: Residual Scatterplots of independent Variables

## References

Boyle, K.J. M.P. Welsh, R.P. Bishop ans R.M. Baumgartner. (1987). Analysing the Effects of Glen Canyon Dam Releases on Colorado River Recreation Using Scenarios of Unexperienced Flow Conditions.

Brown, W.G. and F. Nawas. (May 1973). Impact of Aggregation on Estimation of Recreational Demand Functions. American Journal of Agricultural Economics., (55), pp. 246-49.

Burt, O. and D. Brewer. (1971). Estimation of Net Social Benefits for Outdoor Recreation. Econometrica. (39), pp. 813-27.

Cesario, F. (1976). The Value of Time in Recreation Benefit Studies. Land Economics, (52(2)), pp. 32-41.

Ciccetti, C.J., A.C.Fisher, and V.K. Smith. (1976). An Econometric Evaluation of a Generalized Consumer Surplus Measure: The Mineral King Controversy. Econometrica, (44), pp. 1259-1276.

Clawson, Marion and Jack Knetsch. (1966). The Economics of Outdoor Recreation. Baltimore MD:: Johns Hopkins University Press.

Daubert, J.T., R.A. Young and S.L. Gray. (1979). Economic Benefits from Instream Flow in a Colorado Mountain Stream. (Tech. Rep.). U.S. Dept. of Interior Wash. D.C.., A report to the Office of Water Reaserch and Technology.

Daubert, J.,T. and R.A. Young. (1981). Recreational Demands for Maintaining Instream Flows: A Contingent Valuation Approach. American Journal of Ag. Econ., (63), pp. 666-676.

Duffield, John. (1987a). Heteroscedasticity and Functional Form in a Regional Travel Cost Demand Model of Montana Cold Water Stream Fishing. (Tech. Rep.). Helena: Montana Department of Fish, Wildlife and Parks,

Duffield, John. (1987b). Sensitivity Analysis of Montana Cold Water Stream Benefit Estimates. (Tech. Rep.). Helena: Montana Department of Fish, Wildlife and Parks.,

Duffield, J., and S. Allen. (1987). Contingent Valuation of Montana Trout Fishing by River and Angler Subgroup. (Tech. Rep.). Montana Department of Fish Wildlife and Parks,

Duffield, John, John Loomis and Rob Brooks. (1987). The Net Economic Value of Cold Water Fishing in Montana: A Regional Travel Cost Model. (Tech. Rep.). Helena: Montana DFWP,

Dwyer, John, John Kelly, and Michael Bowes. (1977). Improved Procedures for Valuation of the Contribution of Recreation to National Economic Development (Tech. Rep. 128). Water Resources Center, University of Illinois, Gum, R. and W.E. Martin. (November 1975). Problems and Solutions in Estimating the Demand for the Value of Rural Outdoor Recreation. American Journal of Agricultural Economics, (57), pp. 558-66.

Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lutkepohl and T.-C. Lee. (1980). The Theory and Practice of Econometrics.. John Wiley and Sons.

Just, R., D. Hueth, and A. Schmitz. (1982). Applied Welfare Economics and Public Policy.. New Jersey.: Prentice Hall.

Knetsch, Jack L. (1963). Outdoor Recreation Demands and Benefits. Land Economics, (37), pp. 387-96.

Koutsoyiannis, A. (1978). Theory of Econometrics.. Barnes and Noble.
Loomis, John B. (1987). The Economic Value of Instream Flow: Methodology and Benefit Estimates for Optimum Flow. Journal of Envirionmental Economics and Management, Vol. 24.

Menz, F. and D. Wilton. (1983). Alternative Ways to Measure Recreational Values by the Travel Cost Method. American Journal of Agricultural Economics, (65(2)).

Narayanan, Rangesan. (1986). Evaluation or Recreational Benefits of Instream Flows. Journal of Leisure Research, 18(No. 2), 116-128.

Sorg, C., and J. Loomis. (1985). An Introduction to Wildife Valuation Techniques. Wildlife Society Bulletin, 13, 38-46.

Vaughan, W.J. and Clifford Russell. (1982). Valuing a Fishing Day: An Application of a Systematic Varying Parameter Model. Land Economics. 58(4), 450-463.

Walsh, R., L. Sanders and J. Loomis. (1985). Wild and Scenic River Economics: Recreation use and Preservation Values.. Denver, CO.: American Wilderness Alliance.

Weisbrod, B.A. (1964). Collective Consumption Services of Individual Consumption Goods. Quarterly Journal of Economics, 77, 71-77.
U.S.Water Resources Council. (December 1979). Procedures for Evaluation of National Economic Development (NED) Benefits and Costs in Water Resource Planning. Final Rule. Federal Register, Vol. vol 44(No. 242).

## U.S. Water Resources Council. (1983). Economic and Environmental Principles and Guidelines for Water and Related Land Resources. .


[^0]:    ${ }^{1}$ See Weisbrod (1964) and Krutilla (1967) for an excellent discussion of non-use values attached to natural resources.

[^1]:    ${ }^{2}$ See R. Walsh et al., Wild and scenic river economics: recreation use and preservation values, 1985

[^2]:    ${ }^{3}$ See Duffield, Loomis and Brooks."The Net Economic Value of Fishing in Montana".Montana DFWP. 1987 and Duffield and Allen (1987).

[^3]:    ${ }^{4}$ Owyer, Kelly, and Bowes (1977) provide an exhaustive analysis of the travel cost method and its application.

[^4]:    ${ }^{5}$ For an excellent discussion of alternative regional model specifications and the rational behind them see Sorg and Loomis (1985).

[^5]:    ${ }^{6}$ For a more rigorous discussion of the derivation of standard variable and reported costs see Duffield, Loomis and Brooks (1987) Appendix B.

[^6]:    ${ }^{7}$ It must be noted that Daubert and Young's study was conducted in a drought year in which the maximum observed flows on the Poudre river were onlv a fourth of normal maximum flows.

[^7]:    ${ }^{8}$ This problem with the 1984 data was discussed with Bob MacFarland, statistician for the DFWp. His recommendation was that the 84 estimates were downwardly biased and should be excluded from the study.

[^8]:    ${ }^{\prime}(1)=$ FREE FLOWING; (2)=SLIGHT DAM CONTROL; (3)=HEAVY CONTROL

[^9]:    ${ }^{1}$ Average number of trips predicted for each origin/destination pair.
    ${ }^{2}$ Standard error of the predictions.

[^10]:    IF (ORDER=83.01) CLUSPOP=151700 if (ORDER=83.2) CLUSPOP=155100 IF (ORDER=83.3) $\mathrm{CLUSPOP}=76600$ IF (ORDER=83.4) CLUSPOP=16600 IF (ORDER=83.5) CLUSPOP $=21600$ IF (ORDER=83.6) CLUSPOP $=134900$ IF (ORDER=83.7) CLUSPOP=21300 IF (ORDER=83.8) CLUSPOP=89000 IF (ORDER=83.9) $C L U S P O P=46400$ IF (ORDER=83.10) CLUSPOP=6900 IF (ORDER=83.11) CLUSPOP $=27600$ IF (ORDER=83.12) $C L U S P O P=76500$

    ```
    IF (ORDER=84.1) CLUSPOP=102200
    IF (ORDER=84.2) CLUSPOP=234100
    IF (ORDER=84.3) CLUSPOP=55400
    IF (ORDER=84.4) CLUSPOP=133300
    IF (ORDER=84.5) CLUSPOP=3300
    IF (ORDER=84.6) CLUSPOP=75400
    IF (ORDER=84.7) CLUSPOP=220500
    ```

    IF (ORDER=85.01) CLUSPOP=102200
    IF (ORDER $=85.2$ ) $C L U S P O P=163500$
    IF (OROER=85.3) CLUSPOP $=21600$
    IF (ORDER=85.4) CLUSPOP=13200
    IF (ORDER=85.5) CLUSPOP=1000
    IF (ORDER=85.6) CLUSPOP=11500
    IF (ORDER=85.7) $\mathrm{CLUSPOP}=118700$
    IF (ORDER=85.8) $\mathrm{CLUSPOP}=8600$
    IF (ORDER=85.9) CLUSPOP=6000
    IF (ORDER $=85.10$ ) $\mathrm{CLUSPOP}=75400$
    IF (ORDER=85.11) CLUSPOP=302300
    IF (ORDER $=86.01$ ) CLUSPOP $=31600$
    IF (ORDER=86.2) CLUSPOP=74300
    If (ORDER $=86.3$ ) $\mathrm{CLUSPOP}=506400$
    IF (ORDER=86.4) CLUSPOP $=8700$
    IF (ORDER $=86.5$ ) $\mathrm{CLUSPOP}=35200$
    IF (ORDER=86.6) CL USPOP $=11200$
    IF (ORDER=86.7) $\mathrm{CLUSPOP}=27600$
    IF (ORDER $=86.8$ ) $\mathrm{CLUSPOP}=6900$
    IF (ORDER $=86.9$ ) $C L U S P O P=45800$
    IF (ORDER $=86.10$ ) CL USPOP $=76500$
    IF (ORDER $=87.01$ ) CL USPOP $=18700$
    IF (OROER $=87.2$ ) CLUSPOP $=53900$
    IF (ORDER=87.31 CLUSPOP $=234100$
    IF (OROER=87.4) CLUSPOP $=192000$
    IF (ORDER $=87.5$ ) CLUSPOP $=187900$
    IF (ORDER=87.6) CLUSPOP $=27600$
    IF (ORDER=87.7) CLUSPOP $=76500$
    IF (ORDER=87.8) CLUSPOP=3700

