IMPACTS OF CLIMATE CHANGE ON NATIONAL PARK VISITATION AND LOCAL ECONOMIES IN THE US SOUTHWEST

by

Xudong Ma

A Thesis Submitted to the Faculty of the

DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

2009

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permissions for extended quotation form or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interest of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED:________________

APPROVAL BY THESIS DIRECTOR

This Thesis has been approved on the dates shown below

Professor George Frisvold

Agricultural & Resource Economics

Date

ACKNOWLEDGEMENTS

I would like to thank my advisor and mentor, Dr. George Frisvold for his constructive guidance throughout this study. I am truly grateful for his advice and constant encouragement that helped me complete this thesis.

I would also like to express gratitude to my committee members, Dr. Dennis Cory and Dr. Gary Thompson for providing me invaluable insight into different aspects of this study.

I would especially like to thank Dr. Frisvold for generously providing me with financial support and the data used in this thesis. This work was part of Climate Assessment for the Southwest (CLIMAS) program, supported by National Oceanic and Atmospheric Administration (NOAA). I would also like to thank the Department of Agricultural and Resource Economics for funding part of this project.

Finally, this thesis is dedicated to my parents, who taught me to be a good learner and support me always.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This study examines how climate affects recreational visits to national parks in the Southwest and how climate-induced changes in park visits affect spending, employment and personal incomes in local communities. A fixed effects vector decomposition (FEVD) regression method is used to estimate how climate affects annual visitation, controlling for other factors (such as park attributes, population, and gasoline prices). The study uses 25 years of visitation data for 42 sites in the national park system. The study combines the regression results with input-output modeling to examine how climate change affects local economies near parks.

Results suggest that warmer July temperatures discourage visits. Warmer January temperatures increase visits, except in low desert parks with relatively high temperatures. The net effect of projected temperature increases is an increase in visits in Northern Arizona, Northern New Mexico, and Southern Utah, and a decrease in visits in the southern parts of Arizona, New Mexico, and California. As previous studies of climate change and outdoor recreation have found, the net economic benefits of warmer temperatures are estimated to be positive. However, results suggest that lower lake elevations and large fires (such as the Cerro Grande fire) can have significant negative impacts on visits. Thus, a complete assessment of climate change impacts on outdoor recreation would also account for changes in reservoir levels and the probability of large wildfire events.

CHAPTER ONE: INTRODUCTION

Climate change has been proved to be one of the significant factors related to national park visitation (Loomis and Crespi, 1999; Mendelsohn and Markowski, 1999). Since Industry Revolution in 19th Century, global climate changes with a much faster pace. Various industries contribute to the faster change of global climate, chemistry industry, steel industry, etc. Some of the most notable changes in global environment include global warming, water pollution, desertification, etc. As the environment changes, interactively it influences people's lifestyle. However, it is often difficult to examine how the climate change influences people's life directly. One reason is that climate change takes place in a glacial pace compared to people's daily life. Besides, climate change is often in a macro-level context while people's daily life is usually in a micro-level context so that the connection is hard to observe. National park visitation, however, provides a convenient way to evaluate how the climate change influences people's lifestyle. How temperature change is related to park visitation? How lake elevation change is related to park visitation? Evaluation on these relationships can give us some insights into the impacts of climate change on people's life. This paper reveals how the change of some critical environmental variables, such as temperature, humidity, etc, can influence park visitation, which may provide some useful insights into the impacts of climate change on human beings.

National Park Service (NPS) has a 16-category typology of public lands designations (Weiler and Seidl, 2004). This paper includes many national-level parks in Southwest, especially inside and around Arizona. To simplify, national park is used in the following descriptions to represent all kinds of national-level parks. The National Park designated by NPS in the 16-category typology is expressed as National Park NPS Designated instead.

National parks in the U.S. Southwest attract more than 26 million visitors per year while water-based recreation accounts for over 40% of visits. Visitors spend more than $1.3 billion in gateway communities that surround parks and such spending is an important source of jobs and income in many rural communities of the Southwest. Good knowledge of how climate change may influence visits also sheds light on how local economy and employment would be influenced. With appropriate input-output analysis, a connection between climate change and local employment can be set up, which reveals how closely the environment is related with our daily life and helps people realize the importance of environment protection. The result can also provide useful information to local government and the gateway communities to the park.

Time-series data can reflect visitation trend continuously and is usually applied for national park visitation analysis (Johnson and Suits, 1983; Wicks, Uysal and Kim, 1994; Weiler and Seidl, 2003; Weiler, 2006; Pergams and Zaradic, 2006). Some paper chooses several points in time-series and applies the specific-point data for analysis (Hanink and White, 1997; Hanink and Stutts, 2002). This method is also valid according to different topics and focuses. However, climate change is usually based on a temporal dimension and time-series data can better describe the visitation trend induced by climate change. Therefore, time-series data is used in our paper.

Sometimes time trend biases the results and makes the interpretation too ambiguous to convey valid information. However, with cross-section data, time trend would not be a factor (Johnson and Suits, 1983). Thus it is generally suggested that cross-section data be used when the data for individual parks is available. In our analysis, all the national parks are inside and around Arizona and the visitation data can be found on NPS website. In a word, typical time-series cross-section data is used in our study. Then, the following sections of the paper are organized as follows:

Many literatures talk about the analysis of time-series cross-section data. Which method should be regarded as best and adopted in this paper? Therefore, a brief literature review is set up to evaluate different methods in different cases. In order to determine the optimal method for this paper, at first we need describe the data. The visitation to a national park can be influenced by variables of three categories: demand factors, park attributes and climate/environmental variables. We explain about the variable selection to set up the model and then, based on the data, the optimal method can be specified. After applying the method to the data, this paper analyzes the empirical results. To testify our results and show the advantages of the new technique, the results are compared with those from other analysis methods. To study the relationship between climate change and park visitation, sensitivity analysis and visits decomposition method are applied. Some important findings are noted in that section as well as in the conclusion. With our results, we also carry out input-output analysis to evaluate how the climate change would influence local economy and employment. At last we summarize our findings in the research and we point out the significance as well as shortages for future studies.

CHAPTER TWO: LITERATURE REVIEW OF TIME-SERIES CROSS-SECTION (TSCS) DATA ANALYSIS

As Beck (2001) pointed out, "TSCS data are characterized by repeated observations (often annual) on the same fixed political units (usually states or countries)". In this paper, TSCS data are characterized by repeated annual observations on the same national park unit. Because of the repetition, one of TSCS data's advantages is they can considerably expand the number of observations and thus are statistically important for econometric analysis and results. For example, data in this paper are from 42 national parks inside or around Arizona, from 1997 to 2003. In the angle of individual units, only 42 observations are available while in the angle of time-series data, only 25 specific time points exist. However, when the data is organized as TSCS data, 1,050 observations are available for analysis. The considerable expansion of number of observations, on the other hand, incurs some problems. Firstly, are TSCS data the same as panel data? If the answer is no, methods applied to panel data cannot be simply taken over for TSCS data analysis. Secondly, among the repeated observations, TSCS data typically have variables with constant values during the research period. For example, in our data, area for different parks stays the same during the 25 years. Thus how to deal with the time-invariant variables? What is the optimal method to deal with the fixed effects in a model? Last but not least, there are some problems related to error term. Different from Gauss-Markov assumption, error term $\in_{i,t}$ (i for different units, t for different time points) in TSCS data may have the following features: (i) panel heteroskedasticity, (ii) contemporaneous correlation of the errors and (iii) serially correlated errors (Beck, 2001). Thus, what is the optimal method to deal with the error term?

Admittedly, TSCS data and panel data are similar in many ways. For example, the two types of data are both featured with observations from time points and different units. However, they have some significant differences. Beck (2001) pointed out the differences between TSCS data and panel data: “Panel data are repeated cross-section data, but the units are sampled (usually they are survey respondents obtained in some random sampling scheme), and they are typically observed only a few times. TSCS units are fixed; there is no sampling scheme for the units, and any “resampling” experiments must keep the units fixed and only resample complete units (Freedman & Peters, 1984). In panel data, the people observed are of no interest; all inferences of interest concern the underlying population that was sampled, rather than being conditional on the observed sample. TSCS data are exactly the opposite; all inferences of interest are conditional on the observed units.” In a summary, there are three essential differences between TSCS data and panel data. First of all, the individual units in TSCS data are fixed while the units in panel data is sampled, namely, can be randomly selected or asymptotically randomly selected with Heckman correction. The next, TSCS data usually use continuous time-series data. Number of specific units N can be larger than, fewer than or equal to the number of time points T. In panel data, much fewer observations in time points exist for each individual unit. Researchers usually select a few time points for analysis, 3, 5 points for example, while these time points are usually scattered (Hanink and White, 1997; Hanink and Stutts, 2002). The last difference between TSCS data and panel data is the different purpose of inferences. TSCS models attempt to draw conditional inferences on the observed unit in a temporal dimension, usually conveying insights into the future time points. In addition, they can also shed light on common rules across sections. Models dealing with panel data, on the other hand, attempt to expand the conclusions to other unobserved units in a spatial dimension only while time issue is not the focus. Because of these differences, methods for panel data analysis cannot be simply applied in TSCS data analysis. Thus, TSCS data need adjusted method or new method for analysis.

Generalized least squares (GLS) method described by Parks (1967) is designed to deal with some common problems that occur in TSCS data (Beck and Katz, 1995). In practice, feasible generalized least squares (FGLS) method is usually in use because the true covariance matrix （Σ） of the errors is often unknown. However, “the FGLS approach can do considerable harm with TSCS data; this is because it is possible to estimate some error properties of TSCS data that cannot be estimated with either a single time series or cross section.” (Beck, 2001). Ordinary Least Squares with Panel Corrected Standard Errors method presented by Beck and Katz in 1995 solves some problems in FGLS and improves the model performance. However, it has nothing to do with the fixed effects and time-invariant variables in the model, while they are critical issues in TSCS data analysis. “Hsaio shows that fixed effects are appropriate if one wants to make inferences to the observed units, whereas the random effects model (which assumes that the effects are drawn from some distribution) is appropriate if one thinks of the observed units as a sample from a larger population and if one wants to make inferences about the larger population”. (Beck, 2001). “The Random Coefficients Model (RCM) is an interesting compromise between assuming complete homogeneity and assuming complete heterogeneity.” However, “the RCM is seldom used in TSCS applications. One reason is that the RCM estimates of the β will be similar to the OLS estimates, in that OLS is still consistent, albeit inefficient if the coefficients really are random. Given the large sample sizes of typical TSCS data, this inefficiency may not be important. The OLS standard errors will of course be wrong.” (Beck, 2001). Obviously, the fixed effects model is more appropriate to address TSCS data.

The simplest way to allow for heterogeneity is to assume that each unit has its own intercept. Such a model is called the fixed effects model (FE) (Beck and Katz, 2004). Yet the use of fixed effects model comes with costs too. “Fixed effects are clearly collinear with any independent variables that are unchanging attributes of the units, so they force us to drop such unchanging variables from the specification.” (Beck, 2001). However, some time-invariant variables are especially interesting to researchers, like the relationship between national park area and national park visitation, or whether a country with harbors or not is related to its economic growth, etc. We not only want to keep those variables, but also would like to estimate their effect in econometric models. So how to estimate the coefficients for time-invariant variables is a both interesting and challenging topic for researchers. Beck and Katz (2004) presented a model called Fixed Effects with Lagged Dependent Variables (FELDV) in the following, attempting to address such issues in TSCS data analysis:

$$y_{i,t} = \phi y_{i,t-1} + \beta x_{i,t} + \alpha_i + \varepsilon_{i,t} \tag{1}$$

They claimed a natural first approach to estimate this equation is to including a series of dummy variables for each unit in the pooled model, leading to the Least Squares Dummy Variable Estimator (LSDV) (Beck and Katz, 2004). They compared the LSDV with other popular estimators used in TSCS data, like Anderson-Hsiao estimator as well as Kiviet estimator. Their Monte Carlo analysis shows that for typical TSCS data, the FELDV performs about as well as the much more complicated Kiviet estimator, and better than the Anderson-Hsiao estimator (both designed for panel data) (Beck and Katz, 2004). Thus, they suggest the method of OLS with fixed effects and a lagged dependent variable be used in TSCS data analysis. However, there are problems about the validity of the lagged dependent variable (LDV) as an explanatory variable. Generally speaking, the inclusion of LDV has the following three faults: (i) we explain current y by lagged y, which generally is not the purpose of regression analysis; (ii) LDV causes logic confusion if we want to trace back to model lags; (iii) LDV is likely to cause perplexity when interpreting coefficient of LDV and coefficients of other explanatory variables, which is a serious econometric problem (Beck and Katz, 2004). Although Beck and Katz argued that the inclusion of LDV has helped a lot to improve the model performance, because of the potential misguiding interpreting problems, a better approach should be explored and proposed without LDV.

Another reason that FELDV model is not good enough for TSCS data analysis is because a general problem in fixed effects model. As *Plümper* and Troeger (2004) pointed out, "One danger of fixed effects models is that many researchers believe that the inclusion of unit dummies precludes problems with omitted variables." However, "on the one hand, unit dummies do not necessarily eliminate omitted variable bias; on the other hand, fixed effects models cannot estimate the coefficients of theoretically interesting time-invariant variables". To fix the problems in fixed effects model, *Plümper* and Troeger (2004) presented a systematic approach called Fixed Effects Vector Decomposition technique (FEVD) in TSCS data analysis. This technique consists of three stages: "the first stage runs a fixed-effects model without time-invariant variables, the second stage decomposes the unit-effects vector into a part explained by the time-invariant variables and an error term, and the third stage re-estimates the first stage by pooled-OLS including the time invariant variables plus the error term of stage 2." (*Plümper* and Troeger, 2004). The following is a clear mathematical map of the three steps (*Plümper* and Troeger, 2007).

The original function of the pooled model: $Y_{it} = \alpha + \sum_{k=1}^{K} \beta_k X_{kit} + \sum_{m=1}^{M} \gamma_m Z_{mi} + \mu_i + \varepsilon_{it}$ (2)

In this model, Y is the dependent variable; α is the intercept; X represents the time-varying explanatory variable while β is the coefficient; Z represents the time-invariant variable while γ is the coefficient; i represents the observed specific units; t is the observed time points; k is the kth time-varying variable; m is the mth time-invariant variable; μ represents the fixed effect for a specific unit i; ε is the error term for each pooled TSCS observation.

In the first step, we average function (1) over the observed period T so that we get the following function $\bar{Y}_i = \alpha + \sum_{k=1}^{K} \beta_k \bar{X}_{ki} + \sum_{m=1}^{M} \gamma_m Z_{mi} + \mu_i + \bar{\varepsilon}_i$ (3)

where $\bar{Y}_i = \frac{1}{T} \sum_{t=1}^{T} Y_{it}$, $\bar{X}_i = \frac{1}{T} \sum_{t=1}^{T} X_{it}$ and $\bar{\varepsilon}_i = \frac{1}{T} \sum_{t=1}^{T} \varepsilon_{it}$

Subtracting (2) from (1) and we get

$$Y_{it} - \overline{Y}_i = \alpha - \alpha + \sum_{k=1}^{K} \beta_k (X_{kit} - \overline{X}_{ki}) + \sum_{m=1}^{M} \gamma_m (Z_{mi} - Z_{mi}) + \mu_i - \mu_i + \varepsilon_{it} - \overline{\varepsilon}_i$$

$$\equiv \ddot{Y}_{it} = \sum_{k=1}^{K} \beta_k \ddot{X}_{kit} + \ddot{\varepsilon}_{it} \qquad (4)$$

In function (3), $\ddot{Y}_{it} = Y_{it} - \overline{Y}_i$, $\ddot{X}_{kit} = X_{kit} - \overline{X}_{ki}$ and $\ddot{\varepsilon}_{it} = \varepsilon_{it} - \overline{\varepsilon}_i$, which are demeaned variables of the fixed effects transformation. By regressing $\ddot{Y}$ on $\ddot{X}$, we can get $\hat{\beta}_k^{FE}$. Then we calculate estimated fixed effects: $\hat{\mu}_i = \overline{Y}_i - \sum_{k=1}^{K} \hat{\beta}_k^{FE} \overline{X}_{ki}$.

In the second stage, we assume $\hat{\mu}_i$ can be explained by time-invariant variables Z_{mi}. Thus we construct a function: $\hat{\mu}_i = \omega + \sum_{m=1}^{M} \gamma_m Z_{mi} + \eta_i$, in which, error term η_i theoretically explains the varying part of each unit's fixed effect while ω and γ_m jointly explain the common part of each unit's fixed effect. By regressing $\hat{\mu}_i$ on Z_{mi}, we can get $\hat{\omega}$ and $\hat{\gamma}_m$. Thus we can calculate $\hat{\eta}_i = \hat{\mu}_i - \hat{\omega} - \sum_{m=1}^{M} \hat{\gamma}_m Z_{mi}$.

In the last stage, we put $\hat{\eta}$ back into the function of the pooled model and get

$$Y_{it} = \alpha + \sum_{k=1}^{K} \beta_k X_{kit} + \sum_{m=1}^{M} \gamma_m Z_{mi} + \delta \hat{\eta}_i + \varepsilon_{it} \qquad (5)$$

Because of the two previous processes, $\hat{\eta}$ is no longer correlated with the vector of Z's. With pooled-OLS method, we can get coefficients for both time-varying and time-invariant variables. Worth noting, the coefficient of $\hat{\eta}$ should approximately equals to 1, since $\hat{\eta}$ attempts to explain the varying part of each unit's fixed effect.

To testify the validity of FEVD, *Plümper* and Troeger conduct a series of Monte Carlo simulations to compare their method with other popular techniques like pooled-OLS, random effects model (RE), Hausman-Taylor model (HT) as well as ordinary fixed effects model (FE). HT essentially employs instrumental variables to correct RE inconsistency when RE is applied in TSCS data analysis. *Plümper* and Troeger first showed their Monte Carlo results in comparison of FEVD, poold-OLS, RE and HT in five cases: "when unit effects are uncorrelated with the time-varying variables all four estimators are unbiased. When unit effects are uncorrelated with time-invariant variables, pooled OLS, random effects and fixed effects vector decomposition give unbiased, Hausman-Taylor biased estimates. When unit effects are correlated with the time-varying variables, pooled OLS and random effects model perform poorly, fixed effect vector decomposition and Hausman-Taylor are unbiased, Hausman-Taylor is less efficient. Finally, when unit effects are correlated with time-invariant variables, all procedures are about equally biased, but Hausman-Taylor is by far least efficient. In cases where both time-varying and time-invariant variables are correlated with the unit effects, fixed effects decomposition technique clearly outperforms its competitors." (*Plümper* and Troeger, 2004). Obviously, in most cases, FEVD is much superior to other approaches. *Plümper* and Troeger showed their refined results of Monte Carlo experiments after adding FE into comparison: "estimation by pooled-OLS or RE models is only appropriate if unit effects do not exist or if the Hausman test suggests that existing unit effects are uncorrelated with the regressors. If either of these conditions is violated, the FE model and the vector decomposition model compute more reliable estimates for time-varying variables." (*Plümper* and Troeger, 2007). In regard to estimates for time-invariant variables, FEVD performs better than the HT, pooled-OLS and RE while FE does not compute coefficients of time-invariant variables. (*Plümper* and Troeger, 2007). When comparing FEVD and FE, "the FE model performs best if the within variance of all regressors of interest is sufficiently large in comparison to their between variance. Otherwise, the efficiency of the fevd model becomes more important than the unbiasedness of the FE model." (*Plümper* and Troeger, 2007). Just as *Plümper* and Troeger summarized: “unit fixed effects vector decomposition technique produces the least biased and most efficient coefficients under a wide variety of data generating processes” (*Plümper* and Troeger, 2004) and “Under specific conditions, the vector decomposition model produces reliable estimates for time-invariant and rarely changing variables in panel data with unit effects than any alternative estimator of which we are aware of” (*Plümper* and Troeger, 2004), FEVD seems to be the most appropriate method for the analysis of national park visitation. Yet we will further check its validity with our data in the next section: Empirical Data and Variable Selection.

As is noted above, GLS can do considerable harm with TSCS data by giving wrong estimates (Beck, 2001). Another danger of applying GLS in TSCS data analysis is that the standard errors produced by FGLS usually lead to extreme overconfidence, often underestimating variability by 50% or more (Beck and Katz, 1995). In order to solve the problem of overconfidence in GLS, Beck and Katz applied a method called OLS with panel corrected standard errors (PCSE). In this approach, one of the main adjustments is that they use covariance matrix $(X'X)^{-1}\{X'\Omega X\}(X'X)^{-1}$ in PCSE instead of using $\{X'\Omega^{-1}X\}^{-1}$ from GLS. Their Monte Carlo results show that “The combination of OLS with PCSEs allows for accurate estimation of variability in the presence of panel error structures without inducing the severe problems caused by the Parks method” (Beck and Katz, 1995). However, they also realized “Ordinary least squares, with corrected standard errors, will not always prove to be superior to more complicated techniques” (Beck and Katz, 1995). In 2004, Beck and Katz further realized the main contribution of PCSE is to “improve on OLS standard errors, but they do not fix any other problems, and only fix the OLS standard errors with respect to panel heteroskedasticity and contemporaneous correlation of the errors” (Beck and Katz, 2004). In other words, PCSE mainly addresses two features related to error terms in TSCS data analysis: (i) panel heteroskedasticity; (ii) contemporaneous correlation of errors. In regard to the third feature, AR1 procedure can be applied to address serially correlated errors.

Now let us trace back to the three questions in the beginning of the literature review. Although TSCS data and panel data are similar in many ways, the significant differences between them caution us methods applied for panel data analysis may not be valid for TSCS data analysis. One key issue for TSCS data analysis is fixed effects and time-invariant variables. FEVD seems a reliable approach in many cases. Yet we need further check if this method is optimal for our data. In the three features of error terms in TSCS data analysis, panel heteroskedasticity and contemporaneous correlation of errors can be addressed by PCSE while serially correlated errors can be addressed by AR1 procedure.

CHAPTER THREE: EMPIRICAL DATA AND VARIABLE SELECTION

The data in this paper are from 42 national parks inside or around Arizona, from 1979 to 2003. Thus, 1,050 observations are available for regression analysis. The national parks are comprised of 24 National Monuments, 10 National Parks NPS Designated, 3 National Historical Parks, 2 National Historic Sites, 2 National Recreation Areas and 1 National Memorial. A brief summary table is displayed below:

Type	Number of Observations	Number of Specific Units
National Historical Park	75	3
National Historic Site	50	2
National Monument	600	24
National Park	250	10
National Recreation Area	50	2
National Memorial	25	1

Table 1: Park Category

Variables in consideration are mainly from the following three categories: demand factors, park attributes and climate/environmental variables.

Demand factors include variables denoted as gas_price, exchange_rate, population, cindex and mpi. Gas_price is the unleaded regular gas price which has been widely used in many other researchers' analysis (Johnson and Suits, 1983; Pergams and Zaradic, 2006). This variable has been proved to be significantly related with visitation (Pergams and Zaradic, 2006). Exchange_rate attempts to be a proxy for international travel cost. When the international travel cost is very high, people may choose travel to national parks. When the international travel cost is relatively cheap, people may prefer travel to other countries. In this sense, exchange_rate may have impact on national park visits. Population is a widely used explanatory variable in visits analysis too. In our data, population is the total number of residents in Arizona, New Mexico, southern Nevada, southern Utah and southern California. Cindex, competing destination index for short, accounts for distance effect. Distance effects are notable in many visits analyses (Hanink and White, 1997; Hanink and Stutts, 2002; Pergams and Zaradic, 2006). Research results show that competing destination index can significantly influence national park visits (Hanink and White, 1997; Hanink and Stutts, 2002). Variable mpi (market potential index) is also regarded as a significant factor for visitation (Hanink and White, 1997; Hanink and Stutts, 2002). Thank Srinivasa Ponnaluru's hard work, his research on "Visitation to the National Parks in the Southwest: the Influence of Economic and Climate Variables" (2005) provides very important information and valuable data for us. Our data of gas_price, exchange_rate, cindex and mpi get a lot of information from his research. "The national average price of unleaded regular gasoline was used in the study. The gasoline prices were deflated with adjusted Consumer Price Index for the West Region to suit the local trends. Gasoline price data were obtained from the website of the Energy Information Administration, Department of Energy." (Ponnaluru, 2005). "The exchange rate used is the Broad Index exchange rate. The Board of Governors of the Federal Reserve System constructs the Broad Index of the Foreign Exchange Value of the Dollar. This Broad Index exchange rate is a weighted average of the foreign exchange values of the U.S. dollar against the currencies of a large group of major U.S. trading partners. This is a real index: exchange rates used in the calculations are adjusted for aggregate price inflation in the markets of partner countries (Goldberg, 2004). Data is available from the New York Federal Reserve Industry Specific Exchange Rate Historic Data Files." (Ponnaluru, 2005). Cindex is expressed as $Cindex_i = \sum_i^n \frac{1}{D_{ij}}$, where D_{ij} is the distance of park *i* from park *j*. "Driving distances between the parks were considered and the data were collected piece by piece from the driving directions obtained from the website of the Arizona chapter of American Automobile Association (AAA).” (Ponnaluru, 2005). Mpi is expressed as $mpi_i = \sum_j \frac{Y_j}{D_{ij}}$, where Y_j is the real personal income in metro- or micropolitan statistical area (i.e. city) *j* and D_{ij} is the distance between park *i* and metro- or micropolitan statistical area *j*. “Annual data on personal income are obtained from the Regional Economic Information System of the Bureau of Economic Analysis. Personal incomes were deflated using the Consumer Price Index for the West Region which was collected from the Bureau of Labor Statistics, U.S. Department of Labor. Driving distances from the AAA website are used for distances between the parks and cities.” (Ponnaluru, 2005). The data of population in our research are from Bureau of Economic Analysis.

Park attributes have a number of variables seeking to account for the features of park. Area for individual parks is such a variable and has been proved to be one of the significant factors to influence national park visitation (Hanink and White, 1997; Weiler and Seidl, 2003; Weiler, 2006). To explain non-linear effect, variable areasq=area*area is included. “The data is obtained from the NPS Acreage Reports database.” (Ponnaluru, 2005). Estab, a variable expressing when the national park was established, is similar to park age to account for park’s history effect. In Ponnaluru’s work (2005), we got the data for park age and then we converted it to the year of establishment for each park. The park age data is “obtained from the National Park Service History website” (Ponnaluru, 2005). Typogz is a variable taking topographical diversity of parks into consideration. The data is from Economic Research Service (ERS). In addition, groups of dummy variables are included to better account for national park attributes. Variable lodging includes information of whether a national park has concessionary lodging facilities or not. Conversion of a park from National Monument to National

Park NPS Designated can exert much impact on park visitation (Weiler and Seidl, 2003; Weiler, 2006). Weiler and Seidl (2004) pointed out this signaling effect can attract as many as 11,642 new long-term visitors for the eight national parks in their research. Their results show that different national park designation may have impact on park visits, so two dummy variables are included for designation effect, or signaling effect. Np stands for whether a park is the National Park NPS Designated in the 16-category designation while Nhp stands for National Historical Park designation. The last group of dummy variables takes park's anomalies into consideration. Tonto indicates the year when fire occurred in Tonto National Forest (tonto=1); Cerro indicates the year when Cerro Grande fire took place in Bandelier National Monument (cerro=1); Petrified, sunset and casa show the year when accounting method for visits changed respectively in Petrified National Forest, Sunset Crater Volcano National Monument and Casa Grande Ruins National Monument (petrified=1; sunset=1 and casa=1).

The third category of variables is from climate/environmental variables. Jatemp and jutemp represent the average January and July temperature during the 25 years for each national park respectively. The data is from Western Regional Climate Center (WRCC). Jasun is average sunshine time in January for every national park during the 25 years, which uses hour as the unit. Juhumid, on the other hand, measures average humidity in July for each park during the 25 years. The data is from ERS. Just like areasq; jatempsq, jasunsq, jutempsq and juhumidsq are included to explain non-linear effect. Dry_extreme is a dummy variable deriving from the 12 month Standard Precipitation Index (SPI). SPI, from National Oceanic and Atmospheric Administration (NOAA) measures precipitation in an area in relation to historical averages and is standardized to range -4 to 4, with 0 as a normal year. When SPI≤-2, dry_extreme is set to 1. Many visitors go for lake recreation in US southwest and two parks are very famous for their lake visiting: Lake Mead NRA and Glen Canyon NRA. Thus a reservoir elevation index (REindex) is created to explore how the change of reservoir elevation may influence visits. Choose the maximum reservoir elevation during the 25 years of each lake as the denominator while each year's reservoir elevation as the numerator. The following equation introduces the conversion:

$REindex = \frac{RE_i}{RE_{\max}}$, in which, RE_i is every year park's reservoir elevation, while $RE_{\max}$ is the park's maximum reservoir elevation during the 25 years. Finally, an interaction term jutemp*REindex as juRE is included to show joint effect of July temperature and reservoir elevation. Figure 1 and 2 displays the change of reservoir elevation during the 25 years for Lake Mead and Glen Canyon respectively, which may imply future trend to some extent.

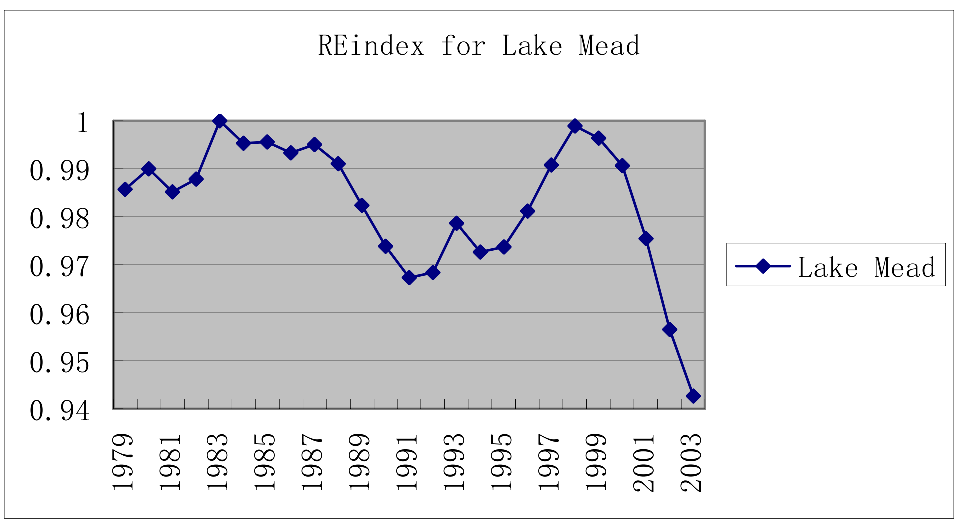


Figure 1: Change of Reservoir Elevation for Lake Mead NRA from 1979 to 2003

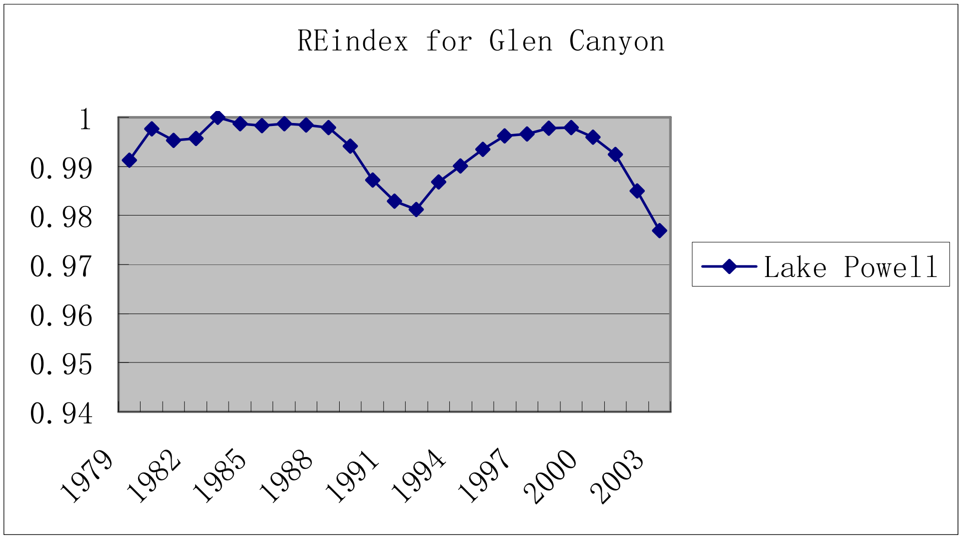


Figure 2: Change of Reservoir Elevation for Glen Canyon NRA from 1979 to 2003

Undoubtedly, both graphs show a tendency of declining reservoir elevation from 1979 to 2003. The lowest points of REindex both appear in 2003 and it may suggest a future trend of declining reservoir level.

The following tables display some variables' descriptive statistics. Table 2 includes the whole sample while Table 3 considers statistics for each park:

	Visits	Gas_Price($)	Population(1,000s)	Cindex	Area(sq. miles)	Estab(year)
Mean	630,867	1.73	38,298	0.16	155,796	1929
Median	164,988	1.52	38,781	0.16	6,412	1927
Maximum	9,838,702	2.86	47,515	0.28	1,495,664	1966
Minimum	4,912	1.19	29,363	0.09	40	1906
Std.Dev.	1,362,520	4.68	5,388	0.04	354,001	19
N=1050; Cross-section=42						

	JaTemp(^{o}F)	JaSun(hours)	JuTemp(^{o}F)	JuHumid(%)	REindex(lake only)
Mean	36.20	220.21	76.60	27.31	0.988
Median	33.65	230.00	76.70	26.00	0.991
Maximum	54.70	266.00	94.20	68.00	1
Minimum	21.70	141.00	60.10	14.00	0.943
Std.Dev.	8.29	32.22	7.59	8.68	0.012
N=1050; Cross-section=42					

Table 2: Descriptive Statistics for the Whole Sample

Worth noting, visits have a large variation with maximum 9,838,702 and minimum 4,912. Area also has a large variation with maximum 1,495,664 and minimum 40. It seems time-invariant variable area is correlated with the time-varying dependent variable visits. How can it be? Fixed effect in our case is the factor that explains the visitation for each park constantly. While visitation itself is a time-varying variable, thus, fixed effect must be correlated with time-varying variable; on the other hand, area as a typical time-invariant variable is probably related with the fixed effect. Correlation by correlation, therefore area seems to be correlated with visits. The assumption serves to help decide which may be the optimal method in our analysis and we will discuss more in the next section.

Park Name	Average Visits	Area(sq. miles)	Estab(year)
Arches NP	608,547	76,519	1949
Aztec Ruins NM	66,017	318	1928
Bandelier NM	291,934	33,677	1932
Bryce Canyon NP	841,217	35,835	1923
Canyon de Chelly NM	645,312	83,840	1931
Canyonlands NP	282,586	337,598	1964
Capitol Reef NP	505,097	241,904	1937
Capulin Volcano NM	53,771	793	1916
Carlsbad Caverns NP	648,792	46,766	1930
Casa Grande Ruins NM	142,522	473	1918
Cedar Breaks NM	489,036	6,154	1933
Chaco Culture NHP	71,673	33,960	1907
Chiricahua NM	73,474	11,985	1924
Coronado Nmem	68,039	4,750	1952
El Morro NM	60,635	1,279	1906
Fort Bowie NHS	7,790	999	1964
Fort Union NM	15,781	721	1966
Gila Cliff Dwellings NM	47,594	533	1907
Glen Canyon NRA	2,501,760	1,254,429	1958
Grand Canyon NP	3,625,148	1,217,403	1919
Hovenweep NM	24,190	785	1923

Hubbell Trading Post NHS	185,021	160	1965
Joshua Tree NP	1,001,726	789,745	1928
Lake Mead NRA	7,852,648	1,495,664	1947
Montezuma Castle NM	726,287	858	1906
Natural Bridges NM	101,572	7,637	1908
Navajo NM	70,816	360	1909
Organ Pipe Cactus NM	232,350	330,689	1937
Pecos NHP	41,126	6,670	1965
Petrified Forest NP	764,013	93,533	1906
Pipe Spring NM	44,293	40	1923
Rainbow Bridge NM	208,032	160	1910
Saguaro NP	680,732	91,440	1933
Salinas Pueblo Missions NM	36,746	1,071	1909
Sunset Crater Volcano NM	357,610	3,040	1926
Tonto NM	69,663	1,120	1907
Tumacacori NHP	57,760	361	1908
Tuzigoot NM	106,521	812	1939
Walnut Canyon NM	118,010	3,579	1950
White Sands NM	562,118	143,733	1933
Wupatki NM	226,998	35,422	1924
Zion NP	1,981,467	146,598	1918

N=1050; Cross-section=42

Park Name	JaTemp(^{o}F)	JaSun(hours)	JuTemp(^{o}F)	JuHumid(%)
Arches NP	30.1	178	82.6	21
Aztec Ruins NM	30.5	183	75.9	21
Bandelier NM	28.1	202	67.3	37
Bryce Canyon NP	21.7	141	62.5	19
Canyon de Chelly NM	32	230	77.2	26
Canyonlands NP	28	178	77.8	21
Capitol Reef NP	27.9	178	76.7	21
Capulin Volcano NM	31.2	211	70	42
Carlsbad Caverns NP	44.1	246	77.9	35
Casa Grande Ruins NM	50.7	248	90.4	20
Cedar Breaks NM	28.2	141	71.1	19
Chaco Culture NHP	28.1	183	72.7	21
Chiricahua NM	44	260	74.8	28
Coronado Nmem	45.3	260	75.4	28
El Morro NM	28.8	233	68	28
Fort Bowie NHS	46.2	260	82.6	28
Fort Union NM	29.8	202	60.1	37
Gila Cliff Dwellings NM	36.3	258	70	36
Glen Canyon NRA	34.7	230	85.5	26
Grand Canyon NP	30.4	230	67	26
Hovenweep NM	28.1	178	75.8	21

Hubbell Trading Post NHS	32.6	230	73.7	26
Joshua Tree NP	50	224	88.4	68
Lake Mead NRA	47.2	244	94.2	14
Montezuma Castle NM	42.8	230	81.9	31
Natural Bridges NM	28.3	178	73.1	21
Navajo NM	29.8	230	71.9	26
Organ Pipe Cactus NM	54.7	266	89.2	23
Pecos NHP	31.4	202	68.7	37
Petrified Forest NP	34.9	230	76	26
Pipe Spring NM	34.8	230	76.7	31
Rainbow Bridge NM	32.2	178	81.4	21
Saguaro NP	51.7	260	86.5	28
Salinas Pueblo Missions NM	37.3	233	78.4	28
Sunset Crater Volcano NM	27.5	230	65.8	26
Tonto NM	42.6	248	77.9	20
Tumacacori NHP	48.4	260	81.1	28
Tuzigoot NM	45.2	230	83.1	31
Walnut Canyon NM	29.7	230	66.1	26
White Sands NM	39.1	246	79.8	29
Wupatki NM	35.7	230	80.1	26
Zion NP	40.4	210	82.1	20

N=1050; Cross-section=42

Table 3: Descriptive Statistics for Each Park

In Table 3 we can see the correlation between area and visits is even more obvious. Another point worth noting is the singularities in our data. For example, Bryce Canyon has the lowest January temperature of 21.7°F, 5.8°F lower than the second lowest January temperature in Sunset Crater Volcano. Joshua tree has the moistest weather in summer with a humidity of 68%, much higher than the second moistest weather in Capulin Volcano. The explanatory variable has two categories: time-varying and time-invariant. Practice and Research have proved time-varying variables (gas_price for example) as well as time-invariant variables (area for example) can both impact visits significantly. Thus, in this case, which model should be chosen for our regression analysis? The following section attempts to answer the question.

CHAPTER FOUR: MODEL SPECIFICATION

Data in our research are from 42 national parks during 25 years. "In cases where both time-varying and time-invariant variables are correlated with the unit effects, fixed effects decomposition technique clearly outperforms its competitors." (*Plümper* and Troeger, 2004). As is noted on Table 2&3, it seems FEVD would be the best choice among the four methods: FEVD, pooled-OLS, RE and HT. In comparison of FEVD and ordinary FE, "the FE model performs best if the within variance of all regressors of interest is sufficiently large in comparison to their between variance. Otherwise, the efficiency of the fevd model becomes more important than the unbiasedness of the FE model." (*Plümper* and Troeger, 2007). Because some variables' variances are either too large or too small to display, instead the following table shows a summary of all variables' within and between standard deviation.

Variable	Derivation	Within Std. Dev.	Between Std. Dev.
logV	log(visits)	0.083	2.30
logG	log(gas_price)	0.059	0
logEx	log(exchange_rate)	0.0098	0
logpop	log(population)	0.020	0
petrified	petrified	0.0043	0.0014
sunset	sunset	0.0052	0.0024
casa	casa	0.0018	0.0002
cerro	cerro	0.0009	3.81E-05
tonto	tonto	0.0009	3.81E-05
dry_extreme	SPI≤-2	0.099	0.023

area	area	0	1.28E+11
jatemp	jatemp	0	67.51
jasun	jasun	0	1,062.56
jutemp	jutemp	0	40.83
juhumid	juhumid	0	77.15
areasq	area*area	0	2.24E+23
jatempsq	jatemp*jatemp	0	416,578
jasunsq	jasun*jasun	0	186,590,682
jutempsq	jutemp*jutemp	0	1,021,779
juhumidsq	juhumid*juhumid	0	475,264
cindex	cindex	0	0.0016
typogz	typogz	0	0.44
lodging	lodging	0	0.107
Nhp	Nhp	0	0.107
estab	estab	0	365.6
mpi	mpi	679,632	2,676,315
Np	Np	0.011	0.157
REindex	REindex	0.0023	0.2129
juRE	jutemp*REindex	0.22	19.15

Note: Variance=Std.Dev.^{2}

Table 4: Within and Between Standard Deviation for Each Variable

Worth noting, logV is used instead of visits. LogG, logEx and logpop take place of gas_price, exchange_rate and population respectively. According to the table, within variances of logG, logEx, logpop, petrified, sunset, casa, cerro, tonto and dry_extreme are larger than between variances. However, within variances of area, jatemp, jasun, jutemp, juhumid, areasq, jatempsq, jasunsq, jutempsq, juhumidsq, cindex, typogz, lodging, Nhp, estab, mpi, Np, REindex and juRE are smaller than between variances ($\text{Variance}=\text{Std.Dev.}^2$). Within variances of some variables are actually equal to 0. Obviously, the table suggests FEVD should be more appropriate for regression analysis than FE. On the other hand, the table provides a valid classification of time-varying and time-invariant variables. Since within variance is larger than between variance, logG, logEx, logpop, petrified, sunset, casa, cerro, tonto and dry_extreme are time-varying variables; while within variance is smaller than between variance, so that area, jatemp, jasun, jutemp, juhumid, areasq, jatempsq, jasunsq, jutempsq, juhumidsq, cindex, typogz, lodging, Nhp, estab, mpi, Np, REindex and juRE are regarded as time-invariant variables. The classification contributes to a specific econometric model.

$$
\begin{aligned}
\log V_{it} &= \alpha + \beta_1 \log G_t + \beta_2 \log Ex_t + \beta_3 \log pop_t \\
&+ \beta_4 petrified_{it} + \beta_5 sunset_{it} + \beta_6 casa_{it} + \beta_7 cerro_{it} + \beta_8 tonto_{it} + \beta_9 dry\_extreme_{it} \\
&+ \gamma_1 area_i + \gamma_2 jatemp_i + \gamma_3 jasun_i + \gamma_4 jutemp_i + \gamma_5 juhumid_i \\
&+ \gamma_6 areasq_i + \gamma_7 jatempsq_i + \gamma_8 jasunsq_i + \gamma_9 jutempsq_i + \gamma_{10} juhumidsq_i \\
&+ \gamma_{11} cindex_i + \gamma_{12} typogz_i + \gamma_{13} lodging_i + \gamma_{14} Nhp_i + \gamma_{15} estab_i \\
&+ \gamma_{16} mpi_{it} + \gamma_{17} Np_{it} + \gamma_{18} REindex_{it} + \gamma_{19} juRE_{it} + \varepsilon_{it}
\end{aligned}
$$

i represents the specific park unit; t is the observed year. Mpi, Np, REindex and juRE vary a little with year. So it is more valid to treat them as time-invariant variables in the model.

logG is expected to be negatively related with visitation. Coefficient of logEx is also expected to be negative. Coefficient of logpop is expected to be positive while petrified, sunset, casa, cerro and tonto are all expected to be negatively related with visitation. Coefficient of dry_extreme is expected to be negative because extreme dry weather may make the destination an unpleasant place for travel. Among time-invariant variables, cindex and estab bear a negative relationship with logV while mpi, typogz, lodging and Np are expected to have a positive relationship. Similar to the signaling effect of Np, Nhp may negatively influence visits, because National Historical Park may seem to be a more serious site for visits than other national parks, which causes its coefficient to be negative. The overall effect of REindex and Nraju is supposed to positively influence visits. That means the marginal effect of REindex is expected to be positive. However, the respective signs cannot be decided before hand. The same as REindex, the overall effect of area and areasq is believed to be positive for visits but the respective signs are not decided before hand either. Jatemp, jasun, jutemp, juhumid, jatempsq, jasunsq, jutempsq and juhumidsq's relations with visits are not quite clear before hand. Table 5 is a display of expected sign for each variable:

Variable	Derivation	Expected Sign
logV	log(visits)	NA
logG	log(gas_price)	-
logEx	log(exchange_rate)	-
logpop	log(population)	+
petrified	petrified	-
sunset	sunset	-
casa	casa	-
cerro	cerro	-

tonto	tonto	-
dry_extreme	SPI≤-2	-
cindex	cindex	-
mpi	mpi	+
estab	estab	-
typogz	typogz	+
lodging	lodging	+
Np	Np	+
Nhp	Nhp	-
REindex	REindex	NA
juRE	Jutemp*REindex	NA
area	area	NA
jatemp	jatemp	NA
jasun	jasun	NA
jutemp	jutemp	NA
juhumid	juhumid	NA
areasq	area*area	NA
jatempsq	jatemp*jatemp	NA
jasunsq	jasun*jasun	NA
jutempsq	jutemp*jutemp	NA

juhumidsq	juhumid*juhumid	NA

NA: Not Applicable

Table 5: Expected Sign of Coefficient for Each Variable

To verify our choice of FEVD, a latter section will compare regression results with those from pooled-OLS, pooled-OLS with PCSE as well as ordinary FE.

CHAPTER FIVE: ECONOMETRIC RESULTS

A variety of forms have been tested, including log-log, log-linear and linear-linear. Finally a combination of log-log and log-linear form is accepted as the best for regression analysis. Table 6 displays the econometric results:

Variable	Coefficient	Standard Error
logV=		
Intercept	34.27	1.56
logG	-0.24	0.07
logEx	-0.43	0.10
logpop	0.83	0.11
petrified	-0.10^	0.06
sunset	-0.93	0.03
casa	-0.44	0.09
cerro	-0.24	0.10
tonto	-0.29^	0.22
dry_extreme	-0.08^	0.05
cindex	16.74	0.09
mpi	3.40E-08	1.75E-09
estab	-0.02	0.0005
typogz	0.18	0.01
lodging	1.56	0.02
Np	1.02	0.01

Nhp	-0.38	0.03
REindex	-47.76	0.66
juRE	0.57	0.01
area	1.02E-05	9.34E-08
jatemp	0.45	0.02
jasun	-0.07	0.002
jutemp	-0.08	0.03
juhumid	0.18	0.003
areasq	-7.37E-12	7.23E-14
jatempsq	-0.005	0.0002
jasunsq	9.73E-05	5.56E-06
jutempsq	0.0004	0.0002
juhumidsq	-0.002	4.44E-05
Eta (η)	0.99987	0.008
Statistics		
F statistic	71,092.91	
R-squared	0.9824	
Adjusted R-squared	0.9819	
Root MSE	0.202	

^: The null hypothesis can be rejected at 95% confidence level in a two-tailed test.
Table 6: Regression Results (log / log&linear)

According to the results of R-squared value and F statistic, the model makes sense. Most of the variables are significant at the level 0.005 (two tails) except petrified, cerro, tonto, dry_extreme and jutempsq. Yet cerro and jutempsq are significant at the level 0.05 (two tails). Among the demand-factor variables, results of logG, logEx and logpop are consistent with our prediction. However, in the same category, cindex is positively related with visitation rather than negatively related. The direct interpretation for cindex implies the closer two national parks, the more visits attracted to both parks. A possible explanation is that the units we include are various kinds of national parks, so that parks emphasize on different themes by name, which may send positive signals to tourists and attract their visitation at the same time. Market potential index indicates how many potential visits for a park. Thus, the positive coefficient of mpi is no surprise. The result is consistent to Hanink and his colleague's researches.

In variables of park attributes, estab conforms to our expectation with a negative sign. Based on Table 2, the earliest year for a park's establishment is 1906, while the latest year is 1966. Combined with $\hat{\beta}$ for estab, $(1966-1906)*(-0.02)=-1.2$, which suggests the park visits can be reduced as much as 1.2% for the latest national park because of its late establishment. The result is reasonable because a park with longer history usually enjoys wide-spread fame accumulated by time and can be put into visits' schedule more frequently. The positive sign of typogz is also consistent with our prediction, since it is easy to envision a larger diversity in the national park can contribute to a larger visitation. Dummy variable lodging shows positive relationship with national park visits. However, interpretation for lodging should be very cautious. It could be that the lodging facility has a positive effect on visiting; it could also be the case that a popular park attracts investors to build lodging facilities inside. The group dummy variables of signaling effect show consistent results with our expectation. Np has a positive signaling effect on visitation, whose result verifies Weiler and Seidl's conclusion (Weiler and Seidl, 2003; Weiler, 2006). The effect of Nhp on visitation is negative as expected. The anomaly dummy variables conform to our expectation too. All the dummy variables:

petrified, sunset, casa, cerro and tonto, are negatively related with visitation. In addition, cerro's significance as well as tonto's insignificance may suggest conclusions should be cautious when the impact of fires on national park visits is in discussion. Of course, a bigger fire like Cerro Grande fire is more likely to influence visitation. Coefficient of area is positive while coefficient of areasq is negative, which reveals visitation is increasing (first-derivative positive) at a decreasing rate (second-derivative negative) as national park area enlarges. Figure 3 shows the effect of area for each park.

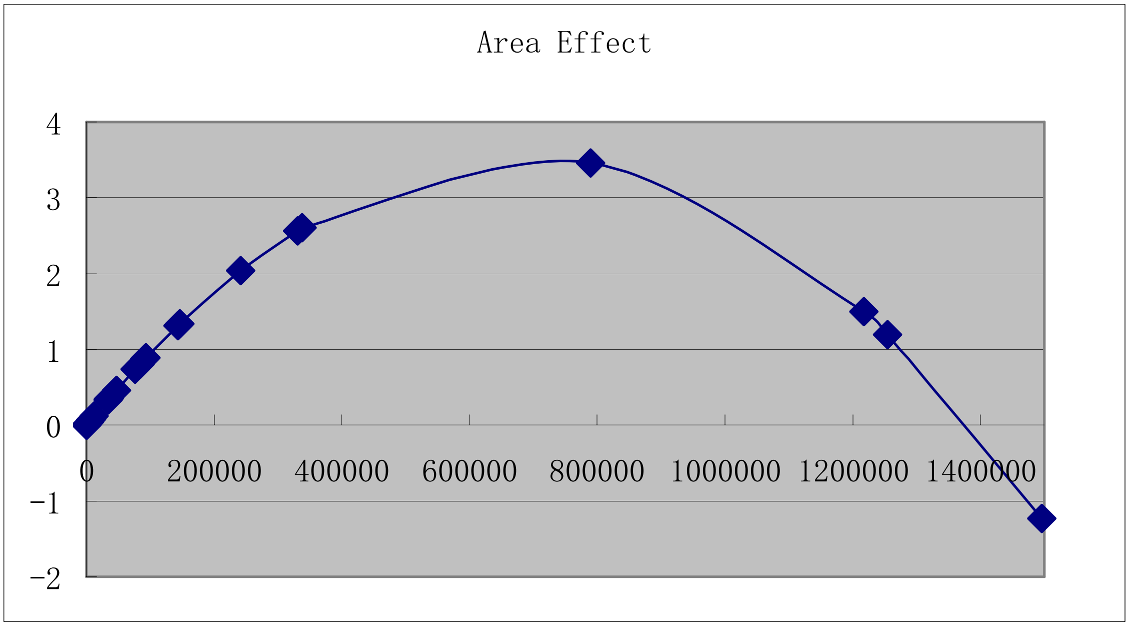


Figure 3: Effect of Area on LogV for Each Park

It is usually considered true that visits increase when park area increases. However, according to our result, we can see visits' increase with the increase of park area is not always valid. For some especially large parks, they may have more visits because of their fame or special attributes (lake recreation, etc), but the vast area may not be a key factor. In our case, when park area exceeds 800,000 square miles, it has no more positive effect on visits. The three parks exceeding 800,000 square miles are Lake Mead NRA, Glen Canyon NRA and Grand Canyon NP. Worth noting, since Lake Mead and Glen Canyon are characterized by lake recreation, if water area is excluded from the total area and only land area is left for regression, we are not sure whether the conclusion would change, which waits for future studies.

In the category of climate/environmental variables, the marginal effect of REindex expresses as $(-47.76 + 0.57 * jutemp)$, which is expected to be positive. Two parks with lake recreation are included in our analysis. Their average July temperature is 85.5^{0}F and 94.2^{0}F respectively. So that the marginal effect of REindex is 0.975 for Glen Canyon NRA and 5.934 for Lake Mead NRA, which conforms to our expectation, although the number for Lake Mead is really large. In order to observe the relationship between REindex and visits clearly, take Lake Mead NRA from 2002 to 2003 as an example. The change of REindex is 0.9427-0.9566=-0.0139. According to the marginal effect, the visits should drop by (0.0139*5.934)*100%≈8.25%, namely about 653,035 visits. The number will be used in latter part of input-output analysis. We notice the possible over estimation of marginal effect for Lake Mead. So we exclude the variables associated with lakes and re-run the full model without the data of Glen Canyon and Lake Mead. Surprisingly, the econometric results are very consistent with those of full model with all observations. We show the table of econometric results without Glen Canyon and Lake Mead at the end of this section.

Like Figure 3 for area, the following graphs display the effect of jatemp, jasun, jutemp and juhumid on visits for each park respectively.

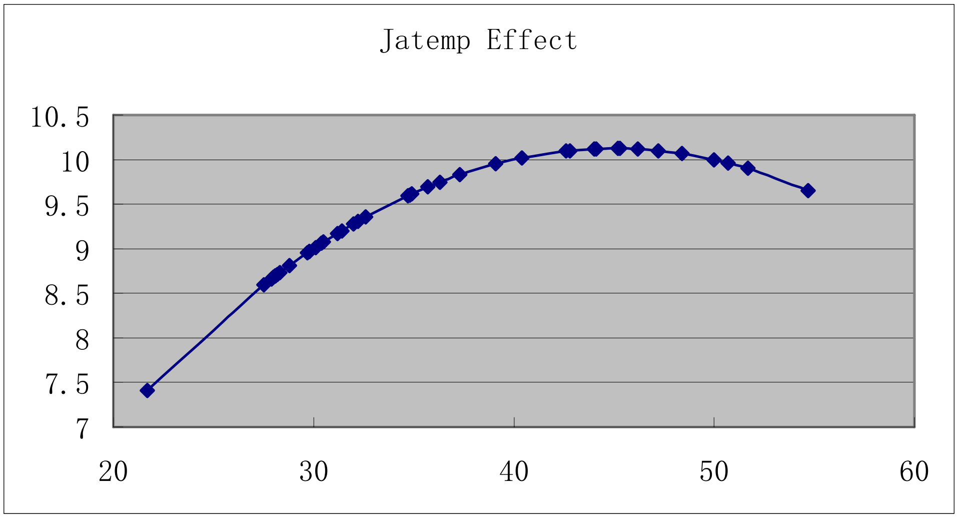


Figure 4: Effect of January Temperature on LogV for Each Park

From the graph, it is very clear the visits increase when January temperature increases at first. Yet when temperature exceeds about 45°F, the marginal effect of January temperature becomes negative. In the US southwest, the conclusion is reasonable. Arizona, New Mexico and southern California are famous for its warm winter. When warmer winter comes, more visits would like to go to the resorts. However, high temperature does not always mean good. When temperature reaches its turn point, some people may prefer a little cold weather in the winter. That is why at first visits increase and then decrease as January temperature goes up.

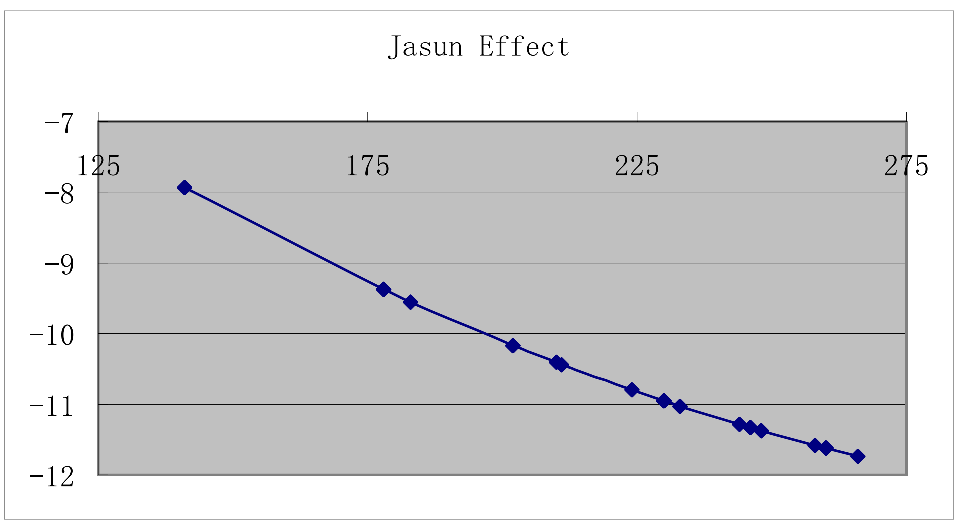


Figure 5: Effect of January Sunlight on LogV for Each Park

Visits decrease as January sunlight increases and the marginal effect of January sunlight is always negative according to the graph. National parks in US southwest are featured with warm weather and good sunlight during the winter. The conclusion on January sunlight seems conflict with common sense. The possible reason may be the correlation between January temperature and sunlight reaches 0.74. In addition, if we consider the effect of both January temperature and sunlight, it may be the case that temperature is more decisive in visitation considerations than sunlight.

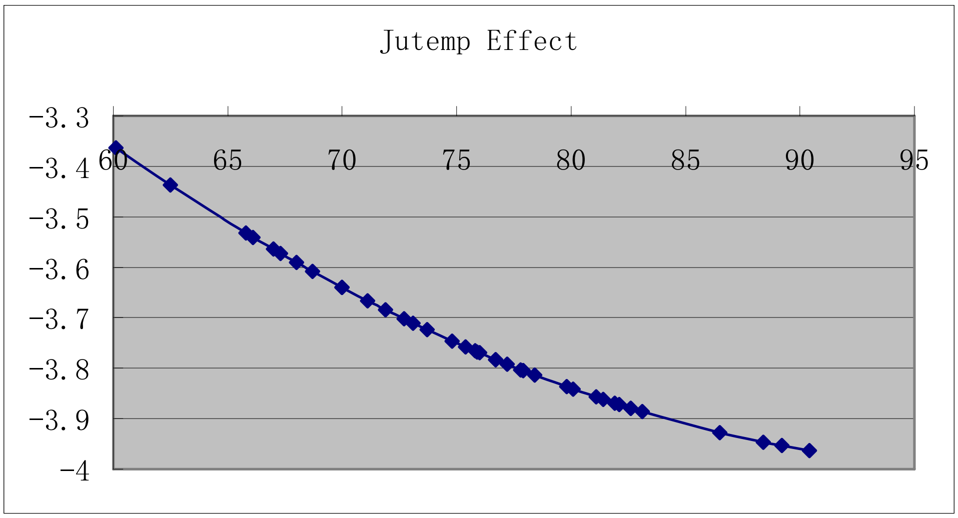


Figure 6: Effect of July Temperature on LogV for Each Park
(Except Glen Canyon NRA and Lake Mead NRA)

According to the graph, the marginal effect of July temperature is always negative and visits decrease as July temperature increases. National Parks in US Southwest are usually very hot during the summer. Thus, it is reasonable that people would like to stay with air-conditioning at home rather than go to parks. The marginal effect of July temperature for Glen Canyon NRA and Lake Mead NRA are different because of the interaction term. Their marginal effect is always positive according to our calculation. It means more visitors would come when July temperature is higher. Namely, the hotter, more visitors. For the parks featured with lake recreation, our conclusion is consistent with other research findings (Mendelsohn and Markowski, 1999; Loomis and Crespi, 1999).

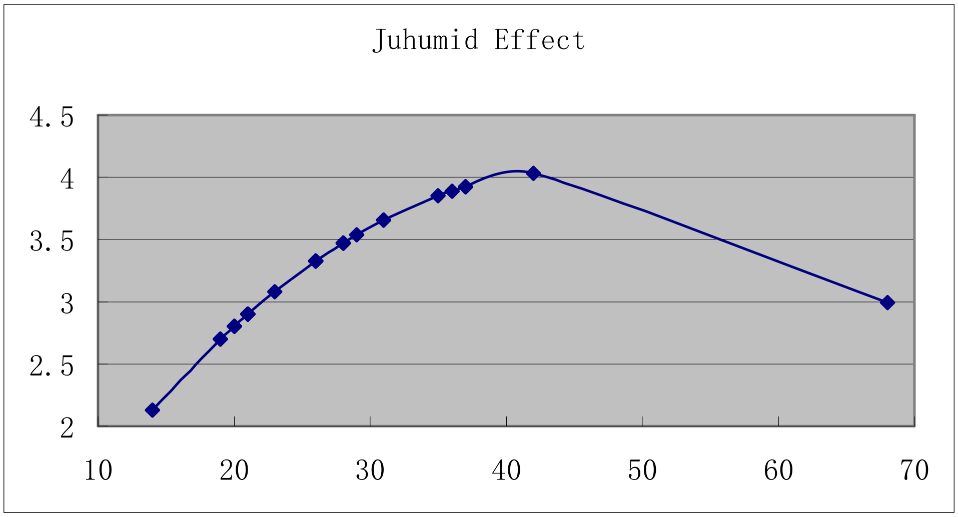


Figure 7: Effect of July Humidity on LogV for Each Park

It is not strange to find the visitation increases as July humidity increases. US southwest are usually hot and dry during the summer. When humidity increases, more people would like to go to national parks for recreation and rest. There is a notable outlier in the graph, which is Joshua Tree NP located in California. Its average humidity reaches 68% during the summer. Just like January temperature, we can see it is not always good with a high humidity in the summer. A higher humidity may make some visitors feel uncomfortable and thus reduce visits during the summer.

The following table shows the econometric results without Glen Canyon and Lake Mead.

Variable	Coefficient	Standard Error
logV=		
Intercept	34.14	1.56
logG	-0.24	0.06
logEx	-0.44	0.1
logpop	0.84	0.11
petrified	-0.10^	0.06
sunset	-0.93	0.03
casa	-0.44	0.09
cerro	-0.24	0.1
tonto	-0.29^	0.22
dry_extreme	-0.08^	0.05
cindex	16.74	0.09
mpi	3.41E-08	1.75E-09

estab	-0.02	0.0005
typogz	0.18	0.01
lodging	1.56	0.02
Np	1.02	0.01
Nhp	-0.38	0.03
area	1.02E-05	9.32E-08
jatemp	0.45	0.02
jasun	-0.07	0.002
jutemp	-0.08	0.03
juhumid	0.18	0.003
areasq	-7.37E-12	7.22E-14
jatempsq	-0.005	0.0002
jasunsq	9.74E-05	5.56E-06
jutempsq	0.0004	0.0002
juhumidsq	-0.002	4.44E-05
Eta (η)	1.000116	0.008
Statistics		
F statistic	65,017.58	
R-squared	0.9783	
Adjusted R-squared	0.9777	
Root MSE	0.203	

^: The null hypothesis can be rejected at 95% confidence level in a two-tailed test.
Table 7: Regression Results (log / log&linear) without Glen Canyon and Lake Mead

By comparing the coefficient of two tables, we can see the econometric results are almost the same, no matter whether we include parks with lake recreation or not. It makes us believe our interpretation on common climate indicators like temperature, sunlight as well as humidity holds water. However, the result may also imply we need sufficient data to treat parks with lake recreation separately. New findings about parks with lake recreation may be discovered if we have enough data for analysis.

The discovery of climate/environmental variables' impact on visitation is meaningful. Firstly, the research studies how the change of reservoir elevation can influence tourists. The result shows very positive relationship between reservoir level and visits. In our example, the decrease of Lake Mead elevation from 2002 to 2003 can cause a decrease of 653,035 visitors in 2003. The figure may be larger than our expectation. However, the result strongly points out how the environmental change can influence people’s living style. Secondly, it verifies temperature can significantly influence people’s traveling. For US southwest that is usually featured with warm winter and hot summer, increase of January temperature and decrease of July temperature can both increase visits for most parks, which is understandable. Warm winter not just attracts local residents but usually attracts tourists nation wide and world wide, especially those from US Northeast and US Northwest. When hot summer becomes cooler, more residents and far-away visitors would like to go to parks for recreation and rest. Worth mentioning, the parks featured with lake recreation show opposite response to the change of temperature in summer. To them, the result is “the hotter, more visits”. Thirdly, the research takes January sunlight into consideration and the variable has a significant effect on visits. Although the interpretation shows conflict with common sense, it presents a topic of how to understand the relationship between January sunlight and visits for future studies. Last but not least, the conclusion of July humidity on visits conforms to our expectation. Summer in the US southwest is usually dry and hot. When the weather becomes more humid for national parks, more people would like to go for a visit. However, Joshua Tree is an exception, since much humid weather can make people uncomfortable.

CHAPTER SIX: COMPARISON WITH OTHER REGRESSIONS

To testify the validity of FEVD in our analysis, we also apply methods of pooled-OLS, pooled-OLS with PCSE as well as ordinary FE and compare the results.

Variable	Coefficient	Standard Error
logV=		
Intercept	46.53	8.041
logG	-0.52	0.170
logEx	-0.61^	0.248
logpop	0.27^	0.288
petrified	-0.39^	0.317
sunset	0.02^	0.280
casa	0.31^	0.546
cerro	1.07^	0.742
tonto	-0.75^	0.743
dry_extreme	-0.27^	0.233
cindex	16.09	0.951
mpi	4.03E-08	1.03E-08
estab	-0.02	0.002
typogz	0.17	0.056
lodging	1.58	0.133
Np	1.03	0.081

Nhp	-0.38	0.086
REindex	-46.96	3.562
juRE	0.56	0.042
area	1.02E-05	6.28E-07
jatemp	0.49	0.068
jasun	-0.07	0.016
jutemp	-0.10^	0.110
juhumid	0.17	0.022
areasq	-7.39E-12	5.20E-13
jatempsq	-0.01	0.001
jasunsq	9.71E-05	3.82E-05
jutempsq	4.62E-04^	0.001
juhumidsq	-2.12E-03	2.69E-04
Statistics		
F statistic	124.36	
R-squared	0.773	
Adjusted R-squared	0.767	
Root MSE	0.736	

^: The null hypothesis can be rejected at 95% confidence level in a two-tailed test.
Table 8: Pooled-OLS Regression Results (log / log&linear)

By comparing the results of FEVD and pooled-OLS, we can see logEx, logpop, dry_extreme and all the anomaly dummy variables are insignificant at the level 0.05 in pooled-OLS. That means, only logG among the time-varying variables has significant effect on visits. In addition, jutemp and jutempsq are insignificant at the level 0.05 among time-invariant variables. However, symbols of the significant variables from FEVD and pooled-OLS are the same. The F value and adjusted R-squared value in FEVD are much larger than those in pooled-OLS, which suggests FEVD is superior to pooled-OLS.

Variable	Coefficient	Standard Error
logV=		
Intercept	36.10	5.84
logG	-0.19	0.07
logEx	-0.41	0.14
logpop	0.81	0.21
petrified	-0.07^	0.07
sunset	-0.67	0.18
casa	-0.30^	0.17
cerro	-0.23^	0.17
tonto	-0.30	0.08
dry_extreme	-0.07	0.04
cindex	16.04	0.58
mpi	3.64E-08	1.38E-08
estab	-0.02	1.26E-03
typogz	0.13	0.05
lodging	1.66	0.09
Np	0.61	0.10

Nhp	-0.43	0.09
REindex	-51.59	2.60
juRE	0.61	0.03
area	1.16E-05	5.87E-07
jatemp	0.54	0.05
jasun	-0.07	0.01
jutemp	-0.11^	0.08
juhumid	0.19	0.01
areasq	-8.33E-12	4.52E-13
jatempsq	-0.01	6.43E-04
jasunsq	1.08E-04	2.93E-05
jutempsq	5.24E-04^	5.67E-04
juhumidsq	-2.36E-03	1.98E-04
Statistics		
Wald Chi-square statistic	16,249.43	
R-squared	0.963	

^: The null hypothesis can be rejected at 95% confidence level in a two-tailed test.
Table 9: Pooled-OLS with PCSE Regression Results (log / log&linear)

In comparison of the results of FEVD and pooled-OLS with PCSE, several differences are noted. First, tonto and dry_extreme are insignificant in FEVD while they are significant at the level 0.05 in pooled-OLS with PCSE. For casa and cerro, they are just the inverse cases. Casa and cerro are significant in FEVD but are not significant in pooled-OLS with PCSE. Second, just like pooled-OLS, jutemp and jutempsq are insignificant in pooled-OLS with PCSE either. These differences would cause differences in interpretation. However, again, among the significant variables, the symbols are consistent between the two methods. Since F statistic and R-squared value in FEVD are larger than those in pooled-OLS with PCSE, FEVD is still better than pooled-OLS with PCSE.

Variable	Coefficient	Standard Error
logV=		
Intercept	8.47	0.59
logG	-0.22	0.03
logEx	-0.61	0.07
logpop	0.43	0.11
petrified	-0.14^	0.11
sunset	-0.98	0.10
casa	-0.45	0.11
cerro	-0.24	0.09
tonto	-0.30	0.09
dry_extreme	-0.07	0.03
mpi	-2.00E-08^	2.15E-08
Np	0.04^	0.08
REindex	52.02^	68.13
juRE	-0.55^	0.73
Statistics		
F statistic	25.75	
R-squared Within	0.26	

R-squared Between	0.085
R-squared Overall	0.94

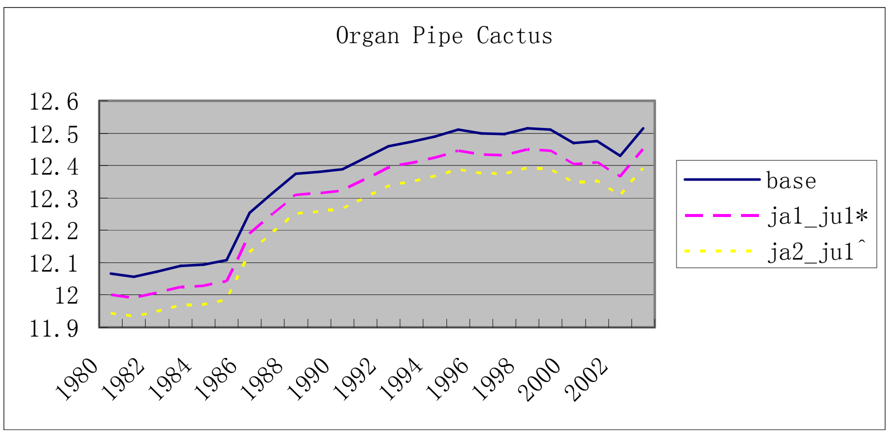
^: The null hypothesis can be rejected at 95% confidence level in a two-tailed test.
Table 10: Ordinary Fixed Effects Regression Results (log / log&linear)

According to the rule of fixed effects model, all the time-invariant variables should be thrown out. Yet since mpi, Np, REindex and juRE have slight variation with time, they are retained in the model. Two obvious differences are noticed. Tonto and dry_extreme are insignificant in FEVD but now all the time-varying variables except petrified are significant at the level 0.05 in ordinary FE. Besides, mpi, Np, REindex and juRE are all significant in FEVD but now all become not significant in ordinary FE. These differences would cause quite different interpretation of the explanatory variables. Worth noting, all the significant variables still have the same sign. Since F value in FEVD is far larger than that in ordinary FE, it is reasonable to consider FEVD a more valid model in our case.

The comparison of FEVD to pooled-OLS, pooled-OLS with PCSE as well as ordinary FE verifies our choice of FEVD model. Interestingly, the same sign and close coefficients for significant variables strengthen the validity of our results from FEVD and our interpretation.

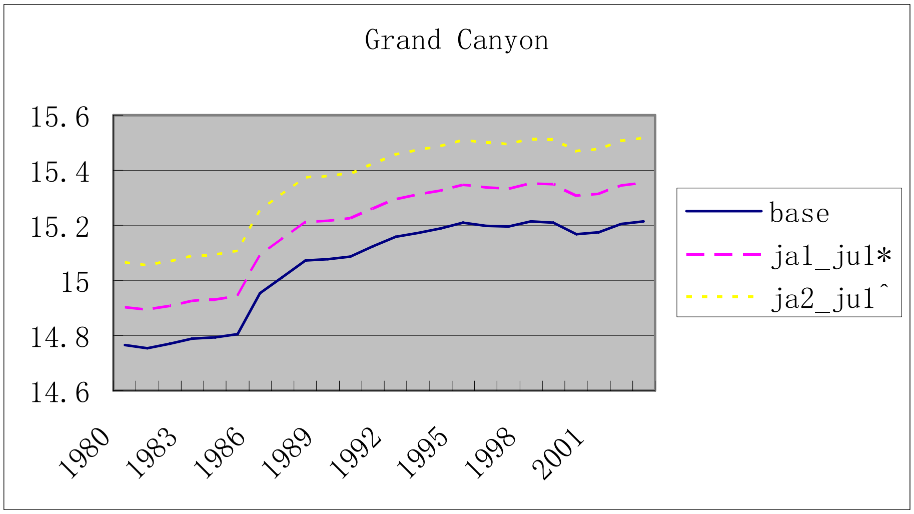
CHAPTER SEVEN: SENSITIVITY ANALYSIS

Global warming is a fact in our life. How will the global warming influence national park visitation? To study the possible impact of global warming, we conduct the sensitivity analysis for January and July temperature based on our regression results. Basically we have three cases: (i) the base case; (ii) both January and July temperature increase by 1; (iii) January temperature increases by 2 while July temperature increases by 1. We calculate the predicted visits for three cases and compare the results. Figures 6, 7 and 8 show our findings.



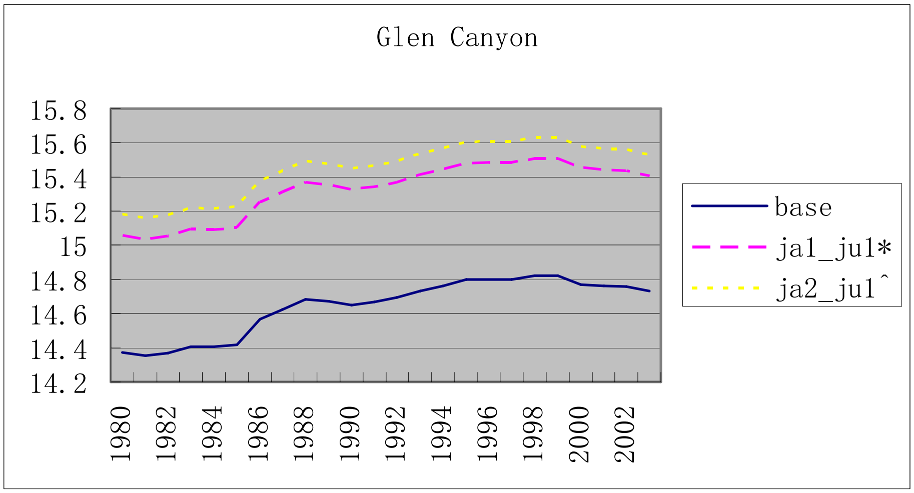
*: January and July temperature increase by 1 respectively
^: January temperature increases by 2 and July temperature increases by 1
Note: Y axis represents the predicted value of logV
Figure 8: Sensitivity Analysis of January and July temperature for Organ Pipe Cactus NM

Organ Pipe Cactus NM is located along the southern border of Arizona and next to Mexico. According to the graph, the increase of temperature causes visits to decrease. Considering the usually hot and dry weather in southern Arizona, southern New Mexico and part of southern California, the result is reasonable. Global warming is harmful for visitation not just to Organ Pipe Cactus, but to most parks in this area, like Saguaro NP, Joshua Tree NP, etc.



*: January and July temperature increase by 1 respectively
^: January temperature increases by 2 and July temperature increases by 1
Note: Y axis represents the predicted value of logV
Figure 9: Sensitivity Analysis of January and July temperature for Grand Canyon NP

Parks in Northern Arizona, Northern New Mexico and part of Utah tell a different story. Take Grand Canyon NP for an example. The graph indicates visits are likely to increase as temperature goes up. Considering climate features in such area, the conclusions make sense. Northern Arizona, Northern New Mexico and the part of Utah along the Colorado River are wet and very cold in the winter. It is dry and very hot in the summer, but not so much compared to the south of Arizona and New Mexico. In a sense, the winter is a little harsh while the summer is not very bad. When winter temperature increases, we can expect more people would like to go out for visitation, recreation and rest in winter. Even if the high temperature reduces some visits in summer, the positive effect in winter can overcome the negative effect since the summer is not quite bad compared to southern Arizona and southern New Mexico, such that the overall effect for parks in this area is positive. Global warming seems positive for national park visitation in this area. Let us trace back to the graphs showing effects of January and July temperature (Figure 4 and 6). Without surprise, most parks exceeding 45°F are located in the south. For them, marginal effect of January temperature is negative. For northern parks, marginal effect of January temperature is positive and it can offset the negative marginal effect of July temperature.



*: January and July temperature increase by 1 respectively
^: January temperature increases by 2 and July temperature increases by 1
Note: Y axis represents the predicted value of logV
Figure 10: Sensitivity Analysis of January and July temperature for Glen Canyon NRA

Parks with lake recreation, however, distinguish themselves from other parks in this area. Visits to Lake Mead and Glen Canyon soars up when July temperature increases while respond just mildly to the change of January temperature. The graph above displays the result of sensitivity analysis for Glen Canyon. The conclusions are reasonable. The best time for lake recreation such as boating, fishing is summer. When temperature increases, more people would like to go boating and fishing, not just for recreation, but a good way to stay away from the broil in urban areas. Conversely, in winter days the lake surface usually frosts and it is very hard for fishing, even less boating. So that we can expect the increase of January temperature does not have as much effect as the increase of July temperature. Of course, if the temperature increases so much that the lake stays flowing during the winter, the conclusions may change. Worth noting, the effect of global warming may not necessarily be optimistic for parks with lake recreation. The increase of July temperature also implies the drop of reservoir elevation. When the lake elevation drops very sharply, the negative effect may exceed positive effect and make the overall effect negative.

To prepare for latter input-output analysis, we carry out sensitivity analysis for all parks based on year 2003. This time we had some changes. We increased January temperature by 1 and July temperature by 2 instead of increasing January temperature by 2 and July temperature by 1. In another case, we increased January temperature by 2 and July temperature by 4. The reason is many climate models show a trend of more summer warming than winter warming. Worth noting, the result for parks with lake recreation is consistent with our above conclusion in sensitivity analysis. However, we got very big marginal effect for Lake Mead that may bias our results in input-output analysis. Therefore, we would like a table without parks with lake recreation, which produces more valid results. So we applied the estimates without Lake Mead and Glen Canyon for our sensitivity analysis. The following table shows our results.

Park name	year	$\log \hat{V}$	$\log \hat{V}^{\#}$	$\log \hat{V}^{\&}$
Arches NP	2003	13.38	13.51	13.63
Aztec Ruins NM	2003	11.04	11.16	11.27
Bandelier NM	2003	12.72	12.85	12.97
Bryce Canyon NP	2003	13.76	13.94	14.11
Canyon de Chelly NM	2003	13.46	13.57	13.67
Canyonlands NP	2003	12.61	12.75	12.89
Capitol Reef NP	2003	13.17	13.31	13.45
Capulin Volcano NM	2003	11.09	11.20	11.30
Carlsbad Caverns NP *	2003	13.41	13.40	13.40
Casa Grande Ruins NM *	2003	11.56	11.52	11.47
Cedar Breaks NM	2003	13.28	13.41	13.54
Chaco Culture NHP	2003	11.33	11.46	11.59
Chiricahua NM *	2003	11.32	11.31	11.30
Coronado Nmem *	2003	11.29	11.27	11.25
El Morro NM	2003	11.14	11.26	11.38
Fort Bowie NHS *	2003	9.05	9.04	9.02
Fort Union NM	2003	9.77	9.87	9.97
Gila Cliff Dwellings NM	2003	10.96	11.02	11.07
Grand Canyon NP	2003	15.22	15.33	15.43
Hovenweep NM	2003	10.18	10.32	10.45
Hubbell Trading Post NHS	2003	12.40	12.50	12.59
Joshua Tree NP *	2003	13.94	13.90	13.85
Montezuma Castle NM	2003	13.69	13.71	13.72
Natural Bridges NM	2003	11.58	11.72	11.85
Navajo NM	2003	11.29	11.40	11.52
Organ Pipe Cactus NM *	2003	12.52	12.44	12.35
Pecos NHP	2003	10.62	10.72	10.81
Petrified Forest NP	2003	13.52	13.60	13.67
Pipe Spring NM	2003	10.89	10.97	11.05
Rainbow Bridge NM	2003	12.27	12.39	12.49
Saguaro NP *	2003	13.50	13.45	13.38
Salinas Pueblo Missions NM	2003	10.63	10.69	10.75
Sunset Crater Volcano NM	2003	12.11	12.24	12.37
Tonto NM	2003	11.23	11.24	11.24
Tumacacori NHP *	2003	11.11	11.07	11.03
Tuzigoot NM	2003	11.75	11.75	11.74

Walnut Canyon NM	2003	11.76	11.87	11.98
White Sands NM	2003	13.13	13.18	13.22
Wupatki NM	2003	12.40	12.48	12.55
Zion NP	2003	14.73	14.77	14.80

#: January temperature increases by 1 and July temperature increases by 2
&: January temperature increases by 2 and July temperature increases by 4

*: the predicted value of logV is smaller than $\hat{\log V}^{\#}$ and $\hat{\log V}^{\&}$.

Table 11: Sensitivity Analysis for All Parks except Lake Mead and Glen Canyon (Estimates without Lake Mead and Glen Canyon Are Applied)

According to the table, the results are consistent with our conclusions. Most parks in the south have fewer visits when temperature increases, while global warming has a positive effect for most parks in the north.

CHAPTER EIGHT: VISITS DECOMPOSITION

One way of evaluating how climate change influences national park visitation is to attribute the change of visits to changes of different factors. So it is called visits decomposition. Consider the average visits to 42 national parks in 2003 as the average level. Table 11 analyzes and explains the deviation of the five most popular parks' visits from the average level. In other words, the deviation of visits from the average level is explained by cross section effect of explanatory variables, fixed effects and residuals.

	Lake Mead NRA	Grand Canyon NP	Glen Canyon NRA	Zion NP	Joshua Tree NP
Year	2003	2003	2003	2003	2003
Deviation of log(Visits)	2.512	1.860	1.073	1.343	0.693
Explained by changes in					
Demand factors	-0.637	0.186	0.443	-0.102	-1.056
Explained by					
logG	0	0	0	0	0
logEx	0	0	0	0	0
logpop	0	0	0	0	0
cindex	-0.654	0.204	0.465	-0.086	-1.097
mpi	0.018	-0.018	-0.022	-0.016	0.042
Park attributes	-0.471	3.688	1.638	2.082	4.273
Explained by					
area*	-1.702	1.021	0.725	0.850	2.981
estab	-0.331	0.198	-0.539	0.217	0.028
typogz	-0.001	-0.111	-0.111	-0.001	0.248
lodging	1.564	1.564	1.564	0	0
Np	0	1.016	0	1.016	1.016
Nhp	0	0	0	0	0
petrified	0	0	0	0	0
sunset	0	0	0	0	0
casa	0	0	0	0	0
cerro	0	0	0	0	0
tonto	0	0	0	0	0

Climate/environmental variables	5.072	-0.533	0.914	0.365	-0.354
Explained by					
jatemp*	1.124	-0.560	0.115	0.751	1.156
jasun*	-0.607	-0.322	-0.322	0.150	-0.189
jutemp*	-0.372	0.267	-0.227	-0.155	-0.281
juhumid*	-1.001	0.083	0.083	-0.381	-1.040
dry_extreme	0	0	0	0	0
REindex	-45.028	0	-46.661	0	0
juRE	50.956	0	47.927	0	0
Unknown fixed effects (η)	0.258	0.118	-0.018	0.630	-0.684
Explanatory variables in all	4.223	3.459	2.977	2.975	2.179
Residual	-1.710	-1.599	-1.904	-1.632	-1.486

*: the effect of squared term is included

Table 12: Visits Decomposition across Section

From the table, we can see the total deviation of log(Visits) is 1.073 for Glen Canyon, in which, variables of demand factors account for 0.443; variables of park attributes account for 1.638; climate/environmental variables account for 0.914; fixed effect η accounts for -0.018 and the rest is explained by residual effect. Actually it is not hard to see the variables of park attributes explain the most part of the deviation in cross section analysis. However, Lake Mead is an exception, whose deviation is mainly caused by climate/environmental variables. For all parks, the deviation explained by log(gas_price), log(exchange_rate) as well as log(population) is equal to 0. Since Gas_price, exchange_rate and population are constant cross section and are only time variant, they only show their effect in cross year analysis but all equal to 0 in cross section analysis.

Year	Lake Mead NRA 1999-2003	Grand Canyon NP 1999-2003	Glen Canyon NRA 1999-2003	Zion NP 1999-2003	Joshua Tree NP 1999-2003
Deviation of log(Visits)	-0.131	-0.104	-0.341	0.004	-0.025
Explained by changes in					
logG	-0.052	-0.052	-0.052	-0.052	-0.052
logEx	-0.015	-0.015	-0.015	-0.015	-0.015
logpop	0.071	0.071	0.071	0.071	0.071
mpi	0.008	0.004	0.004	0.004	0.009
Np	0	0	0	0	0
petrified	0	0	0	0	0
sunset	0	0	0	0	0
casa	0	0	0	0	0
cerro	0	0	0	0	0
tonto	0	0	0	0	0
dry_extreme	0	0	0	0	0
REindex	2.563	0	1.003	0	0
juRE	-2.881	0	-1.023	0	0
Unknown fixed effects (η)	0.314	-0.005	-0.069	-0.005	-0.011
Explanatory variables in all	0.007	0.003	-0.082	0.003	0.002
Residual	-0.138	-0.107	-0.259	0.0004	-0.028

Table 13: Visits Decomposition across Year

Table 13 shows the deviation of visits comparing 2003 and 1999. Since all the time-invariant variables are constant within section except mpi, NP, REindex and juRE, we can expect they have no effect in cross year analysis. So here only lists time variant variables plus mpi, NP, REindex and juRE. We can see the deviation of log(Visits) for Grand Canyon NP is -0.104, in which, log(Gas_price) explains -0.052; log(exchange_rate) accounts for -0.015; log(population) and mpi have positive effect of 0.071 and 0.004 respectively; the unknown fixed effects account for -0.005 and the residual explains the remaining part. According to the table, the demand factors and the unknown fixed effects account for the most important part of visits' variability in cross year analysis. Worth noting, Lake Mead becomes an exception again. The deviation of visits to Lake Mead is mainly explained by climate/environmental variables and the unknown fixed effects.

CHAPTER NINE: INPUT-OUTPUT ANALYSIS

In order to show the impact of the climate change on local economy, input-output analysis is carried out. We have shown that from 2002 to 2003, the decline of reservoir level in Lake Mead NRA can cause a decrease of 653,035 visitors. Based on the 2003 public report of Lake Mead NRA from National Public Service, we estimate the possible impact of the decreased amount of visitors on local economy. The following table shows the result, calculated proportionally to the figures presented by NPS.

*Sectors	Sales	Personal Incomes	Jobs	Value Added
Direct Effects				
Motel, Hotel, B&B and Cabins	$2,234,099	$648,450	57	$985,874
Campsites	$1,533,674	$444,675	39	$676,500
Restaurants & Bars	$4,353,522	$1,371,149	136	$1,909,874
Admissions & Fees	$2,555,848	$878,624	75	$1,437,974
Retail	$2,934,523	$1,497,374	104	$2,338,048
Others	$975,974	$369,600	17	$537,075
Total	$14,587,640	$5,209,872	429	$7,885,345
Secondary Effects	$4,605,147	$1,545,224	71	$2,850,373
Total Effects	$19,192,787	$6,755,096	500	$10,735,718

*: Decrease of 653,035 Visitors with 1.47% Drop of Reservoir Elevation in Lake Mead NRA
Table 14: Input-Output Analysis of Reservoir Elevation Decline for Lake Mead NRA, 2003

With a brief view of the table, we can see when the reservoir level drops by 1.47%, it can finally cause a decrease of $19,192,787 in sales, a decrease of $6,755,096 in personal incomes and a decrease of $10,735,718 in value added industry to local economy. In addition, the drop of the reservoir level can even cause a loss of 500 jobs in communities around. Considering the income of communities around the national parks mainly depend on travel industry, the numbers are amazing. Worth noting, since temperature is regarded as time-invariant variable in our model, while sensitivity analysis suggests the change of July temperature should have much influence on visits for parks with lake recreation, our result only indicates how the drop of reservoir elevation can influence local economy, which might not be the overall effect of climate change on local economy and community employment. However, the number still tells us how closely the economy is related with environment and why it is important for parks and communities around to monitor lake levels and maintain a good environment for lake recreation.

Continuing from Table 11 of sensitivity analysis, we use the result of January temperature increase by 1 and July temperature increase by 2 to conduct our input-output analysis. Economic data also come from NPS 2003 public report and we calculate the impacts of climate change on local economy proportionally. The following table displays our results for all parks except Lake Mead and Glen Canyon.

	Recreation Visits			**Spending**	**Direct Effects**			**Total Effects**		
Park Name	Visits Change (%)	Visits Change	Party Days	Total Visitor Spending (millions)	Sales (millions)	Jobs	Income (millions)	Sales (millions)	Jobs	Income (millions)
Arches NP	13.3%	100,410	49,207	$8.45	$7.17	198	$2.47	$9.53	233	$3.27
Aztec Ruins NM	11.9%	5,047	1,755	$0.14	$0.11	3	$0.04	$0.16	3	$0.06
Bandelier NM	12.8%	36,627	16,210	$1.42	$1.15	29	$0.42	$1.66	36	$0.60
Bryce Canyon NP	17.8%	161,262	94,210	$7.26	$5.98	171	$2.02	$7.95	201	$2.68
Canyon de Chelly NM	10.7%	92,810	47,759	$4.26	$3.53	88	$1.27	$5.14	110	$1.85
Canyonlands NP	14.4%	55,852	33,571	$2.37	$1.86	53	$0.63	$2.47	62	$0.84
Capitol Reef NP	14.4%	76,876	43,948	$3.33	$2.73	78	$0.93	$3.63	92	$1.23
Capulin Volcano NM	10.4%	6,353	2,359	$0.14	$0.10	3	$0.04	$0.14	3	$0.05
Carlsbad Caverns NP	-0.2%	-769	-366	-$0.03	-$0.03	-1	-$0.01	-$0.04	-1	-$0.01
Casa Grande Ruins NM	-4.3%	-3,748	-1,499	-$0.10	-$0.07	-2	-$0.03	-$0.09	-3	-$0.03
Cedar Breaks NM	13.2%	75,398	37,771	$2.33	$1.88	54	$0.64	$2.48	63	$0.85

Chaco Culture NHP	13.6%	11,485	4,763	$0.24	$0.19	5	$0.07	$0.28	6	$0.10
Chiricahua NM	-0.5%	-245	-136	-$0.01	-$0.01	0	$0.00	-$0.01	0	$0.00
Coronado Nmem	-1.6%	-1,460	-730	-$0.04	-$0.04	-1	-$0.01	-$0.05	-1	-$0.02
El Morro NM	12.2%	7,079	2,935	$0.22	$0.17	5	$0.06	$0.22	6	$0.08
Fort Bowie NHS	-1.4%	-115	-58	-$0.01	$0.00	0	$0.00	-$0.01	0	$0.00
Fort Union NM	10.1%	1,311	556	$0.03	$0.02	1	$0.01	$0.03	1	$0.01
Gila Cliff Dwellings NM	5.7%	2,739	1,119	$0.06	$0.05	1	$0.02	$0.06	2	$0.02
Grand Canyon NP	10.6%	438,246	170,522	$35.92	$31.71	640	$13.09	$46.11	830	$18.04
Hovenweep NM	14.0%	4,175	2,128	$0.19	$0.15	4	$0.05	$0.22	5	$0.08
Hubbell Trading Post NHS	9.6%	15,795	7,899	$0.70	$0.58	14	$0.21	$0.84	18	$0.30
Joshua Tree NP	-3.9%	-50,564	-23,664	-$1.89	-$1.54	-35	-$0.58	-$2.39	-45	-$0.89
Montezuma Castle NM	1.6%	10,292	5,147	$0.45	$0.37	9	$0.14	$0.55	12	$0.20
Natural Bridges NM	13.5%	13,300	6,819	$0.59	$0.49	12	$0.18	$0.71	15	$0.26
Navajo NM	11.9%	9,124	4,619	$0.40	$0.33	8	$0.12	$0.48	10	$0.18
Organ Pipe Cactus NM	-8.1%	-22,444	-11,643	-$1.00	-$0.82	-20	-$0.30	-$1.20	-26	-$0.43
Pecos NHP	10.0%	3,822	1,316	$0.07	$0.06	2	$0.02	$0.07	2	$0.03
Petrified Forest NP	7.9%	46,297	25,936	$2.85	$2.41	59	$0.86	$3.52	75	$1.26
Pipe Spring NM	8.1%	4,560	2,280	$0.20	$0.17	4	$0.06	$0.24	5	$0.09
Rainbow Bridge NM	11.2%	11,036	5,519	$0.49	$0.40	10	$0.15	$0.58	13	$0.21
Saguaro NP	-5.8%	-37,141	-20,815	-$2.29	-$1.93	-47	-$0.69	-$2.82	-60	-$1.01
Salinas Pueblo Missions NM	6.1%	2,057	857	$0.07	$0.05	2	$0.02	$0.07	2	$0.02
Sunset Crater Volcano NM	13.1%	20,800	10,402	$0.92	$0.76	19	$0.27	$1.10	24	$0.40
Tonto NM	1.2%	707	353	$0.03	$0.03	1	$0.01	$0.04	1	$0.01
Tumacacori NHP	-3.6%	-1,880	-940	-$0.06	-$0.05	-1	-$0.02	-$0.06	-2	-$0.02
Tuzigoot NM	-0.4%	-440	-220	-$0.02	-$0.02	0	-$0.01	-$0.02	-1	-$0.01
Walnut Canyon NM	11.1%	12,399	6,201	$0.55	$0.45	11	$0.16	$0.66	14	$0.24
White Sands NM	4.7%	22,916	8,544	$1.01	$0.85	24	$0.28	$1.13	28	$0.38
Wupatki NM	7.8%	20,732	10,368	$0.91	$0.76	19	$0.27	$1.10	24	$0.40
Zion NP	3.8%	94,039	36,082	$3.09	$2.46	62	$0.90	$3.55	77	$1.29

Table 15: Input-Output Analysis for All Parks except Lake Mead and Glen Canyon with January temperature increase by 1^{o}F and July temperature increase by 2^{o}F

Communities in US southwest are usually characterized with low-density population and those around national parks mainly depend on travel industry and visitor spending. From the table, we can see among parks with visits' increase, Grand Canyon has an increase of 46.11 million dollars in sales and an increase of 18.04 million dollars in income. Besides, global warming brings about 830 more job positions for Grand Canyon. For Fort Union, who benefits the least from global warming, receives $30,000 more in sales and $10,000 more in income. In an average level, global warming contributes 3.16 million dollars in sales and 1.17 million dollars in income to parks. It helps add about 66 job positions for parks with visits' increase averagely. If global warming means good news for parks in the northern area, the southern parks become the losers. Saguaro sacrifices the most with 60 job positions, a decrease of 2.82 million dollars in sales and a decrease of 1.01 million dollars in income. Averagely, the southern ten parks lose 14 job positions, $670,000 in sales and $240,000 in income. The loss is relatively small compared to the gain. Thus, the overall effect of global warming is positive for southwestern national parks. However, this table provides a clear view of what role the climate/environmental change can serve in the change of visits, which is especially meaningful for those parks who are suffering decreasing visits.

CHAPTER TEN: CONCLUSION

It is common sense that climate change would influence people's life. This paper specifically analyzes how climate/environmental change impacts people's visitation to southwestern national parks from 1979 to 2003.

The data, from 42 southwestern national parks spanning from 1979 to 2003, has typical features of TSCS data. Thus, the method to be applied in the study is discussed. Summarizing Beck, Katz, *Plümper* , Troeger and other researchers' findings, FEVD seems to be the best method for our analysis. In the latter section, the comparison of FEVD to pooled-OLS, pooled-OLS with PCSE as well as ordinary FE verifies our choice. There are 28 independent variables from three large categories in the model: five from demand factors, 12 from park attributes and 11 belonging to climate/environmental variables. By view of time varying or not, the 28 explanatory variables are classified as 9 time-varying variables and 19 time-invariant variables, which is a premise for FEVD analysis.

In all categories of variables, the results of the study are consistent with other researchers' findings (Johnson and Suits, Pergams and Zaradic, Hanink and White, Hanink and Stutts, Weiler and Seidl, Mendelsohn and Markowski, Loomis and Crespi). The study explores more variables and has new findings in each category. For demand factors, close distance between national parks does not necessarily mean competition and has negative effect for visitation. For some parks, the close distance means more complementary effect rather than competition effect. Referring to the section of visits decomposition analysis, we can see not all competing destination index plays a negative role in the increase of visits. For park attributes, we explore dummy variables accounting for anomaly and add a new signaling dummy variable for National Historical Park. The results show that anomalies in a park do not necessarily influence visits. A bigger fire like Cerro Grande fire is more likely to influence visitation. Otherwise small fires are negligible. The signaling variables prove to not only have positive effect but also have negative effect. The last but most important, the study reveals how climate/environmental change can influence southwestern national park visitation. Mendelsohn and Markowski as well as Loomis and Crespi both prove the positive effect of July temperature on visits to parks featured with lake recreation. Our study not only verifies their conclusion, but further reveals some new rules. (i) For parks with very hot and dry climate, global warming has negative effect for visits in both winter and summer, and causes tourists to decrease. (ii) For parks with cold and wet winter but also with hot and dry summer, visits are likely to increase with the increase of winter temperature. Although the increase of summer temperature can reduce visits, the overall effect of global warming is positive on visits. (iii) For parks featured with lake recreation, visits rise a lot when summer temperature increases but have only mild response to the increase of winter temperature. These rules show how climate/environmental change can have significant impact on visits. In addition to temperature, humidity is another key factor to influence tourists. US southwest are much drier than other parts of the country, especially during the summer. It possibly explains why visits are likely to increase with the increase of July humidity. Another important part of the research studies the relationship between reservoir elevation and visits. U.S. southwest is usually very dry and lack of water. Thus, lakes become very popular for recreation. Our results indicate when lake’s reservoir elevation decreases, visits also decrease. Besides, it seems reservoir level has much more influence on visits for Lake Mead than for Glen Canyon. Based on the results, we further carry out input-output analysis for Lake Mead and it is a surprise to see how the drop of reservoir elevation caused by climate/environmental change is closely related with local economy. The table of input-output analysis for all parks without lake recreation tells a similar story. Contrary to people’s common sense, global warming actually benefits the US southwestern national park visitation and local economies as a whole. However, for southern parks in this area, if their loss cannot be offset by the positive effect of demand factors and park attributes, climate change throws them a problem of how to cope with the decreasing visits, sales and income as well as the loss of job positions, while their communities mainly depend on travel industry and visitor spending.

Admittedly, there are several regrets and shortages for the study. (i) The effect of January sunlight on visits is negative, which is not consistent to common sense. Up to now, we are not quite clear about the relationship between January sunlight and visits to US southwestern national parks. (ii) For many parks, the model works well; for a few others, the model does not have a good prediction, especially for parks with lake recreation. The comparison of the econometric results with and without Glen Canyon and Lake Mead possibly indicates sufficient data is needed to get more findings about parks with lake recreation. Referring to Table 12 and 13 in the section of visits decomposition analysis, error terms also give us some insights. So it seems the parks with lake recreation need more studies. However, due to limited time and knowledge, it waits for future studies and other researchers' discovery.

REFERENCES

1. Beck N. and J. Katz. "What to Do (and Not to Do) with Time-Series Cross-Section Data." *American Political Science Review* 89, no. 3 (1995):634-647.

2. Beck N. "Time-Series-Cross-Section Data: What Have We Learned in the Past Few Years?" *Annual Review of Political Science* 4 (2001):271-293.

3. Beck N. and J. Katz. "Time-Series-Cross-Section Issues: Dynamics, 2004." *Draft of July 24, 2004* (2004):1-35.

4. Freedman D. and S. Peters. "Bootstrapping a Regression Equation: Some Empirical Results." *Journal of the American Statistical Association* 79 (1984):97-106.

5. Frisvold G. and K. Ingram. "Sources of Agricultural Productivity Growth and Stagnation in Sub-Saharan Africa." *Agricultural Economics* 13 (1995):51-61.

6. Goldberg L. "Industry-Specific Exchange Rates for the United States." *Economic Policy Review* 10, no. 1 (2004):1-16.

7. Hanink D. and K. White. "Distance Effects in the Demand for Wildland Recreational Services: the Case of National Parks in the United States." *Environment and Planning A* 31 (1997):477-492.

8. Hanink D. and M. Stutts. "Spatial Demand for National Battlefield Parks." *Annals of Tourism Research* 29, no. 3 (2002):707-719.

9. Jia S. and S. Rathi. "On Predicting Log-transformed Linear Models with Heteroscedasticity." *SAS Global Forum 2008: Statistics and Data Analysis* (2008):1-6.

10. Johnson R. and D. Suits. "A Statistical Analysis of the Demand for Visits to U.S. National Parks: Travel Costs and Seasonality." *Journal of Travel Research* 22 (1983):21-23.

11. Loomis J. and J. Crespi. "Estimated Effects of Climate Change on Selected Outdoor Recreation Activities in the United States." *The Impact of Climate Change on the United States Economy* (1999):289-311.

12. Mendelsohn R. and M. Markowski. "The Impact of Climate Change on Outdoor Recreation." *The Impact of Climate Change on the United States Economy* (1999):267-288.

13. Parks R. "Efficient Estimation of a System of Regression Equations When Disturbances Are Both Serially and Contemporaneously Correlated." *Journal of the American Statistical Association* 62 (1967):500-509.

14. Pergams O. and P. Zaradic. "Is Love of Nature in the US Becoming Love of Electronic Media? 16-Year Downtrend in National Park Visits Explained by Watching Movies, Playing Video Games, Internet Use and Oil Prices." *Journal of Environmental Management* 80 (2006):387-393.

15. *Plümper* T. and V. Troeger. "The Estimation of Time-Invariant Variables in Panel Analyses with Unit Fixed Effects." (2004):1-33.

16. *Plümper* T. and V. Troeger. "Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects." *Political Analysis* 15 (2007):124-139.

17. Ponnaluru S. "Visitation to the National Parks in the Southwest: the Influence of Economic and

Climate Variables." *Master Thesis, Department of Agricultural and Resource Economics, University of Arizona* (2005):1-65.

18. USNPS. "Visitor Spending and Economic Impacts for National Park Units" http://web4.canr.msu.edu/mgm2/. 2003.

19. Western Regional Climate Center. "Western US COOP Map" http://www.wrcc.dri.edu/coopmap/. 2008.

20. Weiler S. "A Park by Any Other Name: National Park Designation as a Natural Experiment in Signaling." *Journal of Urban Economics* 60 (2006):96-106.

21. Weiler S. and A. Seidl. "What's in a Name? Extracting Econometric Drivers to Assess the Impact of National Park Designation." *Journal of Regional Science* 44, no. 2 (2003):245-262.

22. Weiler S., J. Loomis, R. Richardson and S. Shwiff. "Driving Regional Economic Models with a Statistical Model: Hypothesis Testing for Economic Impact Analysis." *Review of Regional Studies* 32 (2000):97-111.

23. Wicks B., M. Uysal and S. Kim. "The Effect of Lodging Prices on Visitors' Demand: Everglades National Park." *Journal of Hospitality & Tourism Research* 17 (1994):51-61.