A Revealed Preference Approach to the Measurement of Congestion in Travel Cost Models

by

Christopher Timmins

and

Jennifer Murdock

PRELIMINARY AND INCOMPLETE

March 8, 2005

Abstract

Travel cost models are regularly used to determine the value of recreational sites or particular site characteristics, yet a key site attribute, congestion, is typically excluded from such analyses. One of several reasons is that congestion is determined in equilibrium by the sorting process itself, and thus presents significant endogeneity problems. This paper illustrates the source of endogeneity, describes how previous research has dealt with it by way of stated preference techniques, and describes an instrumental variables approach to address endogeneity in a revealed preference empirical setting. We apply this technique to the valuation of a large recreational fishing site in Wisconsin (Lake Winnebago) which, if eliminated, would induce significant re-sorting of anglers amongst remaining sites. We demonstrate that failing to address the endogeneity of congestion is likely to lead to the understatement of its disutility or possibly to the mistaken recovery of agglomeration benefits. In this application to Lake Winnebago, ignoring congestion leads to an understatement of its value of more than 34 percent.

Keywords: Congestion, Random Utility Model, Travel Cost, Discrete Choice, Instrumental Variables, Quantile Regression

---

The authors gratefully acknowledge the helpful insights received from Han Hong. The authors also wish to thank Triangle Economic Research for the use of the Wisconsin fishing data, and Doug MacNair and Bill Desvousges in particular.
Contact Information–Timmins: *Department of Economics, Duke University, PO Box 90097, Durham, NC 27708. E-mail: christopher.timmins@duke.edu*. Murdock: *Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7. E-mail: jennifer.murdock@utoronto.ca*.

1. Introduction

Random utility models (RUMs) of recreation deεεmand exploit the information in the trade-offs individuals make between travel time and site attributes to value the latter. Consider the case of recreational fishing. Applications typically include data on attributes such as expected fish catch, urban and industrial development, water quality, and amenities like paved boat ramps and fishing piers. The RUM has become a staple of the legal and policy communities because it provides a convenient tool for attaching values to non-marketed commodities (e.g., water quality) that might be the subject of litigation or environmental policy debates.

One important attribute that is conspicuously absent from nearly every such study and particularly those based on revealed preference techniques is congestion. Measures of congestion describe the number of other individuals encountered during the recreation experience.[1] For activities like hunting, hiking, camping, fishing, and beach use, congestion is likely an important attribute of the recreation experience. When congestion is not included in the estimation of a RUM, three important things happen. (i) The role of congestion as an effective rationing device is ignored. This can have implications for the proper design of policy. (ii) Congestion becomes an omitted variable that will lead to biased estimates of the value of other attributes with which it is correlated. (iii) The ability to accurately value entire sites is compromised, especially when those sites are large and their closure induces significant resorting among remaining sites.

This paper represents the first attempt to address congestion empirically using revealed preference techniques without basing identification on functional form

[1] There are a number of papers that deal specifically with the question of how to define congestion. We describe these briefly in Section 2.

assumptions. It does so by relying on a previously unexploited source of variation in the data – the isolation of alternative sites in exogenous attribute space. This proves useful in properly controlling for congestion, which is otherwise a difficult task. What makes this problem difficult is that variables describing the equilibrium behavior of other individuals in the site-choice problem are typically endogenous. Without properly accounting for that source of endogeneity, there is a natural tendency to understate the role of congestion and to even improperly recover estimates of agglomeration effects. In this paper, we describe the source of this endogeneity, cast it as a simple instrumental variables problem in a familiar linear regression context, and demonstrate how it can be solved in an application to Wisconsin recreational fishing. We then use our estimates to demonstrate how ignoring congestion can lead to significant biases in measuring the value of a large site.

After a brief review of the literature on the role of congestion in travel cost models in Section 2, we describe our model of site selection with congestion in Section 3. Here, we make clear the source of difficulty in estimating the marginal disutility of congestion, and we show how the problem can be solved by casting it in a familiar instrumental variables framework. In Section 4, we describe the data set we use in an application of this technique. In Section 5 we discuss an econometric complication that arises when we model congestion effects depending on whether it is a weekday versus weekend. Section 6 reports model estimates, and Section 7 shows how congestion impacts a site valuation exercise. Section 8 concludes.

2. Previous Literature

That congestion costs could be an important determinant of behavior in models of site selection has long been recognized. We categorize papers on the topic into three groups – one theoretical and two that are primarily empirical. The set of theoretical papers describe important issues that will motivate our modeling exercise. Anderson and Bonsor (1974) are one of the first to discuss the implications of congestion for measuring willingness to pay, while Fisher and Krutilla (1972) note that optimal management of a recreation site requires a charge that incorporates both marginal congestion and environmental costs. Cesario (1980) introduces the primary issue we address in our empirical application – that one cannot recover unbiased estimates of the value of a recreation site without accounting for the equilibrium resorting. The removal of a recreational site affects the welfare of users of other sites as displaced recreationists re-sort across the remaining sites. If congestion costs are ignored, there is a tendency to understate the value of new site construction. In a more recent paper, Jakus and Shaw (1997) discuss ways to measure congestion, emphasizing the need to use the value individuals expect at the time they make their site decision rather than, for example, an *ex post* realization of congestion. A similar point is made by Schuhmann and Schwabe (2004), who also emphasize the timing of congestion costs. For example, this could entail differentiating between the expected number of other visitors on a weekday versus a weekend. Michael and Rieling (1997) emphasize the role of heterogeneous preferences for congestion in inducing recreators to sort over days of the week.

Empirical work on congestion in site valuation can generally be divided into studies based on stated versus revealed preference data. Cichetti and Smith (1973) measure the effect of "wilderness encounters," which is congestion in the hiking context, on stated willingness to pay, with an application to the Spanish Peaks Primitive Area in Montana. McConnell (1977) employes stated preference techniques to estimate the role of congestion in the demand for beach recreation and uses the results to derive net surplus maximizing projects. Boxall, Rollins, and Englin (2003) similarly use a stated preference model to value congestion in four separate components of a back-country canoeing trip, emphasizing that the estimate of distaste for congestion may be very different depending upon the specific activity under consideration.

In this paper, we adopt a revealed preference approach to measuring the costs of congestion. It is important to comment, however, on the value of stated preference data. In particular, asking someone how much they would hypothetically contribute in order to avoid additional congestion avoids the difficulties associated with the fact that actual congestion is determined by the optimizing decisions of other recreators facing similar choices. This falls into the general class of problems associated with endogenous sorting models. [Bayer and Timmins (2005a)] In such models, congestion is likely to be correlated with unobservables that also drive the behavior of the decision-maker. We will demonstrate in the following section that failing to account for this sort of endogeneity will typically lead to the understatement of congestion costs or possibly even the mistaken diagnosis of agglomeration benefits. Stated preference models avoid this problem by hypothetically varying congestion while holding the unobservables that drive sorting behavior constant. The downside of this solution is, of course, that stated preference models value hypothetical changes, and respondents may not reveal their true preferences in response to hypothetical questions.

There have been very few papers that have addressed the problem of valuing congestion from a revealed preference perspective. Boxall and Adomowicz (2000) conduct both a stated preference analysis (finding small negative effects of congestion) along with estimating a revealed preference model that uses fitted values of perceived congestion from a first-stage estimation procedure. That procedure is based on survey data describing *a priori* perceived congestion and observed site attributes from actual recreation experiences. We show below that, while using fitted values for perceived congestion mechanically breaks the correlation between the congestion variable and unobserved site attributes, the application in Boxall and Adomowicz (2000) does not introduce any determinants of expected congestion that do not already appear in the site selection model. The ability of their approach to identify a congestion effect therefore relies exclusively on the non-linearity introduced by the choice of an ordered logit functional form in the first-stage prediction of expected congestion. As always, the results of a model identified by functional form assumptions can prove to be highly sensitive to those assumptions. Considered in the context of Boxall and Adomowicz (2000), our model can be considered as a data-based as opposed to a functional form driven approach to overcoming this identification problem.

In addition to the role of congestion in models of site selection, this paper also touches on a number of other literatures. Our application to the recreational fishing behavior of Wisconsin anglers builds upon a long line of research using random utility models and travel costs to value site attributes. Bockstael et al (1989) provide one of the earliest published applications of the RUM to recreation demand in their valuation of catch improvements for Florida sportfishing. Subsequent research has considered the sensitivity of the random utility model to a number of data handling and modeling decisions such as the definition of sites, the definition of the choice set, and the assumed error structure. During the last decade researchers have relaxed some of the strict assumptions on the error structure. Nested logit specifications, which allow for correlations among the unobservables for groups of alternatives, and random parameters specifications, which allow individual preferences for site characteristics to be heterogeneous, have become the norm.

Finally, for reasons that will be made clear in Section V, applying our empirical strategy will require the use of instrumental variables techniques adapted to estimation in a quantile regression framework. Recent work has produced a number of approaches to this problem. [Hong and MaCurdy (1999), Chernozhukov and Hansen (2001), Imbens and Newey (2003), Ma and Koenker (2003)] The methods proposed by Hansen and Chernazukov (2001) and Hong and MaCurdy (1999) prove to be particularly well-suited to our context.

3. Model

Modeling congestion in a RUM framework is akin to describing a Nash bargaining model in which individuals make site choices given their expectations about the choices that will be made by other individuals. In equilibrium, all of those expectations are confirmed by other individuals' actual decisions. We therefore begin with the site choice decision of an individual angler *i* on choice occasion *t*. A choice occasion is defined to be a fishing trip, which means that the following is a model of site-choice conditional on the angler choosing to take a trip. The participation decision, the choice of whether or not to take a trip, is not modeled but is conceptually simple to add but computationally challenging. The utility obtained from choosing site *j* in time period *s* is given by

(1) $$U_{ijts} = \delta_{js} + X'_j \Gamma_s(Z_i) + \Phi_s(Z_i)\sigma_{js} + \Theta_s(Z_i)\ln TC_{ij} + \varepsilon_{ijt}$$

where

(2) $$\delta_{js} = X'_j \beta + \alpha\sigma_{js} + \xi_{js}$$

(3) $$\Gamma_s(Z_i) = Z'_i\gamma_s \qquad \Phi_s(Z_i) = Z'_i\phi_s \qquad \Theta_s(Z_i) = \theta_{0,s} + Z'_i\theta_{1,s}$$

and

Z_i	=	observable attributes of angler *i*
X_j	=	observable attributes of fishing site *j*
TC_{ij}	=	travel cost incurred by angler *i* in visiting site *j*
ξ_{js}	=	unobservable attribute of site *j* in time period *s* (common to all anglers)[2]
ε_{ijt}	=	idiosyncratic source of utility for angler *i* at site *j* on choice occasion *t*
σ_{js}	=	expected share of all anglers choosing site *j* in time period s

δ_{js} represents the baseline utility from site *j*, which is what an individual with $Z_i = 0$ would receive, except for the common component of the marginal utility of travel costs, $\theta_0 lnTC_{ij}$).

Anglers are ascribed rational expectations about the behavior of their fellow anglers. This means that the vector of expected shares will be constant across anglers and equal to the actual share. Practically, this assumption is consistent with the idea that anglers have repeatedly played the site-selection game with one another and have reached a Nash equilibrium.

[2] X_j includes the observed site characteristics, which are fixed over time and across anglers in the available data. It is likely, however, that site attributes not described in available data may be very different at different times such as on a weekday versus weekend. We therefore allow for this possibility in the way we define our unobservables.

Simplifying Assumptions

In our application, we distinguish between congestion on weekdays (WD) and weekends (WE). Within each of these groups, we treat each site selection choice made by an angler as an independent event. We therefore calculate two sets of expected shares, $\sigma_{j,WD}$ and $\sigma_{j,WE}$, by taking the share of all visits that are made to each site, and we estimate a separate set of preference parameters for each of these groups. This approach is flexible in that it allows the way in which attributes are combined into utility to differ depending upon whether it is a weekday or a weekend trip. However, we clearly introduce an unrealistic assumption in that we see the same angler make repeated decisions over the course of a fishing season (some of which may fall on weekends and some of which may fall on weekdays), but do not exploit this information. We could, for example, also model the decision about which day of the week to go fishing. Other techniques have been developed to deal with persistence in an individual's site selections. [****Add Cites****] While these techniques could be easily incorporated into the modeling framework presented below, they are not the focus of the current application and are ignored here for simplicity's sake.

We set up the problem as a heterogeneous parameters discrete choice model, allowing preferences for all observable attributes (including congestion and travel cost) to vary with observable individual attributes Z_i. A random parameters logit model, which allows for additional heterogeneity in the taste parameters based on unobserved individual attributes, could be easily incorporated into our modeling framework. Increased flexibility in individual preferences is not the focus of our paper, however, and the random component of preference parameters is omitted in order to simplify the estimation procedure.

Equilibrium

Each angler will maximize his or her utility given expectations about the behavior of other anglers. In equilibrium, those expectations will be validated. We assume that the idiosyncratic unobservable component of utility, ε_{ijt}, is distributed i.i.d. Extreme Value. This means that we can write the probability of seeing angler *i* choose location *j* on choice occasion *t* as:

(4)

$$P(U_{ijts} \geq U_{ilts} \ \forall\, l \neq j) \quad = \quad \frac{EXP\{\delta_{js} + X_j \Gamma_s(Z_i) + \Phi_s(Z_i)\sigma_{js} + \Theta_s(Z_i)\ln TC_{ij}\}}{\sum_{l=1}^{J} EXP\{\delta_{ls} + X_l \Gamma_s(Z_i) + \Phi_s(Z_i)\sigma_{ls} + \Theta_s(Z_i)\ln TC_{il}\}}$$

Integrating over the distribution of angler attributes, $F(Z_i)$, we can predict the share of anglers who will end up choosing each site on a weekday (s = WD) and a weekend (s = WE):

(5) $$\sigma_{js} = \int P(U_{ijts} \geq U_{ilts} \ \forall\, l \neq j)\, dF(Z_i) \quad \forall\, j$$

It is a straightforward application of Brower's fixed point theorem to show that there exists a vector of $\sigma_{j,WD}$'s and $\sigma_{j,WE}$'s that satisfy the contraction mapping implied by (5). Whether these equilibria are unique or not is a more complicated question that depends upon the degree of effective variation in the observed choice attributes.[3] [Bayer and Timmins (2005b)]

Estimation

While important for counterfactual simulations, uniqueness is not necessary to estimate the parameters of equation (1) by maximum likelihood. In particular, we can write the time-period-specific likelihood of observing a vector of site choices:

(6) $$L_s(\bar{\delta}_s, \bar{\gamma}_s, \bar{\phi}_s, \bar{\theta}_s \mid \bar{Z}, \bar{X}, \overline{TC}, \bar{Y}) \;=\; \prod_{i \in N_s} \prod_{t=1}^{T_s} \prod_{j=1}^{J} [P(U_{ijts} \geq U_{ilts} \;\forall\, l \neq j)]^{Y_{ijt}}$$

where N_s represents the set of all angler trips taken in time period s, and Y_{ijt} equals 1 if angler *i* chooses location *j* on choice occasion *t* and equals 0 otherwise. Maximizing equation (6) with respect to the vector $(\bar{\delta}_s, \bar{\gamma}_s, \bar{\phi}_s, \bar{\theta}_s)$ gives us estimates of baseline utility for each site (δ_{js}), along with parameters describing how utility for various site attributes varies with observable angler attributes.[4]

Note the role of the congestion variable at this stage of the estimation procedure. Specifically, one might worry about the potential endogeneity of the share of other anglers choosing a particular site in a particular time period. As will be shown below,

---

3 "Effective variation" in the choice set implies both that choices are different in observable and unobservable dimensions, and that individuals care about those differences – i.e., significant differences in attributes over which individuals are indifferent will do nothing to help achieve uniqueness in the sorting equilibrium.

4 Given the large number of potential alternatives from which individuals can choose (569 in the current application), recovering the full set of δ_{js}'s by searching over the likelihood function can be computationally prohibitive. We therefore employ the contraction mapping technique outlined by Berry (1994) and used in Berry, Levinsohn, and Pakes (1995). The idea of this technique is to choose values for $(\bar{\gamma}_s, \bar{\phi}_s, \bar{\theta}_s)$, and then find the vector of δ_{js}'s that make the predicted share of individuals choosing each alternative exactly equal the actual share. This is easily done by way of a contraction mapping. As the likelihood maximization procedure searches over alternative values of $(\bar{\gamma}_s, \bar{\phi}_s, \bar{\theta}_s)$, the contraction mapping procedure repeatedly updates the corresponding vector of δ_{js}'s.

this is an important concern, but one that is avoided at this stage of the estimation problem. In particular, it will likely be the case that $\sigma_{j,WD}$ and $\sigma_{j,WE}$ will be correlated with unobservable site attributes $\xi_{j,WD}$ and $\xi_{j,WE}$, respectively. Because we control for these attributes non-parametrically with fixed effects $\delta_{j,WD}$ and $\delta_{j,WE}$ at this stage of the procedure, however, this correlation is not a concern. Rather, it becomes an issue when we turn to decomposing the estimates of $\delta_{j,WD}$ and $\delta_{j,WE}$ in order to learn about the determinants of baseline utility.

Consider this decomposition problem:

(7) $$\delta_{js} = X'_j\beta + \alpha\sigma_{js} + \xi_{js}$$

for s = WD, WE. This is nothing but a linear regression problem with ξ_{js} serving as the estimation error. Equilibrium sorting implies a mechanical correlation between σ_{js} and ξ_{js}, COV[σ_{js}, ξ_{js}] > 0. Locations with desirable unobservable attributes will therefore attract more visitors and will have higher baseline utility. Without additional information, the model is unable to tell these two forces apart, and will tend to overstate the value of σ_{js}. There is a natural tendency in estimating (7) by OLS to recover an upward biased estimate of α and therefore either understating the costs of congestion or even finding benefits from agglomeration.

While not presented in this exact framework, the fundamental difficulty faced by all papers seeking to estimate congestion costs is the same. It is worth pausing at this point to consider how the previous literature on site-choice has dealt with this problem. In Section 2, we broke the literature down into two groups of papers – those that rely on stated preference versus those that use revealed preference evidence. The papers that use stated preference evidence essentially avoid this endogeneity problem by hypothetically varying σ_{js} while holding ξ_{js} constant – i.e., by asking "what would you be willing to pay to have less congestion holding everything else about the choice problem (including unobservables) the same?"[5] This is an effective strategy, but does suffer from the usual complaint about respondents' willingness to reveal their true preferences in response to hypothetical questions. The one paper that uses revealed preference data [Boxall and Adomowitz (2000)] instead solves the problem by using fitted values of σ_{js} based on predictions from an ordered logit model.[6] The problem is that those predictions are based on the same information that is contained in X_j, meaning that the α parameter is identified only from the non-linearity inherent in the ordered logit.

An Instrumental Variables Approach

In order to solve this problem, we treat equation (7) as an endogenous-regressors problem and propose an instrumental variables solution. A valid instrument in this case would be some variable that is correlated with σ_{js}, uncorrelated with ξ_{js}, and can reasonably be excluded as a determinant of δ_{js}. We propose such an instrument based on the underlying equilibrium model of sorting across sites. In particular, combinations of the exogenous attributes of sites other than *j* can provide valid instruments for the share of anglers choosing site *j*. Intuitively, this is because anglers look across available

[5] Whether or not it is possible to get respondents to, in fact, hold everything else fixed when making this comparison is a reasonable question to ask in many stated-preference applications.
[6] To be precise, Boxall and Adomowitz (2000) base congestion predictions on information about site attributes that is also used in the site selection model (X_j), as well as on individual attributes. Because individual attributes do not vary with the chosen site, however, they do not provide an independent source of variation in predicted values of congestion. Rather, predicted congestion varies across sites only with the attributes X_j – separately identifying the utility effects of congestion from that of those attributes is therefore done only by functional form assumption.

alternatives for the combination of site attributes that will maximize utility. Having a great many alternative sites with desirable attributes will, for example, reduce the share of anglers choosing a particular site j, *ceteris paribus*. In the decomposition of δ_{js}, however, the attributes of sites other than j can logically be excluded – equation (7) is a structural equation that describes a component of the utility function. There is no reason why the attributes of choices other than j should enter into the expression for the utility derived from choosing j, *except in the way they impact the share of other anglers also choosing j*. Finally, in order to constitute valid instruments, the attributes of choices other than j must be uncorrelated with ξ_{js}. Given that we assume that X_j is uncorrelated with ξ_{js} (i.e., the standard assumption in any kind of hedonic exercise), it is not difficult to further assume that X_{-j} is also uncorrelated with ξ_{js}.

Bayer and Timmins (2005a) (B-T) provide justification for a particular function of the exogenous attributes of the entire choice set as an instrument for σ_{js} in equation (7). In particular, B-T argue for using the predicted share of anglers choosing site j based only on exogenous attributes of all possible choices:[7]

(8) $$\tilde{\sigma}_{js} = \int \frac{EXP\{X_j'\hat{\Gamma}_s(Z_i) + X_j'\hat{\beta}_s + \hat{\Theta}_s(Z_i)\ln TC_{ij}\}}{\sum_{l=1}^{J} EXP\{X_l'\hat{\Gamma}_s(Z_i) + X_l'\hat{\beta}_s + \hat{\Theta}_s(Z_i)\ln TC_{il}\}} dF(Z_i)$$

If exogenous attributes are important determinants of site choice (relative to endogenously determined congestion effects), this instrument will have good power. As

[7] If one were concerned that individuals had sorted geographically in response to ξ_{js} (e.g., retirees choosing to settle close to the best fishing sites), travel cost would be endogenous and should be excluded from the formation of the instrument at this stage. If this is not a concern, however, including travel cost has the potential to greatly increase the power of our instruments.

sites become increasingly similar in exogenous dimensions, the instrument will become increasingly weak.

The obvious problem with using the instrument described in (8) lies in the fact that it requires that we already have in hand estimates of $(\gamma_s, \beta_s, \theta_{0,s}, \theta_{1,s})$, while identifying these parameters is the very goal of the IV strategy. B-T describe a procedure whereby an initial guess at $(\gamma_s, \beta_s, \theta_{0,s}, \theta_{1,s})$ can be found by estimating (6) and (7) and then ignoring the role of σ_{js} in the latter equation. With these estimates, the instruments in (8) are calculated and used in an IV estimation of equation (7) that accounts for the role of both X_j and σ_{js}. B-T provides Monte Carlo evidence on the performance of this instrumental variables strategy in a variety of empirical contexts.

4. Data

Table 1 describes and summarizes the data used in estimation. The data on angler characteristics, travel cost, and fishing site characteristics are discussed in more detail next. Murdock (2002) provides additional details about the data and data collection process.

The 1998 Wisconsin Fishing and Outdoor Recreation (WFOR) survey is a primary source of data. A random digit dial telephone survey recruited anglers willing to complete a fishing diary each month for June through September. Of the anglers completing the telephone interview, 81.0 percent agreed to participate in the diary portion of the survey. This paper focuses on the 512 anglers that reported taking a single day fishing trip. A comparison between all anglers contacted on during the telephone survey and the final sample reveals that they are very similar. These anglers report 3,581 single day fishing trips used in estimation.

Fishing sites are defined using the water body name and quadrangle.[8] Figure 1 shows a map of Wisconsin with the quadrangles marked. Each inland lake visited by an angler constitutes a separate fishing site. In quadrangles containing multiple inland lakes, each unique inland lake forms a separate fishing site. Lake Michigan, Green Bay, Lake Winnebago, and all rivers and streams are divided into quadrangles because of their large size or long length. According to this definition, there are 569 different sites visited by the sample on single day trips.

The fish catch measures vary across fishing sites but not across anglers. The detailed data available for this study allows catch to be identified separately for eight different fish species. Fish catch rates are constructed by combining information from the Wisconsin Department of Natural Resources (WDNR) and the WFOR survey. The WDNR provides information on the surface area, depth, and fish abundance (`abundant', `common', `present', and `not present') for virtually all inland lakes. The bulk of the data was collected in the 1950s and 1960s, which makes it dated, and it excludes Lake Michigan, Green Bay, streams, and rivers. The WFOR fish catch data are detailed and comprehensive: for each day spent fishing, survey participants recorded the number and species of fish they personally caught and the time spent fishing.

A weighted least squares (WLS) procedure is used to combine both sources of data and obtain a catch rate for each species at each site. A separate WLS regression is estimated for each site and species. Each regression includes all sites of similar type

[8] According to the U.S. Geological Survey, Wisconsin contains 1,154 quadrangles and each is roughly seven miles long and five miles wide.

within 50 miles. Weighting allows sites with more observed fishing trips, located nearer the origin site, and with more physical similarities to have more influence in the regression. Because the only right-hand-side variable is the WDNR measure of fish abundance, which is missing for some species and all locations that are not inland lakes, many of the WLS regressions include only a constant term and hence produce a simple weighted average of the WFOR survey data. The predicted value for each species at each site serves as the expected catch.

Table 1 also summarizes the other observed site characteristics. The Wisconsin fishing regulation booklet identifies quality fishing sites, where quality relates to the size of the fish inhabiting the waters. In general, motor trolling is not permitted in Wisconsin's waters except where expressly allowed.[9] Shoreland development may affect choice to the extent that some anglers value a natural and quiet setting. Inspection of the Delorme Atlas and Gazetteer map indicates sites that have at least a portion of their shoreland designated as urban. Map inspection also reveals which fishing sites are contained within a national forest, state forest (or park), county forest, or a wildlife area.

Figure 2 plots the shares of anglers at each fishing site by time period (i.e., weekday versus weekend). It illustrates that expected congestion at a particular site is a function of whether it is a weekday or weekend. While not showing any strong evidence of a systematic bias (i.e., sites that are more crowded on weekdays are not systematically more or less crowded on weekends), the R^2 of a regression line drawn through these data is only ***, indicating that there is a great deal of variation across weekend days and weekdays. It is therefore important that we conduct the analysis separately.

[9] Motor trolling involves trailing a lure or bait from a moving vessel (motor boat or sail boat).

5. Practical Issues in Estimation

The estimation procedure, as described in Section III, relies upon seeing positive shares of anglers choosing to visit each site in each time period in the recovery of the vector of time period specific fixed effects, $\delta_{j,WD}$ and $\delta_{j,WE}$. Practically, these fixed effects play a very important role in the estimation, as they allow for the inclusion of choice-time-period-specific unobservable attributes, $\xi_{j,WD}$ and $\xi_{j,WE}$. By virtue of the way in which the data were collected, we are assured of seeing non-zero shares for all sites across the combined weekday and weekend groups. This is not the case, however, when we consider either time period by itself.

Table 2 shows how the share of trips is spread over the 569 sites when considering only weekday trips, only weekend trips, and weekday and weekend trips combined. In total, 21.6 percent of all sites are not visited on a weekend, while 33.0 percent are not visited on a weekday. This poses a practical problem for the recovery of time period specific fixed effects. In particular, all the data tell us is that these are unattractive choices (i.e., so unattractive as to not induce a single visitor from the sample). The data give no indication, however, of exactly how unattractive these sites are.

We address this problem by first introducing a numerical "patch" that allows the contraction mapping described in Section 3 to function. This simply amounts to adding a small increment (e.g., $\varepsilon = 10^{-6}$) to the total number of visits to each site in each time period before calculating shares. This means that no shares will equal zero, although some will be very small. For very small values of ε, the effect of this patch is seen entirely in the recovered values for $\delta_{j,WD}$ and $\delta_{j,WE}$ for those sites with actual visit shares equal to zero. In particular, the smaller the value of ε that is chosen, the more negative the values of $\delta_{j,WD}$ and $\delta_{j,WE}$ become for those sites. Because very small values of ε have virtually no effect on the relative odds of any two choices with positive numbers of visitors, however, the impact on the remaining values of $\delta_{j,WD}$ and $\delta_{j,WE}$ is negligible.[10] The Appendix reports parameter estimates under three alternative assumptions about ε, and Figure A1 makes this point. In particular, it shows the estimated distribution of $\delta_{j,WD}$ under the assumption that $\varepsilon = (10^{-6}, 10^{-9}, 10^{-12})$. A series of bi-modal distributions emerges. The lower mode reflects values of $\delta_{j,WD}$ determined by the assumption about ε. For smaller values, that mode shifts further to the left. Key to our strategy, the upper portion of the distribution (i.e., that based on visited sites) does not change with alternative assumptions about ε.

We therefore require a second-stage estimator that is robust to the fact that the values of $\delta_{j,WD}$ and $\delta_{j,WE}$ for unvisited sites are arbitrarily negative. A flexible approach is to use a regression technique that does not depend upon the specific values in the lower tail of the δ_{js} distribution – i.e., quantile estimation. As long as a majority of sites have positive numbers of visitors, the median regression is well-suited to this purpose.[11]

Adapting the median regression to deal with endogenous regressors is not as simple as in the case of mean regression (OLS). It has, however, been the focus of recent work in econometric theory. [MaCurdy and Hong (1999), Chernozhukov and Hansen (2001), Imbens and Newey (2003), Ma and Koenker (2003)] This is important in our

[10] It is easy to show with Monte Carlo evidence that as $\varepsilon \to 0$, all the parameters besides $\delta_{j,WD}$ and $\delta_{j,WE}$ for the unvisited sites converge to stable values. The values of $\delta_{j,WD}$ and $\delta_{j,WE}$ for the unvisited sites, however, $\to -\infty$.

[11] Koenker and Bassett (1978) provides the original theory for quantile regression techniques. Koenker and Hallock (2001) provides a convenient summary.

context because of the presence of the endogenous regressor $\sigma_{j,WD/WE}$. We use a simple Smoothed GMM estimation approach based upon the technique described in MaCurdy and Hong (1999). In essence, assuming specifications for the quantiles of structural error distributions conditional upon exogenous or pre-determined instruments, the estimator formulates these conditional quantiles into moment conditions capable of being estimated within a conventional nonlinear instrumental variables or Generalized Method of Moments framework. This apparatus matches the sample analog of the conditional quantiles against their population values, employing a smoothing procedure familiar in various problems found in non-parametric inference and simulation estimation. The analysis applies standard arguments to demonstrate consistency and asymptotic normality of the resulting smoothed GMM quantile estimator. Simulation exercises reveal that this procedure accurately produces estimators and test statistics generated by conventional quantile estimation approaches.

To apply this GMM quantile procedure, let δ_{js} denote baseline utility from site j in time period s, and let (X_j, σ_j) denote our vector of exogenous variables and endogenously determined shares. We are interested in obtaining information about the distribution of δ_{js} conditional upon (X_j, σ_j). We will use $Q_\rho(X_j, \sigma_j)$ to represent the ρ^{th} percentile of this conditional distribution, where $\rho \in (0, 100)$. Our Smoothed GMM quantile estimator makes use of the following moment conditions, which underlie the construction of most quantile estimation procedures:

(9) $$P\left(\delta_{js} < Q_\rho(X_j, \sigma_{js}) \mid X_j, \sigma_{js}\right) = \rho$$

This relation implies the condition:

$$(10) \qquad E\left\lfloor 1\left(\delta_{js} < Q_\rho(X_j,\sigma_{js})\right) - \rho(X_j,\sigma_{js}) \right\rfloor = 0$$

where 1(•) represents the indicator function which takes value 1 when the condition expressed in parentheses is true, and 0 otherwise. The indicator function inside the moment condition is neither continuous nor differentiable. To incorporate this moment condition into the standard framework of nonlinear method of moments estimation, MaCurdy and Hong (1999) propose to use the modified smooth version of this condition:

$$(11) \qquad E\left[\lim_{N\to\infty} \Phi\left(\frac{\delta_{js} - Q_\rho(X_j,\sigma_{js})}{s_N} \right) - (1-\rho) \right] = 0$$

where N represents the sample size (569) and Φ is a continuously differentiable distribution function with bounded symmetric density function φ. The following analysis uses the cumulative standard normal distribution function, but other distributions (e.g., logit) could be used as well. The quantity s_N is a bandwidth parameter that converges to 0 as $N \to \infty$ at a rate slower than that of $N^{1/2}$. Formally, one may choose $s_N = N^{-d}$, where $0 < d < ½$.[12] We choose $s_N = 0.1$ (i.e., $d = 0.363$). Since Φ is a bounded function, one can exchange expectation and limit in (11) to obtain the smoothed moment condition in (9).

[12] This condition is required for the proof of asymptotic normality.

The estimation below relies on the fact that our instrument vector, $(X_j, \tilde{\sigma}_{js})$ will be conditionally independent of the error terms defined by $\left(1\left[\delta_{js} > Q_\rho(X_j, \sigma_{js})\right] - \rho\right)$ in forming a valid set of moment conditions. Practically, this Smoothed GMM procedure can be sensitive to the initial parameter guess. We use the approach proposed by Hansen and Chernazukov (2001) to obtain starting values.[13] Standard errors are those reported by the GMM estimation procedure in any statistical package.

6. Estimation Results

Our estimation results are reported in two groups, reflecting the two-part estimation procedure described above. Table 3 reports estimates of our first-stage (i.e., maximum likelihood) parameter estimates, describing how preferences for certain components of X_j, σ_{js}, and TC_{ij} vary with angler attributes (i.e., education, employment status, and an indicator for boat ownership). The top rows report values for the set of

[13] Taking the expression for the τ^{th} conditional quantile (τ = 0.5 for median regression), $Q_{\delta|\tau}(\tau) = X'_j\beta_\tau + \alpha_\tau\sigma_j$, the parameters ($\beta_\tau$, α_τ) describe the way in which the τ^{th} percentile of the distribution of $\delta_{j,WD/WE}$ evolves with X_j and σ_j. The usual regression framework describes instead the evolution of the mean of the distribution. Hansen and Chernozhukov propose defining a new dependent variable, $\hat{\delta}_{js}(\alpha_\tau) = \delta_{js} - \alpha_\tau\sigma_{js}$, which clearly depends upon some assumed value for α_τ. Alternative values of $\hat{\delta}_{js}(\alpha_\tau)$ are calculated for all of the possible values that α_τ might take, and each is used as the dependent variable in a separate quantile regession:

$$\hat{\delta}_{js}(\alpha_\tau) = X_j\beta_\tau + \lambda_\tau\tilde{\sigma}_{js}$$

producing a range of estimates [$\beta_\tau(\alpha_\tau)$, $\lambda_\tau(\alpha_\tau)$]. $\tilde{\sigma}_{js}$ is the predicted value of the time-period-specific share of visitors at site j based only on exogenous attributes. As in the discussion in Section 2, it functions here as an instrument.[13] In particular, Hansen and Chernozhukov show that the optimal value of α_τ can be found by exploiting the exclusion restriction implied by the instrument, and calculating $\alpha^*_\tau = \underset{\{\alpha_\tau\}}{ARGMIN}\,[\lambda_\tau(\alpha_\tau)]^2$. Practically, this involves performing a grid-search over the possible values that α_τ might take. The quality of the overall estimates depends upon the precision with which this grid-search is carried out. We therefore use it to find good starting values for our smoothed GMM estimator, but derive our standard errors from the latter. In practice, parameter estimates change very little depending upon which estimator is employed.

weekdays, while the bottom rows report values for weekend visits. Given the flexibility introduced by the second stage of the estimation procedure (in particular, the inclusion of the unobserved attribute ξ_{js}, we do not attempt to estimate all possible first-stage interactions. Particularly important is the interaction between boat ownership and our proxy for variables we might expect to be important to boat owners. As a proxy for these factors, we use *PAVED* (i.e., an indicator of a paved boat launch at the site), which is highly correlated with there being no restrictions on motor trolling and there being multiple launches at the site. The interaction between *PAVED* and boat ownership is positive and significant in both time periods. Sites designated as URBAN are also less attractive to boat owners. The unemployed place a significantly higher value (or, alternatively, a lower cost) on congestion on weekdays, and boat owners face significantly higher travel costs on weekdays. Finally, note that travel cost (measured in miles) enters negatively and is very significant for both (i.e., $\theta_{0,WD}$ = -2.671 and $\theta_{0,WD}$ = -2.415 with t-statistics of -19.23 and -21.24, respectively). We will use the disutility of travel cost to convert changes in utility associated with the elimination of a large site into comparable units in the following section.

Table 4 reports estimates from our second-stage IV median regression decompositions of $\delta_{j,WD}$ and $\delta_{j,WE}$. The most important parameter for our purposes is the utility effect of expected share (i.e., congestion). The effect is negative and significant in both time periods. While the magnitude of this congestion cost is greater for weekend visits, it is estimated more precisely for weekdays.

Other second stage parameter estimates have the expected sign. Being located near a wildlife protection area, WILDLIFE (i.e., a variable that is highly correlated with proximity to a state, county, or national forest) and having restroom facilities, RESTROOM (i.e., a variable that is highly correlated with site amenities such as a fishing pier, a paved parking lot, and a picnic area) both enter positively into utility (particularly for weekend visits). The effect of the site being a river is negative for weekends, while the same is true for small lakes (i.e., under 50 acres in area) on weekdays.

Expected catch variables generally enter significantly into utility but with varying degrees of statistical significance. **(****More here on catch variables, and anything else about other parameter estimates****)**

The Role of "IV" in our IV Quantile Estimation

In order to demonstrate the value of the second-stage quantile regression IV strategy, Table 5 reports estimates from a similar set of second-stage regressions that ignore the endogeneity of σ_{js}. Estimates reflect a baseline preference *for* increased congestion (i.e., the expected direction of bias, but extreme enough to produce an agglomeration effect). This will have important implications for site valuation (see Section 7), but also leads to biases in the marginal values we place on specific site attributes. For example, the marginal utility of WILDLIFE falls dramatically (0.700 to 0.451 for weekend visits), and the effect is similar on the value placed on RESTROOMS (0.588 to 0.178). Similar effects are found for the marginal disutility of RIVER and SMALL LAKE.

The Role of "Quantile" in Our IV Quantile Estimation

Table A2 reports estimates of the second-stage utility parameters for different values of the "patch" described in the previous section under IV quantile regression and two-stage least squares procedures. The purpose of these tables is to highlight the role of using quantile regression to control for the arbitrary choice of ε. We find that parameter estimates associated with various site attributes (including congestion) vary dramatically with ε under 2SLS estimation.

7. Valuing of a Large Site

We now examine the role of congestion costs in valuing a large site. We focus on large sites, because the exercise of removing such a site from the choice set will involve significant re-sorting of anglers among the remaining sites. The welfare effects of that re-sorting need to be accounted for in the value ascribed to the site. Ignoring them has the potential to lead to serious under-measurement of value.

A good candidate for such an exercise is Lake Winnebago. Responsible for 9.3% of all trips in our sample, Lake Winnebago is one of Wisconsin's premier sites for fishing and other water activities. It is by far the largest inland lake with over 135,000 acres of surface area and is known for good walleye and perch fishing.

The procedure for valuing Lake Winnebago proceeds as follows. We begin by determining each angler's expected utility under the status quo in a particular time period (we focus here on weekdays, but the exercise could be easily repeated for weekends). In doing so, we first employ the contraction mapping defined in Section 3 to solve for the equilibrium vector of shares under the status quo ($\sigma^0_{j,WD}$).:

(12) $$\sigma^0_{j,WD} = \int \frac{EXP\{\hat{V}_{i,j,WD}\}}{\sum_{l=1}^{J} EXP\{\hat{V}_{i,l,WD}\}} dF(Z_i)$$

where

(13) $$\hat{V}_{i,j,WD} = X'_j \hat{\beta}_{WD} + \hat{\alpha}_{WD} \sigma^0_{j,WD} + \hat{\xi}_{j,WD} + X_j \hat{\Gamma}_{WD}(Z_i) + \hat{\Phi}_{WD}(Z_i)\sigma_{j,WD} + \hat{\Theta}_{WD}(Z_i) \ln TC_{i,j}$$

and a "hat" over a parameter refers to an estimated value recovered in the previous section. By construction, this replicates the shares of weekday anglers choosing each site observed in the data.[14] Based on these shares, we can calculate each angler's expected utility according to the familiar log-sum rule:

(14) $$EU^0_{i,WD} = \ln\left(\sum_{j=1}^{J} EXP\{\delta^0_{j,WD} + X_j \hat{\Gamma}_{WD}(Z_i) + \hat{\Phi}_{WD}(Z_i)\sigma^0_{j,WD} + \hat{\Theta}_{WD}(Z_i) \ln TC_{i,j}\} \right)$$

where

(15) $$\delta^0_{j,WD} = X'_j \hat{\beta}_{WD} + \hat{\alpha}\sigma^0_{j,WD} + \hat{\xi}_{j,WD}$$

[14] Recall that the vector of $\delta_{j,WD}$'s was calculated (with the contraction mapping algorithm adapted from Berry (1994)) to ensure that the share of anglers choosing each site would exactly equal the actual share.

This welfare measure weights the utility the individual would get from each choice by the probability that he or she chooses it. As such, it ascribes positive value to Lake Winnebago for individuals who we observe choosing other sites; the magnitude of that value, however, will depend upon how close a substitute Lake Winnebago is for the chosen site.

Next, we eliminate the sites associated with Lake Winnebago from the choice set and re-calculate the equilibrium share of trips to each of the remaining sites according to (12) and (13).[15] This yields a new vector of equilibrium shares ($\sigma^1_{j,WD}$) from which we can calculate new values of expected utility ($EU^1_{i,WD}$). Different types of individuals' expected utilities are not directly comparable, so we divide by each individual's marginal disutility of travel cost (evaluated at 20 miles), so as to convert all measures into miles.[16] This yields the following measure of foregone expected utility:

$$(16) \qquad \Delta EU^{1-0}_{i,WD} = \frac{20\,(EU^1_{i,WD} - EU^0_{i,WD})}{\Theta_{WD}(Z_i)}$$

[15] Recall that our data consider sites to be composed of evenly sized grid cells, and that large sites (e.g., Lake Winnebago, Lake Michigan, Green Bay) will contain many of these cells. Taking Lake Winnebago out of the choice set involves eliminating 9 of these cells.

[16] Because of complications introduced by the endogenous sorting process, we measure changes in welfare by considering changes in expected utility under the two scenarios (i.e., with and without Lake Winnebago). In this sort of endogenous sorting model, the use of a more traditional measure of welfare (e.g., an equivalent or compensating variation) is not possible because the "currency" in which that variation is denominated (i.e., miles) changes endogenously with the very change that we seek to value (i.e., as anglers choose new preferred sites in the "No Lake Winnebago" equilibrium). Equivalent and compensating variations in income are traditionally used when prices change in a counterfactual scenario, but an individual's income remains fixed. One can then determine what change in income would have yielded an equivalent change utility to that introduced by a price change (in the case of EV). Here, changing miles in an analogous way would induce another resorting of the population, invalidating the EV exercise.

Welfare falls for every angler, with an average cost of 2.25 miles. **(****How should we best add this up to get a total value? Are there any interesting ways to report values for all the anglers?****)**

In order to demonstrate the role of properly measured congestion effects in valuing a large site like Lake Winnebago, we next perform the same exercise, but using parameter estimates from a model that ignores the role of congestion in utility. Tables 6 and 7 report first- and second-stage parameter estimates, respectively, for such a model (for weekdays only).

(**Some discussion of parameter estimates ignoring congestion****)**

Without any role for congestion costs, there is no need to calculate the new equilibrium distribution of anglers without Lake Winnebago in the choice set – the welfare measure expressed in equations (14) and (15) requires only that we know the attributes of the remaining sites. Using those equations, we calculate a comparable set of foregone expected utilities. In line with our intuition, the costs of eliminating Lake Winnebago from the choice set are smaller in the model that ignores congestion costs. The average cost falls from 2.25 miles with congestion to 1.48 miles without it (a 34% reduction). Considering only those who had originally chosen Lake Winnebago, however, their average costs fall from 9.07 miles with congestion to 9.00 miles when congestion is ignored – these individuals suffer a significant loss of utility under either modeling assumption, but their primary welfare cost comes from having to switch to a new fishing site. More important is what happens to the many anglers who did not originally choose to visit Lake Winnebago. Under our model without congestion, they experience only a small cost in expectation (0.57 miles) from the elimination of Lake

Winnebago (i.e., because each had a small probability of actually visiting the lake). Including a role for congestion, however, their average costs rise to 1.42 miles.[17]

(**Any other interpretation or policy analysis to do here?****)**

8. Conclusions and Caveats

Summary:

Congestion is a potentially important site attribute in many travel cost models of recreation demand.

Measuring congestion costs properly involves a difficult endogeneity problem. While stated preference models offer a potential solution based on answers to hypothetical questions, revealed preference approaches require an instrumental variables solution (unless congestion effects are to be identified on the basis of functional form restrictions alone).

[17] Looking at the three sites that see the largest increases in share following the elimination of Lake Winnebago, anglers at Little Lake Butte Des Morts (increase of 0.37% of total trips) see a rise in expected costs from 5.84 miles to 9.99 miles when congestion costs are considered. Expected costs incurred by anglers at Wolf Lake (increase of 0.26% of total trips) and on the segment of the Fox River located closest to Lake Winnebago (increase of 0.24% of total trips) increase from 2.65 to 3.59 miles and from 7.31 to 9.61 miles, respectively. By contrast, anglers originally choosing to visit Sand Lake, located far away in Oneida county, had an increase in expected costs from 0.01 to 0.28 miles with the incorporation of congestion costs. Sand Lake is predicted to see an increase of 0.01% of the total trips with the elimination of Lake Winnebago.

We find evidence of significant congestion effects. Failing to properly account for their endogeneity leads one to incorrectly recover agglomeration benefits (!), and to mis-measure the value of other site attributes.

Lesson for policy is that we will tend to understate the value of large sites if we ignore the role of congestion costs. We could dramatically understate those costs (even find a benefit?) if we improperly measure the costs of congestion.

Caveats:

Our counterfactual simulations don't allow people to opt out of taking a trip (forces all Lake Winnebago users to move to other lakes).

We don't allow for random parameters in utility.

We don't account for the fact that trips are not independent events, but that the repeated decisions of a single angler are correlated.

References

Anderson, F.J. and N.C. Bonsor (1974). "Allocation, Congestion, and the Valuation of Recreational Resources." *Land Economics*. 50:51-57.

Bayer, P. and C. Timmins (2005a). "Estimating Equilibrium Models of Sorting Across Locations." Mimeo, Yale University Department of Economics.

Bayer, P. and C. Timmins (2005b). "On the Equilibrium Properties of Locational Sorting Models." Forthcoming in Journal of Urban Economics.

Bockstael, N.E., K.E. McConnell, and I.E. Strand (1989). "A Random Utility Model for Sportfishing: Some Preliminary Results for Florida." *Marine Resource Economics*. 6: 245-260.

Berry, S. (1994). "Estimating Discrete Choice Models of Product Differentiation." *RAND Journal of Economics*. 25:242-62.

Berry, S., J. Levinsohn, and A. Pakes (1995). "Automobile Prices in Market Equilibrium." *Econometrica*. 63:841-90.

Boxall, P. and W. Adamowicz (2000). "Incorporating Endogenous Perceptions of Environmental Attributes in RUMs: The Case of Congestion." *IIFET 2000 Proceedings*. 1-7.

Boxall P., K. Rollins, and J. Englin (2003). "Heterogeneous Preferences for Congestion During a Wilderness Experience." *Resource and Energy Economics*. 25:177-95.

Cesario, F. (1980). "Congestion and the Valuation of Recreation Benefits." Land *Economics*. 56(3):329-38.

Chernozhukov, V. and C. Hansen (2001). "An IV Model of Quantile Treatment Effects." MIT Department of Economics Working Paper.

Cichetti, C. and V.K. Smith (1973). "Congestion, Quality Deterioration and Optimal Use: Wilderness Recreation in the Spanish Peaks Primitive Area." *Social Science Research*. 2(1):15-30.

Fisher, A. and J. Krutilla (1972). "Determination of Optimal Capacity of Resource-Based Recreation Facilities." *Natural Resources Journal*. 12:417-444.

Hong, Han and Thomas MaCurdy (1999)

Imbens, G.W. and W.K. Newey (2003). "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity." Unpublished Manuscript.

Jakus, P. and W. Shaw (1997). "Congestion at Recreation Areas: Empirical Evidence on Perceptions, Mitigating Behaviour and Management Preferences." *Journal of Environmental Economics and Management*. 50:389-401.

Koenker, R. and G. Bassett (1978). "Regression Quantiles." *Econometrica.* 46(1):33-50.

Koenker, R. and K. Hallock (2001). "Quantile Regression." *Journal of Economic Perspectives*. 15(4):143-156.

Ma, L. and R. Koenker (2003). "Quantile Regression Methods for Recursive Structural Equation Models." Unpublished Manuscript.

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

Michael and Rieling (1997)

Murdock, J. (2002) "Unobserved Characteristics within a Discrete Choice Model: An Application to Wisconsin Recreational Fishing," Dissertation, Yale University.

Schuhmann, P.W. and K.A. Schwabe (2004). "An Analysis of Congestion Measures and Heterogeneous Angler Preferences in a Random Utility Model of Recreational Fishing." *Environmental and Resource Economics*. 27:429-50.

Table 1
Data Summary

Type	Variable	Description	mean	s.d
Angler Characteristics for the 512 anglers reporting single day fishing trips:				
	HS GRAD / DROPOUT	dummy = 1 if angler completed high school at most	0.52	-
	UNEMPLOYED	dummy = 1 if angler not employed full or part time	0.15	-
	BOAT OWNER	dummy = 1 if angler in a household that owns a boat	0.51	-
Trip Characteristics for 3,581 reported single day fishing trips:				
	WE	dummy = 1 if trip on the weekend (Sat. or Sun.)	0.49	-
	WD	dummy = 1 if trip on a weekday (Mon. – Fri.)	0.51	-
	TRAVEL COST	One-way travel distance in miles	19.9	20.8
Site Characteristics for the 569 unique fishing sites:				
	PAVED	dummy = 1 if offers at least one paved boat launch	0.73	-
	URBAN	dummy = 1 if urban area on shoreline	0.18	-
	WILDLIFE	dummy = 1 if site inside a Wildlife Area or Refuge	0.06	-
	FOREST	dummy = 1 if site inside a State or National Forest or a Wildlife Area or Refuge	0.18	-
	RESTROOM	dummy = 1 if restroom available	0.58	-
	RIVER	dummy = 1 if a river fishing location	0.31	-
	SMALL LAKE	dummy = 1 if inland lake surface area < 50 acres	0.17	-
	TROUT	catch rate brook, brown, and rainbow trout	0.09	0.17
	SMALLMOUTH	catch rate smallmouth bass	0.20	0.20
	WALLEYE	catch rate walleye	0.13	0.15
	NORTHERN	catch rate northern pike	0.08	0.06
	MUSKY	catch rate muskellunge	0.01	0.31
	SALMON	catch rate coho and chinook salmon	0.01	0.05
	PANFISH	catch rate yellow perch, bluegill, crappie, sunfish	1.58	0.89
	LARGEMOUTH	catch rate largemouth bass	0.19	0.14

Table 2
Percentage Share of Trips to Each Fishing Site

Share of Trips (%)	obs.	mean	median	s.d.	min.	10^{th} Percentile	90^{th} Percentile	max.
SHARE*100: Weekdays only	569	0.18	0.06	0.38	0	0	0.44	3.50
SHARE*100: Weekends only	569	0.18	0.06	0.32	0	0	0.46	2.69
SHARE*100: All Days	569	0.18	0.06	0.33	0.03	0.03	0.39	3.02

Table 2a
Weighted Percentage Share of Trips to Each Fishing Site

Share of Trips (%)	obs.	mean	median	s.d.	min.	10th Percentile	90th Percentile	max.
SHARE*100: Weekdays only	569	0.18	0.05	0.48	0	0	0.41	7.80
SHARE*100: Weekends only	569	0.18	0.07	0.32	0	0	0.42	2.82
SHARE*100: All Days	569	0.18	0.07	0.37	0.004	0.01	0.44	5.35

Table 3
First-Stage Parameter Estimates and t-statistics
Weekdays: TRAVEL COST (baseline) = -2.671 (-19.23)
Weekends: TRAVEL COST (baseline) = -2.416 (-20.96)

	HS GRAD/ DROPOUT		UNEMPLOYED		BOAT OWNER	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Weekdays:						
PAVED	-0.035	-0.11	1.033	2.04	0.826	2.13
URBAN	-0.170	-0.54	-0.542	1.49	-1.070	-3.47
WILDLIFE	-0.565	-1.08	-0.043	-0.11	-0.754	-1.52
RESTROOM	-0.324	-0.97	0.546	1.23	0.322	0.95
SHARE*100	-0.085	-0.63	0.279	1.92	0.115	0.80
TRAVEL COST	-0.148	-1.14	-0.046	-0.32	0.368	2.72
Weekends:						
PAVED	0.121	0.32	0.586	0.95	0.932	2.51
URBAN	-0.184	-0.63	-0.050	-0.11	-0.807	-2.66
WILDLIFE	-0.124	-0.24	0.723	0.91	-0.439	-0.79
RESTROOM	-0.202	-0.68	-0.214	-0.48	-0.029	-0.10
SHARE*100	0.197	1.09	0.133	0.52	-0.054	-0.32
TRAVEL COST	0.007	0.08	-0.126	-0.68	0.191	1.54

Table 4
Second-Stage Parameter Estimates
IV Median Regression

	Weekdays		Weekends	
	Estimate	t-stat	Estimate	t-stat
CONSTANT	-7.045	-4.84	-3.25	-2.98
PAVED	0.408	1.00	0.541	1.97
URBAN	0.788	1.96	0.304	0.93
WILDLIFE	1.308	1.83	0.887	2.33
RESTROOM	1.405	3.87	0.540	2.69
RIVER	-0.576	-0.47	-1.623	-2.23

SMALL LAKE	-1.889	-3.91	-0.002	-0.01
TROUT	1.611	2.23	-0.139	-0.10
SMALLMOUTH	0.928	1.48	0.148	0.30
WALLEYE	3.509	4.25	2.646	1.85
NORTHERN	3.238	1.27	0.376	0.27
MUSKY	-4.826	-1.90	11.692	1.29
SALMON	7.262	1.80	2.771	0.87
PANFISH	0.845	1.23	-0.234	-0.79
LARGEMOUTH	-0.979	-0.50	-1.651	-1.15
SHARE*100	-2.934	-3.47	-5.445	-2.16

Table 5
Second-Stage Parameter Estimates
Median Regression (No Instruments for Share)

	Weekdays		Weekends	
	Estimate	t-stat	Estimate	t-stat
CONSTANT	-8.522	-4.84	-5.193	-4.40
PAVED	-0.282	-0.50	0.288	0.76
URBAN	0.358	0.61	-0.306	-0.77
WILDLIFE	1.265	1.34	0.431	0.68
RESTROOM	1.189	2.52	0.182	0.58
RIVER	1.030	0.79	-0.092	-0.10
SMALL LAKE	-2.059	-3.13	-0.062	-0.14
TROUT	1.847	1.10	0.175	0.16
SMALLMOUTH	1.438	1.13	0.707	0.83
WALLEYE	2.170	1.20	0.653	0.54
NORTHERN	2.431	0.57	-0.212	-0.07
MUSKY	2.514	0.23	7.809	1.07
SALMON	8.327	1.18	-1.231	-0.26
PANFISH	1.393	2.09	0.237	0.53
LARGEMOUTH	1.181	0.47	0.536	0.32
SHARE*100	1.130	1.56	2.302	4.72

Table 6
Weekdays First-Stage Parameter Estimates
No Congestion Effects
TRAVEL COST (baseline) = -2.675 (-19.39)

	HS GRAD/ DROPOUT		UNEMPLOYED		BOAT OWNER	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
PAVED	-0.048	-0.17	1.107	2.23	0.853	2.38
URBAN	-0.163	-0.53	-0.543	-1.53	-1.071	-3.49
WILDLIFE	-0.525	-1.02	-0.088	-0.18	-0.769	-1.55
RESTROOM	-0.351	-1.08	0.670	1.54	0.372	1.11
TRAVEL COST	-0.146	-1.14	-0.039	-0.27	0.360	2.69

Table 7
Weekdays Second-Stage Parameter Estimates
No Congestion Effects
Median Quantile Estimation

	Estimate	t-stat		Estimate	t-stat
CONSTANT	-8.203	-4.63	SMALLMOUTH	1.435	1.12
PAVED	-0.614	-0.11	WALLEYE	2.893	1.62
URBAN	0.432	0.73	NORTHERN	3.011	0.70
WILDLIFE	1.533	1.61	MUSKY	2.084	0.19
RESTROOM	1.177	2.48	SALMON	9.236	1.34
RIVER	0.334	0.25	PANFISH	1.250	1.86
SMALL LAKE	-2.157	-3.27	LARGEMOUTH	0.592	0.23
TROUT	3.225	1.91			

Figure 1
Map of Wisconsin Showing Quadrangles Used in Defining Fishing Sites

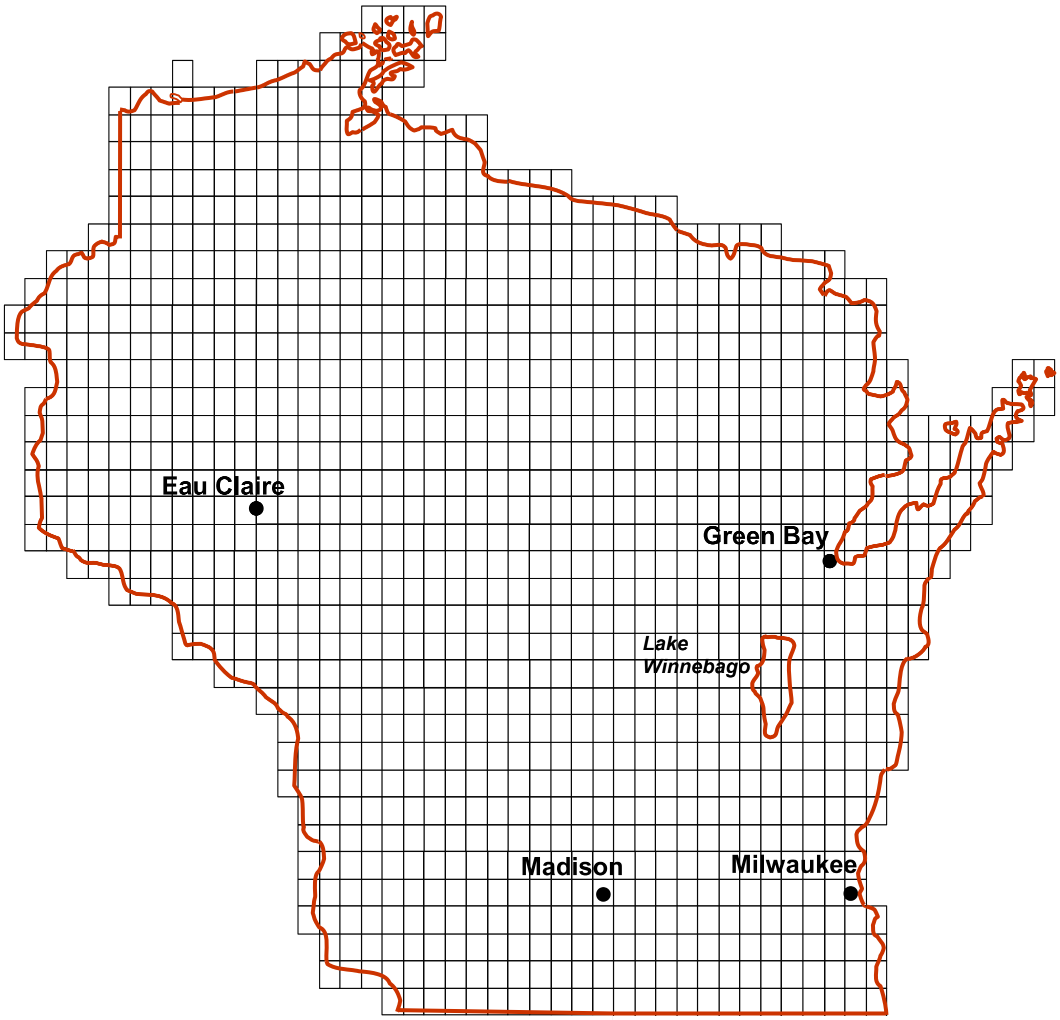
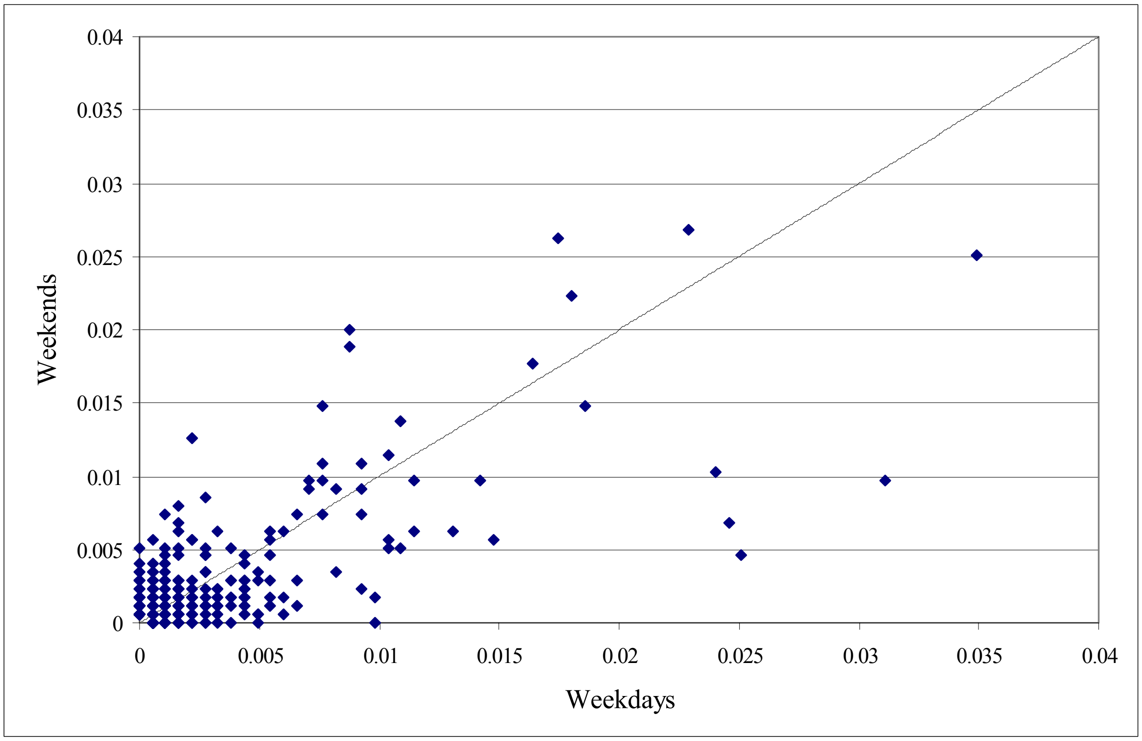


Figure 2
Expected Shares for Weekdays versus Weekends



Appendix: Estimates Under Alternative Assumptions About Zero Shares

Table A1
Weekday First Stage Parameter Estimates
Under Alternative Values of $\varepsilon = (10^{-9}, 10^{-12})$
TRAVEL COST (baseline) = -2.671 (-19.23) in both cases

	HS GRAD/ DROPOUT		UNEMPLOYED		BOAT OWNER	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
$\varepsilon = 10^{-9}$:						
PAVED	-0.035	-0.11	1.033	2.04	0.826	2.13
URBAN	-0.170	-0.54	-0.542	1.49	-1.070	-3.47
WILDLIFE	-0.565	-1.08	-0.043	-0.11	-0.754	-1.52
RESTROOM	-0.324	-0.97	0.546	1.23	0.322	0.95
SHARE*100	-0.085	-0.63	0.279	1.92	0.115	0.80
TRAVEL COST	-0.148	-1.14	-0.046	-0.32	0.368	2.72
$\varepsilon = 10^{-12}$:						
PAVED	-0.035	-0.11	1.033	2.04	0.826	2.14
URBAN	-0.170	-0.54	-0.542	-1.49	-1.069	-3.47
WILDLIFE	-0.565	-1.08	-0.043	-0.11	-0.754	-1.52
RESTROOM	-0.324	-0.97	0.546	1.23	0.322	0.95
SHARE*100	-0.085	-0.63	0.279	1.92	0.115	0.80
TRAVEL COST	-0.148	-1.14	-0.046	-0.302	0.368	2.72

Table A2
Weekday Second Stage Parameter Estimates
Under Alternative Values of $\varepsilon = (10^{-9}, 10^{-12})$
IV Median Estimation and Two-Stage Least Squares

	IV Median Estimation				Two-Stage Least Squares			
	10^{-9}		10^{-12}		10^{-9}		10^{-12}	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
CONSTANT	-7.045	-4.84	-7.045	-4.84	-11.665	-5.16	-15.593	-4.60
PAVED	0.408	1.00	0.408	1.00	-0.053	-0.07	-0.066	-0.06
URBAN	0.788	1.96	0.788	1.96	0.861	1.11	1.150	0.99
WILDLIFE	1.308	1.83	1.308	1.83	1.087	0.90	1.315	0.72
RESTROOM	1.405	3.87	1.405	3.87	1.214	1.98	1.781	1.94
RIVER	-0.576	-0.47	-0.576	-0.47	0.078	0.05	0.512	0.20
SMALL LAKE	-1.889	-3.91	-1.889	-3.91	-1.851	-2.18	-2.561	-2.01
TROUT	1.611	2.23	1.611	2.23	2.426	1.13	2.820	0.88
SMALLMOUTH	0.928	1.48	0.928	1.48	1.587	0.97	2.059	0.84
WALLEYE	3.509	4.25	3.509	4.25	1.735	0.71	1.094	0.30
NORTHERN	3.238	1.27	3.238	1.27	7.133	1.30	10.358	1.26
MUSKY	-4.826	-1.90	-4.826	-1.90	4.276	0.30	3.484	0.17
SALMON	7.262	1.80	7.262	1.80	5.140	0.53	4.469	0.31
PANFISH	0.845	1.23	0.845	1.23	0.695	0.81	0.817	0.64
LARGEMOUTH	-0.979	-0.50	-0.979	-0.50	1.009	0.31	1.881	0.39
SHARE*100	-2.934	-3.47	-2.934	-3.47	4.300	2.34	7.575	2.75

Figure A1
Weekday Distribution of Weekday Fixed Effects ($\delta_{j,WD}$)
Under Alternative Values of $\varepsilon = (10^{-9}, 10^{-12})$