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

by

Christopher Timmins

and

Jennifer Murdock

PRELIMINARY AND INCOMPLETE

June 30, 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 often excluded from such analyses. One of several reasons is that congestion (unlike many other site attributes) is determined in equilibrium by the process of individuals sorting across sites, and thus presents significant endogeneity problems. This paper illustrates this 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 it in a revealed preference context. We demonstrate that failing to address the endogeneity of congestion will likely lead to the understatement of its costs, and possibly to the mistaken recovery of agglomeration benefits. 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. In our application, ignoring congestion leads to an understatement of the lake's value by nearly ½.

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 demand exploit the information in the trade-offs individuals make between travel time and site attributes in order to value the latter. The same models can be used to value bundles of attributes (i.e., entire sites). Consider the case of recreational fishing. Applications typically include data on site 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, or for determining the cost to anglers if a fishing site were to be lost to pollution.

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 to be 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

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

when those sites are large and their closure induces significant resorting over remaining sites.

This paper addresses congestion empirically using revealed preference techniques without basing identification on functional form assumptions, as has been the case in other revealed preference work. It does so by relying on a previously unexploited source of variation in the data – the isolation of alternative sites in exogenous attribute space. Without exploiting this source of variation, controlling for congestion is a difficult task. 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 cost of congestion and to even mistakenly recover estimates of benefits from larger crowds (i.e., agglomeration effects). In this paper, we describe the source of this endogeneity, cast it as a simple instrumental variables problem in a familiar 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. In Section 4, we describe the data set we use in an application of our technique. In Section 5 we discuss an econometric complication that arises when we model different congestion effects depending upon whether they occur on a weekday versus a weekend. Section 6 reports model estimates, and Section 7 illustrates the role of congestion in 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) is one of the first to discuss the implications of congestion for measuring willingness to pay, while Fisher and Krutilla (1972) notes 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 equilibrium resorting. The removal of a recreational site adversely affects the welfare of users of other sites as displaced recreators re-sort across the remaining sites. Conversely, there is a tendency to understate the value of new site construction if congestion costs are ignored. In a more recent paper, Jakus and Shaw (1997) discuss ways to measure congestion, emphasizing 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 highlight the timing of congestion costs. This could entail, for example, differentiating between the expected number of other recreators on a weekday versus a weekend visit. Michael and Rieling (1997) discuss 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" (i.e., congestion in the hiking context) on stated willingness to pay with an application to the Spanish Peaks Primitive Area in Montana. McConnell (1977) employs stated preference techniques to estimate the role of congestion in the demand for beach recreation and uses the results to characterize 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. Consider briefly, however, how stated preference data solve the endogeneity problems associated with congestion. Congestion is determined by the optimizing decisions of recreators; measuring it 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 in question, making it an endogenous attribute. Stated preference models avoid this problem by hypothetically varying congestion while holding constant the unobservables that drive sorting behavior. The downsides of this solution are (i) that stated preference models value hypothetical changes about which respondents may not reveal their true preferences, and (ii) respondents may not actually be able to "hold all else constant" when hypothetically varying the congestion variable – i.e., stated distaste for congestion may reflect preferences for or against unobserved attributes typically associated with congestion.

There have been few papers that have addressed the problem of valuing congestion with revealed preference data. 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 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.

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 5, 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 application.

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 decisions that will be made by other individuals. In equilibrium, those expectations are confirmed by other individuals' actual behavior. 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 (i.e., the choice of whether or not to take a trip) is not modeled. The utility obtained from choosing site j on occasion t in period s = {weekday, weekend} is given by:

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

where

(2) $$\delta_{js} = X_j'\beta_s + \alpha_s\sigma_{js} + \xi_{js}$$

(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[2]
ξ_{js}	=	unobservable attribute of site j in time period s (common to all anglers)[3]
ε_{ijts}	=	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 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, $\theta_0 \ln TC_{ij}$.

Individuals are ascribed rational expectations about the behavior of their fellow anglers. This means that the vector of expected shares will be constant across individuals

[2] We measure travel cost by the angler's imputed opportunity cost of time multiplied by the roundtrip travel time, plus 15¢ per mile. Murdock (2002) describes this imputation in more detail.

[3] 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

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 achieved a Nash equilibrium.

Simplifying Assumptions

In our application, we distinguish between congestion on weekdays (WD) and weekends (WE). Within each of these periods, we treat each site selection choice made by an angler as an independent event. We therefore calculate two sets of expected shares ($\sigma_{j,WD}$, $\sigma_{j,WE}$), and we estimate a separate set of preference parameters for each of these periods. 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 ignore the fact 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). We could, for example, also model the decision about which day of the week to go fishing, or whether to take a fishing trip at all. While these complications could be 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 several 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 also be incorporated into our modeling framework.

different times such as on a weekday versus weekend. We therefore allow for this possibility in the way we define our unobservables.

Equilibrium

Each angler maximizes his or her utility given expectations about the behavior of other anglers. In equilibrium, those expectations are 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* in *s* as:

$$(4) \qquad P(U_{ijts} \geq U_{ilts} \; \forall \, l \neq j) \;=\; \frac{EXP\{\delta_{js} + X_j \Gamma_s(Z_i) + \Phi_s(Z_i)\sigma_{js} + \Theta_s(Z_i)\ln TC_{ij}\}}{\sum_{l=1}^{J} EXP\{\delta_{ls} + X_l \Gamma_s(Z_i) + \Phi_s(Z_i)\sigma_{ls} + \Theta_s(Z_i)\ln TC_{il}\}}$$

Integrating over the distribution of angler attributes, $F(Z_i)$, we can predict the share of anglers who will end up choosing each site in each period:

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

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.[4] [Bayer and Timmins (2005b)] Proving uniqueness in the case of agglomeration effects is difficult,

[4] "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.

and depends upon the idiosyncratic features of the data. In the case of congestion effects, however, one can show that the equilibrium is generically unique.

Estimation

While important for counterfactual simulations, uniqueness is not necessary to estimate the parameters of equation (1) by maximum likelihood.[5] [Bayer and Timmins (2005a)] In particular, we can write the period-specific likelihood of observing a vector of site choices:

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

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 $(\overline{\delta}_s, \overline{\gamma}_s, \overline{\phi}_s, \overline{\theta}_s)$ 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.[6]

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

---

5 This is important, because we do not know *a priori* whether preferences exhibit congestion or agglomeration effects, and we require an estimation technique that is valid under both.

6 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 $(\overline{\gamma}_s, \overline{\phi}_s, \overline{\theta}_s)$, 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

anglers choosing a particular site in a particular time period. As will be shown below, 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 $\delta_{j,WD}$ 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:

$$(7) \qquad \delta_{js} = X_j'\beta_s + \alpha_s\sigma_{js} + \xi_{js}$$

for s = WD, WE. This is simply a linear estimation problem with ξ_{js} serving as the regression error. Equilibrium sorting, however, implies a mechanical correlation between σ_{js} and ξ_{js}, COV[σ_{js}, ξ_{js}] > 0. Locations with desirable unobservable attributes will 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_s, and to therefore either understate the costs of congestion or even find 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. Consider how the previous

---

likelihood maximization procedure searches over alternative values of $(\bar{\gamma}_s, \bar{\phi}_s, \bar{\theta}_s)$, the contraction mapping 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?" – i.e., assuming COV[σ_{js}, ξ_{js}] = 0 within the confines of the stated preference experiment.

The one paper we cite that uses revealed preference data instead solves the problem by employing fitted values of σ_{js} based on predictions from an ordered logit model. 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 predicting values of congestion. Rather, predicted congestion varies across sites only with the other observed attributes X_j. The α_s parameter is therefore identified only from the non-linearity inherent in the ordered logit.

An Instrumental Variables Approach

In order to solve this problem, we propose an instrumental variables estimator for equation (7). A valid instrument in this case would be some variable that is correlated with σ_{js}, uncorrelated with ξ_{js}, and that 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*

procedure repeatedly updates the corresponding vector of δ_{js}'s.

can provide valid instruments for the share of anglers choosing site *j*. Intuitively, this is because anglers look across available 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) provides 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, it argues for using the predicted share of anglers choosing site *j* based only on exogenous attributes of all possible choices:[7]

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

[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 then 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 instruments' power.

If exogenous attributes are important determinants of site choice (relative to endogenously determined congestion effects), this instrument will have good power. As sites become 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. Bayer and Timmins (2005a) 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}. Bayer and Timmins (2005a) also provides Monte Carlo evidence on the performance of this instrumental variables strategy in a variety of empirical contexts.

4. Data

This section describes the data on angler characteristics, travel cost, and fishing site characteristics that we use in our application. Murdock (2002) provides additional details about the data and data collection process.

The 1998 Wisconsin Fishing and Outdoor Recreation (WFOR) survey is our 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 during the telephone survey and the final sample reveals that they are very similar. These anglers report 3581 single day fishing trips (1750 weekend and 1831 weekday) that are used for estimation.

The WFOR survey provides sampling weights that describe the number of anlgers in the general population represented by each of the respondents. These weights are used in the following estimations and counterfactual simulations.

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 were collected in the 1950s and 1960s, making them dated, and they exclude Lake Michigan, Green Bay, streams, and rivers. The WFOR fish catch data are detailed and

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

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 in order to 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 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 summarizes expected fish catch along with other site characteristics. 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, state, or county forest (or park, or within a wildlife area.

Our data also describe a variety of site amenities, including access to boat launches (both paved and unpaved), parking lots, picnic areas, docks, fishing piers, camp sites, and restrooms. Many of these attributes are highly correlated with one another in the sample, making it impossible to include all of them in our estimation. Table 2 describes a number of the most important correlations. While there are relatively high correlations between many site amenities, correlations are low between the expected catches of many fish species. This will prove important, particularly in explaining the choice behavior of weekday anglers.

5. Practical Issues in Estimation

The estimation procedure, as described in Section 3, uses the non-zero share of anglers choosing to visit each site in each period in the recovery of the vector of 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 period-specific unobservable attributes, $\xi_{j,WD}$ and $\xi_{j,WE}$. Given the limited number of site attributes described in even the best data sets, including such unobservables is critically important.[10] 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 periods. This is not the case, however, when we consider either period by itself.

Table 3 shows how the share of trips is spread over the 569 sites when considering only weekday or weekend trips. 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 period-specific baseline utilities. In particular, the data tell us only that these are unattractive choices (i.e., so unattractive as to not induce a single visitor in the sample). The data give no indication, however, of exactly how unattractive these sites are.

---

[9] Motor trolling involves trailing a lure or bait from a moving vessel (motor boat or sail boat).
[10] See Murdock (2004) for evidence on the biases introduced by ignoring unobserved site attributes.

We address this problem by first introducing a numerical "patch" that allows the contraction mapping described in Section 3 to function properly. This simply amounts to adding a small increment (e.g., $\varepsilon = 10^{-6}$) to the total number of visits to each site in each 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 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.[11] The Appendix reports parameter estimates under four alternative assumptions about ε, and Figure A1 makes the point about $\delta_{j,WD}$. In particular, it shows the estimated distribution of $\delta_{j,WD}$ under the assumption that $\varepsilon = (10^{-3}, 10^{-6}, 10^{-9}, 10^{-12})$; results are similar for weekend visits. 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. Quantile estimation is flexible in that it does not depend upon the specific values in the lower tail of the δ_{js}

[11] 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, $\to -\infty$.

distribution. As long as a majority of sites have positive numbers of visitors, the median regression is well-suited to this purpose.[12]

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 context because of the presence of the endogenous regressors $\sigma_{j,WD}$ and $\sigma_{j,WE}$ 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

[12] Koenker and Bassett (1978) provides the original theory for quantile regression techniques. Koenker and

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:

$$\text{(9)} \qquad P\big(\delta_{js} < Q_\rho(X_j, \sigma_{js}) \mid X_j, \sigma_{js}\big) = \rho$$

This relation implies the condition:

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

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:

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

where N represents the sample size (569) and Φ is a continuously differentiable distribution function with bounded symmetric density function φ. The following analysis

---

Hallock (2001) provides a convenient summary.

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 $0 < d < ½$.[13] We choose $s_N = 0.23$, which is implied by $d = 0.23$ as well. Since Φ is a bounded function, one can exchange expectation and limit in (11) to obtain the smoothed moment condition in (9).

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 $\left(1\lfloor\delta_{js} > Q_\rho(X_j, \sigma_{js})\rfloor - \rho\right)$ 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.[14] Standard errors are those reported by the GMM estimation procedure in any statistical package.

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

[14] 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 $\delta_{j,WD}$ or $\delta_{j,WE}$ 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 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. 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

6. Estimation Results

Our estimation results are reported in two groups, reflecting the two-part estimation procedure described above. Table 4 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., young children in household, unemployment status, and boat ownership).[15] 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 an indicator for a paved boat launch at the site, which is highly correlated with there being no restrictions on motor trolling and there being multiple launches and a parking lot. The interaction between this indicator and boat ownership is positive and significant for both weekday and weekend visits. Sites designated as urban, wildlife areas, and managed forests are also less attractive to boat owners, as are small lakes and rivers. Anglers with children in the household under the age of 14 derive more utility from site amenities (proxied for by the presence of restrooms) and from higher rates of panfish catch.[16] Boat owners place a significantly higher value (or, alternatively, a lower cost) on congestion. Finally, note that the natural log of travel cost (measured by the imputed opportunity cost of time x travel time plus 15¢ per mile for the round trip) enters

---

use it to find good starting values for our smoothed GMM estimator, but derive our standard errors from the latter.

[15] Note that the standard errors in our first-stage regression are biased downward by our failure to account for the correlation across trips taken by a particular angler. The estimation algorithm currently treats these as independent events.

[16] Average panfish catch rates are higher than for any other species, and catching panfish requires less expertise and elaborate tackle. This makes them ideal for fishing with children.

negatively and is very significant for both weekdays ($\theta_{0,WD}$ = -2.729) and weekends ($\theta_{0,WE}$ = -2.424). 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 5 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 periods.[17] Other second stage parameter estimates generally have the expected sign. For weekday trips, expected catch variables play the dominant role in determining the utility derived from a site. Of the non-catch attributes, only the presence of restrooms (+) and the site being a small lake (-) are significant. Of the catch rates, bass (both large and smallmouth), walleye, and northern are the most significant determinants of behavior. Note that musky catch enters negatively, suggesting that it may be correlated with some undesirable unobserved attribute. For weekend visits, preferences appear to be quite different. The only significant expected catch rate is that for musky (+), while the presence of a paved boat launch and restrooms and the site being designated wildlife protection area all enter positively. Conversely, the site being on a river or a small lake both enter utility negatively.

The Role of "IV" in our IV Quantile Estimation

In order to demonstrate the value of the IV strategy, Table 6 reports estimates from a similar set of second-stage regressions that ignore the endogeneity of σ_{js}. Estimates reflect a significant baseline preference *for* increased congestion (i.e., the expected direction of bias, and extreme enough to produce an agglomeration effect) in both weekday and weekend trips. This has important implications for site valuation,[18] but also leads to biases in the marginal values we place on specific site attributes. For example, the marginal utility of restrooms falls dramatically (0.44 to 0.09 for weekend visits), while the value of small lakes and rivers becomes positive (although insignificant).

The Role of "Quantile" in Our IV Quantile Estimation

Table A1 reports estimates of the second-stage utility parameters for different values of the "patch" described in the previous section under a two-stage least squares estimation procedure. While the results are identical (and, hence, not reported) under the IV quantile approach for each value of ε, we find that parameter estimates associated with various site attributes (including congestion) vary dramatically with ε under 2SLS estimation. Importantly, congestion enters with a positive sign, even after instrumenting. This is a result of two features of the model: (i) unvisited sites offer very low expected congestion, and (ii) their baseline utility becomes increasingly negative with smaller and smaller values of ε. Treating these artificially low values of δ_j as "real" data in the 2SLS procedure makes it seem that congestion is desirable, even when instruments are employed.

[17] Note that we cannot compare parameter estimate magnitudes directly across weekday and weekend regressions as an arbitrary normalization of utility (in particular, $\delta_{I,WD} = \delta_{I,WE} = 0$) underlies each. This will not, however, limit our ability to calculate the welfare effects of eliminating a site.

[18] In the extreme, the elimination of a popular site could possibly be deemed welfare-improving.

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 – one of Wisconsin's premier sites for fishing and other water activities. Next to Lake Michigan, it is Wisconsin's 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 each period. 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 (σ_{js}^0):

$$\sigma_{js}^0 = \int \frac{EXP\{\hat{V}_{ijs}\}}{\sum_{l=1}^{J} EXP\{\hat{V}_{ils}\}} dF(Z_i) \tag{12}$$

where

$$\begin{aligned} \hat{V}_{ijs} &= X_j'\hat{\beta}_s + \hat{\alpha}_s\sigma_{js}^0 + \hat{\xi}_{js} + \\ &\qquad X_j\hat{\Gamma}_s(Z_i) + \hat{\Phi}_s(Z_i)\sigma_{js} + \hat{\Theta}_s(Z_i)\ln TC_{ij} \end{aligned} \tag{13}$$

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.[19] Based on these shares, we can calculate each angler's expected utility according to the familiar log-sum rule:

$$(14) \qquad EU_{is}^{0} = \ln\left(\sum_{j=1}^{J} EXP\{\delta_{js}^{0} + X_j\hat{\Gamma}_s(Z_i) + \hat{\Phi}_s(Z_i)\sigma_{js}^{0} + \hat{\Theta}_s(Z_i)\ln TC_{ij}\}\right)$$

where

$$(15) \qquad \delta_{js}^{0} = X'_j\hat{\beta}_s + \hat{\alpha}\sigma_{js}^{0} + \hat{\xi}_{js}$$

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).[20] This yields a new vector of equilibrium shares (σ_{js}^{1}) from which we can

[19] Recall that the vector of δ_{js}'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.
[20] 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 eliminates 8 of these cells.

calculate new values of expected utility (EU_{is}^{1}).[21] Different types of individuals' expected utilities are not directly comparable, so we divide by each individual's marginal disutility of travel cost (evaluated at the average values of $17.80 for weekdays and $20.97 for weekends), so as to convert all measures into dollars.[22] This yields the following measure of foregone expected utility:

$$(16) \qquad \Delta EU_{i,s}^{1-0} = \frac{\overline{TC}\,(EU_{i,s}^{1} - EU_{i,s}^{0})}{\Theta_s(Z_i)}$$

Welfare falls for every angler, by an average of 92¢ per weekday trip and $1.07 per weekend trip. Aggregating across all trips and sample weights, this translates into total welfare losses of $1,927,333 for weekday visitors and 2,175,683 for weekend visitors, or a total seasonal cost of $4,103,016.

In order to demonstrate the role of congestion effects in valuing a large site like Lake Winnebago, we next perform the same exercise but use parameter estimates derived from a model that ignores the role of congestion in utility. Tables 7 and 8 report first- and second-stage parameter estimates, respectively, for such a model. Without explicitly accounting for the disutility of congestion, we see that the model recovers smaller utilities for amenities associated with more crowded sites. The marginal utility of restrooms, for

[21] Note that, because we do not model the participation decision, we do not allow anglers to opt out of taking a fishing trip at this stage. This will have the effect of biasing upward our estimate of the total cost of eliminating Lake Winnebago.

[22] Because of complications introduced by the endogenous sorting process, we measure welfare effects by considering monetized changes in expected utility under the two scenarios (i.e., with and without Lake Winnebago). In endogenous sorting models of this sort, more traditional equivalent or compensating variation measures of welfare cannot be used, as they assume that income remains constant in the course of the change that they seek to value. Because anglers select new fishing sites with the elimination of Lake Winnebago (and incur new travel costs), this is not the case in our application.

example, falls from 1.432 to 1.301 (weekdays) and 0.440 to 0.395 (weekends). For weekend visitors, the value of paved boat ramps similarly falls from 0.454 to 0.386 while the value of an urban site falls from 0.370 to 0.269.

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 monetized 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 welfare loss per weekday trip falls 38% from 92¢ with congestion to 57¢ without it, whle the average welfare loss per weekend trip falls 52% from $1.07 to 51¢. The total seasonal costs or eliminating Lake Winnebago fall from $4,103,016 to $2,238,003. Ignoring the role of congestion costs yields an estimate of the value of Lake Winnebago that is only 55% of its value when congestion costs are included.

We conclude by examining how welfare costs, both with and without congestion, are distributed across anglers depending upon their initial site choice. For weekday anglers originally choosing Lake Winnebago, welfare loss per trip from eliminating Lake Winnebago rises from $5.87 to $7.28 (24%) when congestion costs are included. Most of this loss results from these anglers having to accept their second-best site choice. Ignoring congestion costs for these anglers alone would not have a large effect on the total value ascribed to the lake. For anglers at the sites that receive most of the additional traffic because of resorting, however, the percentage of the loss attributable to congestion rises. Little Lake Butte des Morts experiences an increase in its share of weekday anglers from 0.0044 to 0.0080, while Wolf Lake sees a similar increase from 0.0048 to 0.0079. Both of these sites are located near to Lake Winnebago, so a significant fraction of the welfare loss ascribed to those originally choosing them should be a direct result of Lake Winnebago no longer being available. Including congestion costs has the effect of raising the welfare loss per trip from $1.90 to $2.96 (56%) for those originally choosing Little Lake Butte des Morts, while it raises it from $1.34 to $2.04 (52%) for those originally choosing Wolf Lake – the role of congestion costs as a determinant of welfare loss increases. Finally, considering anglers who had originally chosen one of the remaining 559 fishing sites in our sample, we find that inclusion of congestion costs increases their welfare loss from 13¢ to 39¢ (200%). Certainly, these anglers suffer less from the elimination of Lake Winnebago, but the majority of their welfare loss is attributable to congestion effects. When we consider that 99% of all weekday fishing trips fall into this final category, the significance of congestion effects for the proper recovery of site value becomes apparent.

8. Conclusions and Caveats

Congestion is an important site attribute in many travel cost models of recreation demand, but it is typically ignored, particularly in the revealed preference context. This is because properly controlling for congestion costs requires the solution of 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 identification is to be achieved on the basis of functional form restrictions alone).

Implementing such an instrumental variables approach, 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. The practical lesson for policy-makers is that we will tend to understate the value of large sites (by almost ½ in the case of Lake Winnebago) if we ignore the role of congestion costs.

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 (a)
Data Summary – Site Attributes

Variable	Description	Mean	S.D.
URBAN	Dummy = 1 if urban area on shoreline	0.18	0.38
WILDLIFE	Dummy = 1 if site inside a wildlife area or refuge	0.06	0.23
FOREST	Dummy = 1 if site inside a county, state or national forest	0.18	0.38
LAUNCH	Dummy = 1 if site has a boat launch		
NLAUNCH	Number of boat launches available at site	1.58	2.26
PAVED	Dummy = 1 if offers at least one paved boat launch	0.73	0.45
PARKING	Dummy = 1 if parking lot is available	0.79	0.45
PICNIC	Dummy = 1 if picnic area is available	0.52	0.50
DOCK	Dummy = 1 if boating dock is available	0.49	0.50
PIER	Dummy = 1 if fishing pier is available	0.36	0.48
RESTROOM	Dummy = 1 if restroom available	0.58	0.49
RIVER	Dummy = 1 if a river fishing location	0.31	0.46
SMALL LAKE	Dummy = 1 if inland lake surface area < 50 acres	0.17	0.38
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	0.02
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 1 (b)
Data Summary – Angler Attributes

Variable	Description	Weekdays (n=1831)		Weekends (n=1750)	
		Mean	S.D.	Mean	S.D.
KIDS	Dummy = 1 if children under age 14 in household	0.31	0.46	0.39	0.49
UNEMPLOYED	Dummy = 1 if angler not employed full or part time	0.24	0.43	0.12	0.33
BOAT OWNER	Dummy = 1 if angler in a household that owns a boat	0.59	0.49	0.59	0.49
TRAVEL COST	Round-trip travel time x opportunity cost of time + 15¢ per mile	17.55	18.99	20.19	20.63

Table 2 – Correlations Between Site Attributes

Table 2 (a): Urbanization

	FOREST	WILDLIFE	URBAN
FOREST	1.00		
WILDLIFE	-0.11	1.00	
URBAN	-0.21	-0.11	1.00

Table 2 (b): Boating Amenities

	PAVED	MTROLL	LAUNCH	PARKING	NLAUNCH
PAVED	1.00				
MTROLL	0.27	1.00			
LAUNCH	0.75	0.19	1.00		
PARKING	0.46	0.17	0.46	1.00	
NLAUNCH	0.38	0.45	0.32	0.24	1.00

Table 2 (c): Other Amenities

	RESTROOM	PIER	PICNIC
RESTROOM	1.00		
PIER	0.29	1.00	
PICNIC	0.63	0.30	1.00

Table 2 (d): Catch Rates

	TROUT	SMALLMOUTH	WALLEYE	NORTHERN	MUSKY	SALMON	PANFISH	LARGEMOUTH
TROUT	1.00							
SMALLMOUTH	0.24	1.00						
WALLEYE	0.19	0.19	1.00					
NORTHERN	0.16	-0.08	-0.10	1.00				
MUSKY	-0.15	-0.10	0.09	0.21	1.00			
SALMON	0.14	-0.02	-0.15	-0.24	-0.08	1.00		
PANFISH	-0.58	-0.47	-0.40	0.10	0.27	-0.27	1.00	
LARGEMOUTH	-0.52	-0.50	-0.36	-0.05	-0.02	-0.24	0.69	1.00

Table 3

Distribution of Visitor Shares by Period[23]

	Percentile									
	10	20	30	40	50	60	70	80	90	100
Weekdays	0.00	0.00	0.00	2.61×10^{-4}	5.21×10^{-4}	8.55×10^{-4}	1.36×10^{-3}	2.08×10^{-3}	4.08×10^{-3}	7.80×10^{-2}
Weekends	0.00	0.00	2.68×10^{-4}	4.91×10^{-4}	7.44×10^{-4}	1.07×10^{-3}	1.49×10^{-3}	2.29×10^{-3}	4.21×10^{-3}	2.82×10^{-2}

[23] Each row of this table shows the percentage of sites with fewer than a certain share of the total number of trips taken within a particular period. 33% of the sites have no trips taken on a weekday, while 21.6% of sites have no trips taken on a weekend.

Table 4 – First-Stage Parameter Estimates
Maximum Likelihood Estimation

Angler Attribute	Site Attribute	Weekdays (n = 1831)		Weekends (n = 1750)	
		Estimate	Standard Error	Estimate	Standard Error
BOAT OWNER	PAVED	1.078	0.012	0.639	0.016
BOAT OWNER	WILDLIFE	-0.907	0.035	-0.441	0.012
BOAT OWNER	FOREST	-0.181	0.007	-0.638	0.020
BOAT OWNER	URBAN	-1.409	0.114	-1.004	1.231
KIDS	RESTROOM	0.071	0.006	0.004	0.005
BOAT OWNER	RIVER	-0.498	0.016	-0.589	0.008
BOAT OWNER	SMALL LAKE	-0.296	0.023	-0.756	0.023
KIDS	PANFISH	0.037	0.003	0.096	0.011
BOAT OWNER	SHARE (x 100)	0.412	0.011	-0.038	0.004
UNEMPLOYED	Ln(TRAVEL COST)	0.047	0.001	-0.085	0.002
	Ln(TRAVEL COST)	-2.729	0.002	-2.425	0.002

Table 5 – Second-Stage Parameter Estimates[24]
IV Median Regression, Smoothed GMM (σ = 0.23)

	Weekdays (n = 1831)		Weekends (n = 1750)	
	Estimate	Standard Error	Estimate	Standard Error
CONSTANT	-8.018***	2.092	-2.992***	0.945
PAVED	0.211	0.523	0.454***	0.225
WILDLIFE	0.505	0.736	0.666***	0.285
FOREST	-0.065	0.585	0.379	0.319
URBAN	0.180	0.327	0.370*	0.269
RESTROOM	1.432***	0.383	0.440***	0.188
RIVER	-0.014	1.322	-1.748***	0.849
SMALL LAKE	-1.782***	0.626	-0.506*	0.310
TROUT	3.045**	1.834	-0.080	1.110
SMALLMOUTH	1.816***	0.658	0.301	0.441
WALLEYE	4.714***	1.240	3.040***	1.044
NORTHERN	5.601***	2.109	2.529	2.130
MUSKY	-6.252***	3.157	-1.056	11.828
SALMON	10.429*	7.439	-0.635	3.616
PANFISH	1.063	0.845	-0.386*	0.285
LARGEMOUTH	2.202	1.197	0.270	1.541
SHARE (x 100)	-1.919***	0.808	-2.742***	1.223
Test of Overidentifying Restrictions Critical Value: $\chi^2_{(16)} = 23.5$ at $\alpha = 0.1$	10.149		18.869	

[24] Heteroskedastic-consistent standard errors. Instruments for SHARE (x 100) include predicted share based on exogenous attributes, predicted share squared, and predicted share interacted with exogenous attributes. *** = P < 0.05, ** = P < 0.1, * = P < 0.2.

Table 6 – Second-Stage Parameter Estimates (No Instruments for Share)[25]
Median Regression, Smoothed GMM (σ = 0.23, n = 569)

	Weekdays (n = 1831)		Weekends (n = 1750)	
	Estimate	Standard Error	Estimate	Standard Error
CONSTANT	-8.068***	1.224	-4.581***	0.782
PAVED	-0.046	0.693	0.522**	0.274
WILDLIFE	0.532	0.738	0.586***	0.286
FOREST	-0.071	0.671	0.371*	0.282
URBAN	-0.204	0.529	-0.235	0.245
RESTROOM	1.053***	0.401	0.090	0.184
RIVER	1.041	1.120	0.140	0.501
SMALL LAKE	-2.368**	1.406	0.218	0.267
TROUT	3.626***	1.290	-0.015	1.553
SMALLMOUTH	0.658	1.955	0.181	1.253
WALLEYE	2.006***	0.827	0.516	0.651
NORTHERN	2.492	4.540	-1.573	3.429
MUSKY	0.298	2.624	3.483	8.521
SALMON	7.441***	2.885	0.027	3.213
PANFISH	1.424***	0.414	0.225	0.256
LARGEMOUTH	0.421	1.720	-0.586	1.052
SHARE (x 100)	2.878***	0.817	3.287***	0.633

Table 7
First-Stage Parameter Estimates – No Congestion Effects
Maximum Likelihood Estimation

Angler Attribute	Site Attribute	Weekdays (n = 1831)		Weekends (n = 1750)	
		Estimate	Standard Error	Estimate	Standard Error
BOAT OWNER	PAVED	0.740	0.009	0.632	0.017
BOAT OWNER	WILDLIFE	-0.801	0.018	-0.426	0.011
BOAT OWNER	FOREST	-0.342	0.007	-0.628	0.008
BOAT OWNER	URBAN	-1.304	0.183	-1.011	0.331
KIDS	RESTROOM	0.041	0.006	-0.017	0.006
BOAT OWNER	RIVER	-0.570	0.009	-0.580	0.008
BOAT OWNER	SMALL LAKE	-0.664	0.067	-0.739	0.022
KIDS	PANFISH	0.027	0.003	0.094	0.003
UNEMPLOYED	Ln(TRAVEL COST)	0.025	0.001	-0.083	0.002
	Ln(TRAVEL COST)	-2.731	0.002	-2.425	0.002

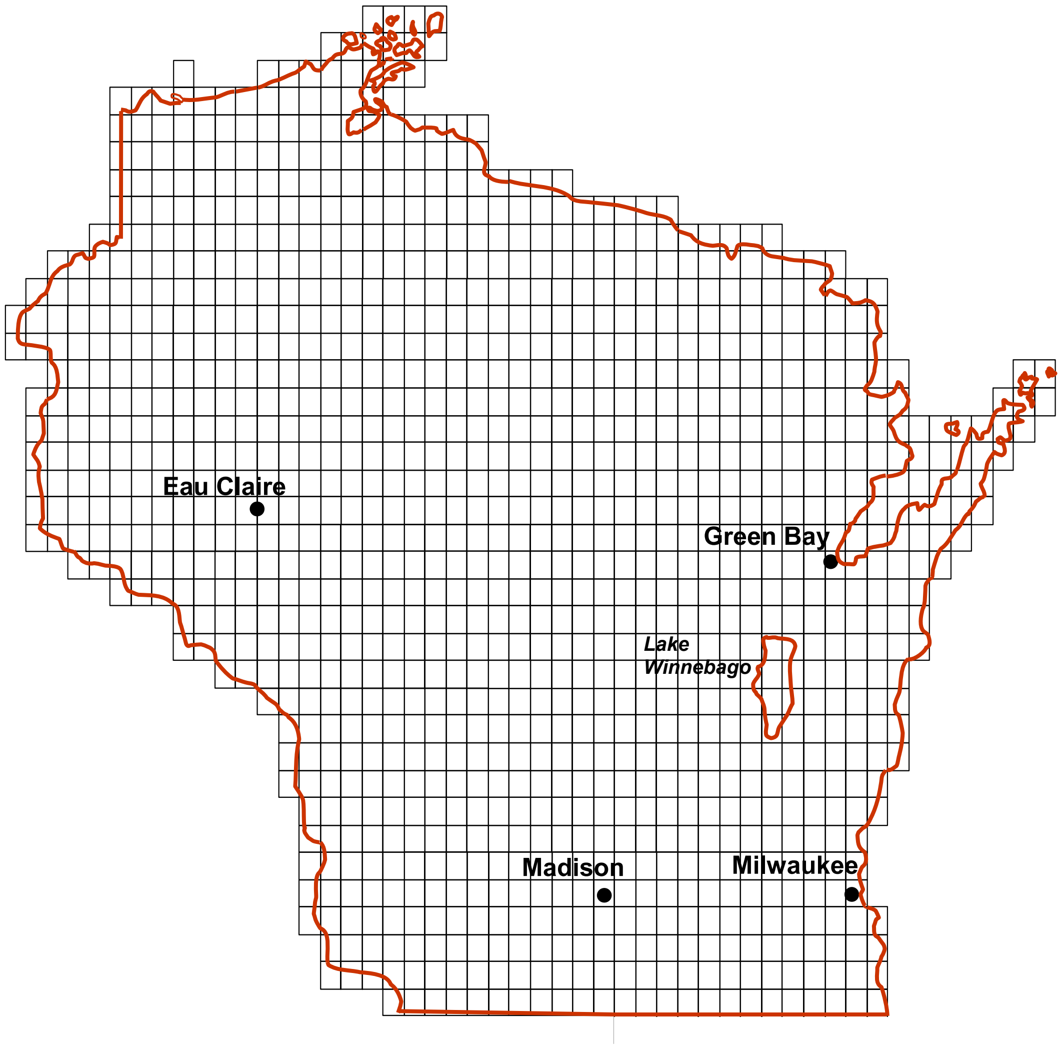
[25] *** = P < 0.05, ** = P < 0.1, * = P < 0.2.

Table 8
Second-Stage Parameter Estimates – No Congestion Effects[26]
Median Regression, Smoothed GMM (σ = 0.23, n=569)

	Weekdays		Weekends	
	Estimate	Standard Error	Estimate	Standard Error
CONSTANT	-8.702***	1.36	-3.567***	0.75
PAVED	0.035	0.53	0.386*	0.24
WILDLIFE	0.892*	0.66	0.526***	0.26
FOREST	-0.287	0.59	0.465*	0.31
URBAN	0.133	0.31	0.269*	0.19
RESTROOM	1.301***	0.41	0.395***	0.18
RIVER	0.915	0.96	-0.642	0.64
SMALL LAKE	-2.113*	1.36	-0.156	0.24
TROUT	3.410***	0.97	-1.141	1.20
SMALLMOUTH	1.736***	0.66	0.162	1.46
WALLEYE	3.759***	0.85	1.671***	0.57
NORTHERN	4.257***	2.10	0.036	4.00
MUSKY	0.929	10.68	11.039***	5.46
SALMON	12.007***	4.26	0.589	3.06
PANFISH	1.527***	0.57	-0.253	0.27
LARGEMOUTH	1.745	1.54	0.545	1.03

[26] *** = P < 0.05, ** = P < 0.1, * = P < 0.2.

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



Appendix: Estimates Under Alternative Assumptions About Zero Shares

Table A1

Weekday Second Stage Parameter Estimates Under Alternative Values of ε

IV Median Estimation and Two-Stage Least Squares (n = 569)

	IV Median		Two-Stage Least Squares							
	All ε		$\varepsilon = 10^{-3}$		$\varepsilon = 10^{-6}$		$\varepsilon = 10^{-9}$		$\varepsilon = 10^{-12}$	
	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err
CONSTANT	-8.02	2.09	-11.76	2.32	-15.75	3.41	-19.75	4.51	-23.746	5.62
PAVED	0.21	0.52	0.37	0.75	0.60	1.11	0.84	1.46	1.076	1.82
WILDLIFE	0.51	0.74	0.61	1.24	0.82	1.83	1.03	2.42	1.237	3.02
FOREST	-0.07	0.59	-0.53	0.78	-0.92	1.14	-1.32	1.51	-1.711	1.88
URBAN	0.18	0.33	0.70	0.81	0.95	1.19	1.21	1.58	1.456	1.97
RESTROOM	1.43	0.38	1.49	0.62	2.18	0.91	2.86	1.21	3.550	1.50
RIVER	-0.01	1.32	0.07	1.73	0.42	2.55	0.77	3.38	1.118	4.21
SMALL LAKE	-1.78	0.63	-1.75	0.87	-2.53	1.28	-3.30	1.69	-4.079	2.10
TROUT	3.05	1.83	3.38	2.24	4.24	3.29	5.10	4.35	5.962	5.42
SMALLMOUTH	1.82	0.66	1.88	1.67	2.23	2.45	2.59	3.25	2.938	4.04
WALLEYE	4.71	1.24	1.92	2.69	1.14	3.95	0.37	5.23	-0.406	6.51
NORTHERN	5.60	2.11	7.59	5.62	10.70	8.26	13.82	10.94	16.939	13.63
MUSKY	-6.25	3.16	1.56	14.37	-0.86	21.13	-3.29	27.98	-5.71	34.86
SALMON	10.43	7.44	12.37	9.07	17.52	13.34	22.67	17.66	27.82	22.00
PANFISH	1.06	0.85	1.05	0.88	1.41	1.30	1.77	1.72	2.12	2.14
LARGEMOUTH	2.20	1.20	-0.22	3.29	-0.50	4.84	-0.78	6.41	-1.06	7.98
SHARE (x 100)	-1.92	0.81	2.49	1.46	4.37	2.14	6.24	2.84	8.12	3.53

Figure A1
Distribution of Weekday Fixed Effects ($\delta_{j,WD}$)
Under Alternative Values of ε