WORKING PAPER SERIES*

DEPARTMENT OF ECONOMICS

ALFRED LERNER COLLEGE OF BUSINESS & ECONOMICS

UNIVERSITY OF DELAWARE

WORKING PAPER NO. 2008-02

STATE DEPENDENCE AND LONG TERM SITE CAPITAL IN A RANDOM UTILITY MODEL OF RECREATION DEMAND*

D. Matthew Massey and George R. Parsons

*http://lerner.udel.edu/economics/workingpaper.htm

State Dependence and Long Term Site Capital in a Random Utility Model of Recreation Demand*

D. Matthew Massey
National Center for Environmental Economics
US Environmental Protection Agency, Washington, DC 2460

GEORGE R. PARSONS
College of Marine Studies and Dept. of Economics
University of Delaware, Newark, DE 19716

JEL: Q26
Key Words: Random Utility Model; State Dependence; Non-Market Valuation

Abstract

Conventional discrete choice Random Utility Maximization (RUM) models of recreation demand ignore the influence of knowledge, or site capital, gained over past trips on current site choice, despite its obvious impact. We develop a partially dynamic RUM model that incorporates a measure of site capital as an explanatory variable in an effort to address this shortcoming. To avoid the endogeneity of past and current trip choices, we estimate an auxiliary instrumental variable regression to purge site capital of its correlation with the error terms in current site utility. Our instrumental variable regression gives a fitted value ranging between 0 and 1 for each alternative for each person – a prediction of whether or not a person visited a site. Results suggest that the presence of accumulated site capital is an important predictor of current trips, and that failure to account for site capital will likely lead to underestimates of potential welfare effects.

1. Introduction and Background

Conventional discrete choice Random Utility Maximization (RUM) models of recreation demand ignore the influence of past trips on current site choice.[1] Yet, there is little doubt that past experiences shape a person's utility on future trips. A person knows more about the characteristics of the sites they have visited in the past – both characteristics observed and unobserved by the researcher. A person knows more about the costs of access, best travel routes, best places for parking, and so forth. The time and search costs needed to plan and access a site visited in the past are no doubt lower than for a site never visited. Also, because of extra site-specific knowledge a person has less uncertainty about what a trip to a site will be like (whether positive or negative). Following this reasoning, failure to account for the effects of knowledge gained during past visits could easily lead to a model that misrepresents behavior.

The process of past choices influencing current choices has been extensively examined in a number of disciplines and has variously been dubbed state dependence, temporal dependence, or habit formation. In his seminal labor market paper Heckman (1981) defined state dependence as the situation where "past experience has a genuine behavioral effect in the sense that an otherwise identical individual who did not experience the event would behave differently." In practice, state dependence can be difficult to model because of its dynamic nature and the fact that unobserved preference heterogeneity can lead to spurious state dependence-like outcomes where individuals repeatedly select the same option. Researchers

[1] For an assortment of applications of the RUM model to recreation demand see Lin, Adams, and Berrens (1996), Loomis (1995), Parsons and Massey (2000), and Landry and Liu (2007).

have predominantly relied on some form of fully or partially dynamic repeated choice models to capture state dependence effects (Pollak 1970; Rust 1987; Smith 2005).[2]

Despite its obvious applicability however, there have been only a handful of attempts to account for state dependence in recreation demand models. Those who have attempted to model the effects of past choices in recreation demand frameworks have motivated their studies with a number of different assumptions and model structures. One branch of the literature has attempted to estimate fully dynamic models. Adamowicz (1994) for example, adapts Pollack's theoretical habit formation model to recreational fishing by assuming that an individuals recreational opportunities may be viewed as stock of goods that is consumed and depreciated over time. In each time period individuals choose whether to consume the available stock or carry some of it over into the next period. Addressing a similar topic in a very different way, Provencher and Bishop (1997) adapt Rust's dynamic optimal stopping model of bus engine replacement to recreational fishing trip demand by assuming that individuals maximize expected daily utility subject to daily budget constraints (derived from a seasonal budget constraint) over the course of a season. Expected daily utility is also assumed to include discounted expected future trip utility conditioned on the current choice. Current choices are influenced by past choices through a variable measuring the days since an individual's last trip and through expected catch predictions that are influenced by past trip catch totals. Both of these studies consider past trips only within a given season.

[2] Fully dynamic models are those models that assume individuals consider both the effects of past decisions on current decisions and the effects of current decisions on future decisions. Partially dynamic models generally only consider the effects of past choices on current choices.

More recently several researchers have turned to partially dynamic model structures in order to capture state dependence and preference heterogeneity. Moeltner and Englin (2004) and Swait et al. (2004) both modify the standard repeated choice logit to incorporate temporal effects. Swait et al. estimate a meta-utility function made up of weighted current and past period utilities. These utilities include previous choices and expected attribute levels constructed of past realizations and current expectations. Moeltner and Englin include variables measuring the total number of times a given option was chosen and the number of consecutive times an option was chosen in order to capture the state dependence effects. The authors also use a random parameters (mixed logit) model structure to deal with the unobserved preference heterogeneity that can lead to spurious state dependence findings.

Not surprisingly, the common finding of all the studies is that the inclusion of past experiences matters in estimation and welfare results. Two common characteristics among these studies are (1) a reliance on large panel data sets in which the researcher knows the order and timing of every decision made (i.e. logbooks or diaries) and (2) a relatively complicated estimation procedure particularly among the fully dynamic models. These two issues are important reasons why none of these methodologies have been fully embraced by practitioners. Recreation demand panel data sets are relatively rare compared to other survey types because they are more time consuming and expensive to collect. Previous dynamic models have been so hard to estimate that they usually require assumptions and concessions that substantially reduces their practical usefulness (Phaneuf and Smith 2004; Swait, Adamowicz et al. 2004). Furthermore, Adamowicz's results suggest that there is little difference in between fully and partially dynamic models empirically, bringing into doubt whether the extra estimation difficulty is even worthwhile.

Although it has not received much attention in the literature to date, researchers also face the task of defining an appropriate measure of alternative specific experience to test for state dependence. In many cases, studies have simply used some version of past trips as a measure of previous experience. The use of past trips is problematic because past trips are likely correlated with unobserved site characteristics that guided choices in the past and that are possibly still present for current choices. To isolate state dependence effects the unobserved correlation must be purged from the measure of past experience.

To avoid these past problems and complications, we propose an alternative partially dynamic modeling method that is relatively easy to estimate and requires little additional data. Similar to previous researchers, we develop a RUM model of site choice that incorporates information on visits to sites in the past. Following Becker and Murphy's (1986) terminology we refer to past visits to a site as 'site capital'. Since we use a dummy variable for weather or a not a person has ever visited a site in a year prior to the current season as our measure of site capital for a site, we refer to it as 'long term' site capital. To avoid the endogeneity of past trips with current trip choice, we estimate an auxiliary instrumental variable regression to purge site capital of its correlation with the error terms in current site utility. Our instrumental variable regression gives a fitted value ranging between 0 and 1 for each alternative for each person – a prediction of whether or not a person visited a site or has any site capital at the site. The fitted value, then, is used in place of the past visit dummy variable and is, in principle, purged of its correlation with the site utility error terms in the model.

We compare four versions of our RUM model: (1) a basic model that ignores past trips, (2) a model that incorporates past trips but does not correct for the endogeneity of site capital, (3) a model that incorporates past trips and corrects for the endogeneity of site capital using a 'short' instrumental variable regression, and (4) a model that incorporates past trips and corrects for the endogeneity of site capital using a 'long' instrumental variable regression. By short and long we are referring to the number of instruments used in the auxiliary regression. The short regression uses a few key instruments and the long regression uses all appropriate available variables. Comparing the results using two different instruments allows us to explore the sensitivity of our results to the choice of instruments. We also estimate all our models in a random parameters framework in order to account for preference heterogeneity over the influence of past trips. Lastly, we consider differences in parameter and welfare estimates across the four models. Our welfare scenarios include the closure of individual beaches, the closure of groups of contiguous beaches, and the narrowing of groups of contiguous beaches.

2. Models and Study Design

In our *Baseline Model* individuals have no memory, and the model is described by the indirect utility functions

$$(1)\quad \begin{array}{l} V_i = b_{tc} tc_i + b_x x_i + e_i \\ V_0 = b_y y + e_0 \end{array},$$

where V_i is the site utility for a trip to site i on a given choice occasion ($i = 1,\ldots,62$) and V_0 is the utility of doing something other than taking a trip on a given choice occasion. There are 62 sites in our application. The arguments in the model are trip cost, tc_i, a vector of site characteristics,

x_i, and a vector of individual characteristics, y. The site characteristics are intended to capture aspects of the site that matter to individuals in selecting a destination and the individual characteristics are intended to capture characteristics of individuals that help predict their probability of taking a trip. b is the coefficient vector to be estimated. b is assumed to vary across the population with the distribution $f(b\,|\,q)$, where q contains the parameters of the b distribution.[3]

If the error term e is assumed to be distributed identically and independently according to the extreme value distribution, then the probability that a participant chooses site k on a particular choice occasion is given by the integral of the logit formula evaluated at all possible values of b,

$$(2) \qquad PR(k) = \int \left[\frac{\exp(V_k)}{\exp(V_0) + \sum_{i=1}^{62} \exp(V_i)} \right] f(b\,|\,q)\,db\,.$$

Due to the analytical difficulty of evaluating multiple integrals, simulation is generally required to obtain results. Equation (2) may be simulated by

$$(3) \qquad SP(k\,|\,q) = \frac{1}{R} \sum_{r=1}^{R} \left[\frac{\exp(V_k^r)}{\exp(V_0^r) + \sum_{i=1}^{62} \exp(V_i^r)} \right],$$

where R is the number of draws of b from $f(b\,|\,q)$, and V^r is indirect utility calculated with draw r of b. The simulated probabilities may then be used to construct a simulated log

[3] Because b is assumed to vary across the population, it is often written with an n subscript. The participant index n is supressed in this case in an effort to make interpretation of the remaining notation more straightforward.

likelihood function that may be maximized to produce estimates of parameters of the b distribution, q .

Models 2 through 4 extend the *Baseline Model* by introducing an individual's long term site capital as an explanatory variable. In all models, site capital enters the utility for each site as an alternative specific constant <u>and</u> as an interaction with the vector of site characteristics. As an alternative specific constant the site capital measure allows site utility to shift depending on whether an individual has visited that site in the past. As an interaction term, it allows the coefficients on the site characteristics to differ for sites with site capital versus those without. Specifically, in models 2 through 4 indirect utility is specified as

$$
\begin{aligned}
V_i &= b_{tc} tc_i + a d_i + d_i \left(b_{capt} x_i\right) + \left(1 - d_i\right)\left(b_{no\,capt} x_i\right) + e_i \\
V_0 &= b_y y + e_0
\end{aligned}, \qquad (4)
$$

where $d_i = 1$ if a person visited site *i* at some time in their adult life prior to the current season, and $d_i = 0$ if not. We refer to d_i as an individual's long term site capital for site *i*. Again, it is long term because it only accounts for the effect of trips in past seasons on current site choice. It does not account for the effect of trips taken earlier in the current season on site choice. In this way our model is like McConnell, Strand, and Bockstael (1990) who consider long terms effects only and unlike Provencher and Bishop (1999), Adamowicz (1994), and Swait et al (2004) who consider short term effects only. While the lack of short term considerations is a shortcoming of the model, focusing solely on long term habit capital greatly reduces the models data requirements. Furthermore, if preferences are thought to be stable over time, then long and short term preferences should be good approximations of one another.

We expect $a > 0$, which indicates (all else constant) that sites with site capital have higher utility than sites without. This implies long term habit formation and is consistent with McConnell, Strand, and Bockstael (1990). A negative coefficient would imply variety seeking. We also expect the site characteristics for sites with capital (past visits) to play a more important role in current site choice than the site characteristics on sites without capital. Individuals are more knowledgeable about the characteristics at these sites and hence are more likely to use this information in determining choice over these sites. For sites without capital, site characteristics are likely to play a smaller role. For many of these sites, individuals may only have rough guesses about site characteristics. This would imply that $\boldsymbol{b}_{capt}$ have greater explanatory power in the site choice model than $\boldsymbol{b}_{no\ capt}$.

Model 2, or the *Exogenous Model*, uses our most basic measure of site capital, which is simply a dummy variable denoting whether or not a person has ever visited a site in the past. The third and fourth models are identical to the second except that that they treat the alternative specific site capital measures as endogenous. Accounting for this endogeneity may be important since past trips (our simple site capital measure in the *Model 2*) are likely to be highly correlated with the unobserved characteristics of current site utility. Or in other words, unobserved characteristics that influence site choice today were likely to have influenced site choice in the past. A model that ignores this endogeneity will yield biased and inconsistent parameter estimates and possibly incorrectly attribute repeated choices to state dependence. Therefore, in the 3rd and 4th models, we purge the past trip variable of its correlation with current error terms using an instrumental variables regression.

Following Imbens and Angrist (1994) and Angrist and Krueger (2001) we estimate the instrumental variable regression using ordinary least squares. A vector of site and participant characteristics z_{in} is regressed on responses to the question whether or not a person has ever visited a site in the past (*PASTT*). The model may be formally written,

$$(5) \qquad PASTT_{in} = f(z_{in}, j\,),$$

where j is a vector of estimated parameters. The model has 562x62 observations -- one beach for each person. The *Endogenous Models* (*Model 3* and *4*) differ by the set of instruments included in z_{in}. *Model 3* uses a short list of key instruments. *Model 4* uses all available appropriate instruments. Comparing *Models 2*, *3*, and *4* allow us to test how sensitive our results are to the choice of instruments.

Welfare effects are calculated for all models by monetizing changes in expected utility due to access or quality changes at one of more sites -- see Phaneuf and Smith (2004) for a presentation of welfare formula in discrete choice random utility models. Because the mixed logit model estimates the distributions of the coefficients, calculating the welfare effects of changes to sites in the choice set again requires simulating integration. For example, the expected welfare change for individual n associated with a change in quality at some or all of the 62 sites would be:

$$(6)\ W_n = \frac{1}{D}\sum_{d=1}^{D}\left[\frac{\ln\{\exp(V_{0n}) + \sum_{i=1}^{62}\exp[V_{in}^{*}(b_d \mid q)]\} - \ln\{\exp(V_{0n}) + \sum_{i=1}^{62}\exp[V_{in}(b_d \mid q)]\}}{b_{tc}}\right]$$

where D is the total number of draws from the estimated distributions, b_d is draw d from the distribution of q, V_{in}^{*} is expected maximum utility calculated with a quality change, b_{tc} is the travel cost coefficient, and the numerator is the difference in the expected maximum utility per choice occasion between the current and changed conditions at some or all of the sites. We use this formula for our beach narrowing scenarios. The formula for the loss of one or more sites is similar and takes the form

$$(7)\ W_n = \frac{1}{D}\sum_{d=1}^{D}\left[\frac{\ln\{\exp(V_{0n}) + \sum_{i=1}^{L}\exp[V_{in}(b_d \mid q)]\} - \ln\{\exp(V_{0n}) + \sum_{i=1}^{62}\exp[V_{in}^{*}(b_d \mid q)]\}}{b_{tc}}\right]$$

where *L (< 62)* is the number of sites that remain open. We use this formula for all our site closure scenarios. Our seasonal measures of loss, reported in a later section, are simply $240 * W_n$, where 240 is the number of choice occasions in the season.

3. Data

In the Fall of 1997, with funding from the National Oceanic and Atmospheric Administration, we conducted a mail survey of Delaware residents over the age of 16. Individuals were asked to report their number of trips to 62 ocean beaches in the Mid-Atlantic region since January 1, 1997 and to indicate which beaches they had visited in past years. The beaches included all of New Jersey, Delaware, and Maryland’s ocean beaches. Assateague Island, which is partially in Virginia, was also included. Figure 1 shows the region covered in our analysis and Table 1 provides a list of beaches by name running from north to south. People were also asked to report household information such a location of hometown, age, family composition, employment, and so forth. Individual characteristic summary statistics are presented in Table 2. In our analysis we consider both participants and non-participants and focus on day-trips. Of the 562 respondents, 397 took at least one day-trip to one of the 62 beaches. The total number of day-trips taken in the sample was 8034.

For each of the 62 beaches, we gathered the characteristic data listed in Table 2. We used a variety of resources to compile the data set including travel guides, field trips, interviews with resource managers in Delaware and New Jersey, and geological maps. The resource managers were particularly helpful; not only in compiling the data but also in deciding what characteristics are likely to matter to individuals in choosing a beach. Table 3 reports summary statistics for all site characteristics used in the m odel. Table 4 reports summary statistics for the individual characteristics. The average trip cost to a Delaware beach was about $50 and to a New Jersey beach was about $150. On average, a person has about 4 beaches with site capital – 10% have zero and 10% have more than 16. For more detailed descriptions of the data and survey design process see Massey (2002) and Parsons et al (2000) .

4. Estimation Results

We begin by estimating a standard baseline travel-cost RUM model that does not account for the effects of past trips on current trips. The parameter estimates on the *Baseline Model* tell a plausible story and are consistent with our earlier work with these data (see Parsons et al. (2000) and Parsons and Massey (2002)). As reported in Table 5, site utility increases with boardwalks, amusements, parks, surfing, park within, and parking. All are features of beaches that we anticipated would improve the desirability of the site. Among these, amusements and park within have the highest relative values. Site utility declines with travel cost, private, narrow, wide, high rise, and facilities. Private beaches tend to have less access for non-residents for day trips due to limited access. Beaches that are too narrow or too wide are generally less desirable. Beaches with high rises (the more developed beaches) tend to have larger overnight and smaller day trip visitation. We also included dummy variables for Atlantic City and New Jersey to capture their distinct character. Atlantic City, a mecca for gambling and nightlife in the area, increases site utility in our model. New Jersey reduces site utility. Both results were expected. Beach length is the only insignificant site characteristic coefficient, but it does have the anticipated positive sign. The only outcome that ran counter to our expectations was the negative coefficient on facilities.

The individual characteristic data in the *Baseline Model* shows that no-trip utility increases with working from home, working part time, and retirement. Conversely, the probability of taking a trip rises with having kids, having flexible work hours, being a student, or being a volunteer. The coefficients on these variables were statistically significantly different from zero in all cases except kids under 10.

Due to the large number of parameters in the estimated models, we only allow two parameters to be random in our mixed logit estimation: no-trip constant and site capital.[4] Site capital is not included in the baseline model and interestingly, the no-trip constant's estimated deviation is not statistically significant as expected. We had expected that there might be considerable unobserved heterogeneity over taking a trip versus not taking a trip.

[4] When we allow more parameters to be free in our repeated logit setting with no-trip included as a choice, we continually ran into convergence and singularity problems. Limiting the model to a few site characteristics or restricting the model to be site choice only without participation helped, but we felt the sacrifice here in terms of a useful model for policy applications was too high. In the trade off between adding unobserved heterogeneity and having a richer behavioral model (with more observed site characteristics, participation, and splitting the sites with and without site capital into two separate groups), we opted for the latter.

Next we extend the *Baseline* model by estimating three new versions of the model that incorporate three different measures of site capital. *Model 2* uses the most basic site capital measure, which is simply a dummy variable indicating whether or not a person has ever visited a site in a previous season. While conceptually appealing and easily implemented in practice, past trip choices are likely to be correlated with current trip choices thereby creating an endogeneity problem. To deal with this endogeneity, *Model 3* and *Model 4* utilize measures of site capital calculated from instrumental variable regressions. In each model, a set of instruments is used to predict past trip visitation (as measured by the past trip dummy from *Model 2*). As shown in Table 6, the two measures differ by the number of instruments used in estimation. This follows Becker and Murphy (1986). The fitted values for site capital in Model 3 only includes a short set of instruments, none of which appear in the *Baseline* RUM model, while the fitted value for site capital in *Model 4* include a long set comprised of the short set plus the explanatory variables used to predict trips in the *Baseline* RUM model. Results of the instrumental variable regressions are plausible and relatively consistent across the two models for the instruments they share. The exception is the distance variable which, as expected, losses size, significance, and even sign when trip cost is included in the long regression. Each measure of site capital is the incorporated into a separate modified version of the *Baseline* model as a regressor. *Model 2* uses the exogenous past visit dummy variable, *Model 3* utilizes the endogenous short instrument, and *Model 4* relies on the endogenous long instrument.

An ideal instrument is one that is correlated with site capital (or taking a trip to a site in the past) but uncorrelated with the current site utility error term. While far from perfect, our set of four variables are plausible – distance to a beach, age, owning a vacation home near the beach, and income. Vacation home is probably the weakest on this list as a pure instrument since choice of site and choice of a location to own a beach home may be governed by similar excluded attributes. As is conventional we also include all the regressors from the original model in the long version instrumental variable regression.

The three site capital RUM models tell much the same story as the *Baseline Model*, however, they do provide significant support to the hypothesis that site capital accumulated on past trips does affect current choices. As we expected the coefficients on sites without site capital have much less explanatory power than the coefficients on sites with site capital – notice the number coefficients with unexpected signs and without statistical significance on the sites without site capital. This stands to reason as people have little experience over the sites they have not visited in the past and hence are less able to base current site choice on site characteristics. On the other hand, for sites they have visited in the past, site characteristics are well know and play an important role in current site choice.

This dynamic may be seen in sign changes that several variables undergo when separated into visited and unvisited sites. For example, in the *Baseline Model*, private beaches reduce average utility, but when site capital is accounted for private beaches actually increase utility in previously visited sites. Similarly, New Jersey beaches decrease utility in the Baseline model while they increase utility if they have been previously visited. Going in the other direction, beach length is insignificant in the basic model, but strongly positive and significant for unvisited beaches. Surfing also makes a noticeable shift in sign and is significant. The results certainly suggest that beachgoers treat the characteristics of sites visited in the past differently from sites never visited in making current site choices -- a reasonable and expected result.

Also as expected, in all versions of the site capital model, the site capital coefficient is positive and significant indicating habit formation for site choice. Surprisingly though, all the models predict very little deviation in site capital preferences. The lack of deviation suggests a fair degree of unobserved homogeneity among beach goers and that few, if any, beachgoers in the data set are variety seekers.

To make direct comparisons across the models, we calculate the implicit prices for each coefficient in each model. In discrete choice models, absolute values across models are not comparable, but values relative to a common coefficient (in our case price) are comparable. These ratios can also be interpreted as implicit prices for the attributes – the value an attribute holds assuming a person is constrained to visit the site. As the Table 7 shows the implicit prices fluctuate significantly at times across the four models. Results also show that the endogenous site capital models (*Models 3* and *4*) predict that site capital is one to one and a half times more valuable than the exogenous site capital (*Model 2*). This rather sizeable increase in the value of site capital between the exogenous and endogenous models indicates a fair degree of correction for endogeneity.

Finally, it is important to note that the trip cost coefficient, which plays a major role in valuation as the marginal utility of income in the denominator of equations (6) and (7), declines once the corrected site capital variable is included in the models – compare *Models 1 & 2* versus *Models 3 & 4*. This implies that trip cost plays a less important role in site choice than conventional models would suggest -- trip cost, in effect, is picking up some site capital effects in conventional models. Once included in the model and corrected, we see the trip cost coefficient fall. This will lead to larger welfare estimates in the site capital versions of the model in the next section.

5. Welfare Estimates

With a few exceptions, travel-cost random utility models are estimated for the purpose of valuing site access or changes in site characteristics. With this in mind, we consider how welfare measures (presented in Table 7) vary across our models. We consider four welfare scenarios: the loss of a group of sites, the loss of beach width, the loss of a few selected single sites, and the loss of site capital.

The most important and striking result is certainly the finding that failure to account for site capital leads to lower welfare estimates. In almost all cases in, the *Baseline* model, which does not account for individuals' accumulated site capital, predicts the smallest welfare effects of all the estimated models. If people have little or no site capital for a given site or sites, as is the case with the least visited sites in the choice set, then the baseline and site capital models return very similar welfare estimates. However, as the level of accumulated site capital increases for a given site, the baseline and site capital models' welfare estimates begin to diverge. At the extreme, the site capital models' welfare estimates for the loss of the most visited beaches range from roughly one and third to two times larger than the baseline model.

The second main result that emerges from the welfare results is that failure to account for endogeneity in the site capital measure will also lead to smaller welfare estimates. In almost all cases, the *Exogenous Model* (Model 2) predicts smaller welfare effects than the two *Endogenous*

Models (3 and 4). It is also obvious from the results that the choice of site capital instruments can have a significant effect on welfare results. The *Endogenous Short Model* (Model 3) returns the largest welfare predictions in every case. While consistently larger than the *Baseline Model* estimates, the *Endogenous Long Model's* welfare effects are actually closer in magnitude to the *Exogenous Model* than they are to the *Endogenous Short Model*.

The results appear to be driven by two factors. As noted above, the coefficient estimate on trip cost is lower in the site capital models implying that models without site capital inadvertently attribute too much explanatory power to trip cost. Indeed, people overwhelming tend to visit closer sites, but when site capital is accounted for we see that much of this is actually due to people having visited close sites in the past. Hence, some of the trips to nearby sites are due, at least in part, to site capital. Second, the coefficient on the site capital term is large in relative terms and increases the utility at sites with already high utility. This, in turn, increases the expected utility of taking a trip to popular beaches relative to other beaches and gives higher welfare losses when the sites are lost or narrowed.

The results also indicate that accumulated site capital is valuable. To measure this value, we estimate a welfare scenario in which we assume that all participants "loose" their accumulated site capital. We find that site capital values range from the mid $600's up to the mid $900's per person.

6. Comments, Caveats, and Conclusions

One of the most attractive features of our application is that it relatively simple to implement compared to previous attempts at modeling state dependence. The past trip information used in the model is easily gathered by a mail or phone survey of the general population. It is not too taxing for individual's to remember whether or not they visited a site in the past. So, it a rather simple adjustment to make to our conventional models, and it appears to matter significantly.

On the downside, our measure of site capital does not account for intensity. For example, our measure treats a site with one trip taken 10 years ago the same a site with 20 trips taken over the past two years. There are a number of ways to improve the measure. For example, one might use the number of past trips to a site, or the number of past years visiting the site, and/or weight recent years more heavily, or even account for quality of the past experience (was site *i* is a beach the person liked or disliked?). Each of these requires information that is more difficult to recall than simply whether or not you have visited the site in the past.

Our measure also fails to account for forward-looking behavior and for any adjustments that may take place over time that may affect the computation of welfare. With forward-looking behavior individuals are viewed as making investments in site capital when they visit a site today. That investment can be used as site capital on future visits to a site, thereby raising future trip utility. If a person visits a site that becomes a favorite, its site utility might increase considerably. We ignore this dynamic completely in our myopic model. Although, as noted, there has been little evidence of forward-looking or variety seeking behavior in past studies. Also, in the computation of welfare when sites are closed or narrowed, people may find themselves visiting new sites and thereb y developing new found site capital. This should work to dampen welfare loses of site closures over time. Our model ignores this dynamic as well and it would seem to be fertile ground for future research in improving models with a dynamic element.

Most importantly, our results suggest that failure to account for past visits and accumulated site capital will likely lead to underestimates of potential welfare effects. Additional research is required to determine whether or not our result will hold in other applications, but intuition and theory suggest they will. Future research may also want investigate ways to formalize the selection of instruments used to purge endogeneity from the past trip variable. The results of this study suggest that estimates are sensitive to instrument choice. Indeed, the validity of the results hinges on the credibility of the instruments successfully purging the endogeneity of past trips. Still, it is move beyond previous research that ignores endogenity entirely.


MID-ATLANTIC REGION
PA
NJ
MD
DE
VA
Delaware Bay
Chesapeake Bay
Sandy Hook
Long Branch
Point Pleasant Beach
Long Beach Island
Atlantic City
Ocean City, NJ
Cape May
Cape Henlopen
Rehoboth Beach
Delaware Seashore State Park
Bethany Beach
Fenwick Island
Ocean City, MD
Assateague Island

Figure 1

Table 1: Mid-Atlantic Beaches from North to South

New Jersey: North Shores

1. Sandy Hook
2. Sea Bright
3. Monmouth Beach
4. Long Branch
5. Deal
6. Asbury Park
7. Ocean Grove
8. Bradley Beach
9. Avon-by-the-Sea
10. Belmar
11. Spring Lake
12. Sea Girt
13. Manasquan

New Jersey: Barnegat Peninsula

14. Point Pleasant Beach
15. Bay Head
16. Mantoloking
17. Normandy Beach
18. Chadwick Beach
19. Ocean Beach
20. Lavallette
21. Ortley Beach
22. Seaside Heights
23. Seaside Park
24. Island Beach State Park

New Jersey: Long Beach Island

25. Barnegat Light
26. Loveladies
27. Harvey Cedars
28. Surf City
29. Ship Bottom
30. Long Beach
31. Beach Haven
32. Holgate

New Jersey: Atlantic City Area

33. Brigantine
34. Atlantic City
35. Ventnor
36. Margate
37 Longport

New Jersey: South Shore

38. Ocean City
39. Strathmere
40. Sea Isle City
41. Avalon
42. Stone Harbor
43. North Wildwood
44. Wildwood
45. Wildwood Crest
46. Cape May

Delaware:

47. Cape Henlopen State Park
48. North Shores
49. Henlopen Acres
50. Rehoboth Beach
51. Dewey Beach
52. Indian Beach
53. Delaware Seashore State Park
54. North Bethany Beaches
55. Bethany Beach
56. Sea Colony
57. Middlesex Beach
58. South Bethany Beach
59. Fenwick Island State Park
60. Fenwick Island

Maryland/Virginia

61. Ocean City, MD
62. Assateague Island

Table 2: Explanatory Variables

SITE CHARACTERISTICS:	
Trip Cost	Travel cost (includes tolls, beach fees, transit costs, and parking fees) + time costs ($.333 \cdot (\text{income} / 2080) \cdot \text{travel time}$)
Length	Length of beach in miles
Narrow	Beach width from dune toe to berm less than 75 feet (1 if yes, 0 if no)
Wide	Beach width from dune toe to berm greater than 200 feet (1 if yes, 0 if no)
Park	State park, federal park, or wildlife refuge (1 if yes, 0 if no)
High Rise	Highly developed (1 if yes, 0 if no)
Private	Private or limited access (1 if yes, 0 if no)
Park Within	Part of the beach is a park area (1 if yes, 0 if no)
Boardwalk	Boardwalk with shops and attractions present (1 if yes, 0 if no)
Amusements	Amusement park, rides, or games available or nearby the beach (1 if yes, 0 if no)
Surfing	Recognized as a good location for surfing (1 if yes, 0 if no)
Facilities	Facilities such as bathrooms, showers, and food available on or just off the beach (1 if yes, 0 if no)
Parking	Presence of adequate parking near beach (1 if yes, 0 if no)
Atlantic City	Beach in Atlantic City, NJ (1 if yes, 0 if no)
New Jersey	Beach located in New Jersey (1 if yes, 0 if no)
INDIVIDUAL CHARACTERISTICS:	
Kids Under 10	Number of children under the age of 10
Kids Between 10-16	Number of children between 10 and 16 years old
Work Part Time	Work part time (1 if yes, 0 if no)
Work at Home	Work at home (1 if yes, 0 if no)
Volunteer	Volunteer (1 if yes, 0 if no)
Flexible Time	Flexible work schedule (1 if yes, 0 if no)
Retired	Retired (1 if yes, 0 if no)
Student	Student (1 if yes, 0 if no)

Table 3: Explanatory Variable Summary Statistics for Beach Characteristics*

	Delaware, Maryland, and Virginia Beaches (16 beaches)	All Beaches (47 beaches)
Continuous Variable Mean Values and Ranges		
*Trip Cost** (1997$)	$49.49	$122.04
	(0.00 to 184.76)	(0.00 to 310.85)
Length (Miles)	1.20 miles	1.86 miles
	(0.40 to 22.00)	(0.40 to 22.00)
Percentage of Beaches With Each Characteristic		
Narrow	6.3%	14.5%
Wide	18.8%	24.2%
Park	25.00%	9.7%
High Rise	6.3%	24.2%
Private	37.5%	25.8%
Park Within	0.0%	14.5%
Boardwalk	6.3%	37.1%
Amusements	12.5%	12.9%
Surfing	43.8%	35.5%
Facilities	50.0%	38.7%
Parking	43.8%	45.2%
Atlantic City	0.0%	1.6%
New Jersey	0.0%	74.2%

* Calculated over 562 people for each beach in the choice set.

Table 4: Explanatory Variable Summary Statistics for Individual Characteristics

Continuous Variable Mean Values and Ranges	
Kids Under 10	.41 kids
	(0 to 6)
Kids Between 10-16	.28 kids
	(0 to 4)
Percentage of Individuals with Each Characteristic	
Work Part Time	10.1%
Work at Home	6.4%
Volunteer	3.2%

Flexible Time	18.5%
Retired	24.6%
Student	5.0%

Table 5: Estimation Results

	MODEL 1: Baseline Model	MODEL 2: Exogenous Model	MODEL 3: Endogenous Model w/ Short IV	MODEL 4: Endogenous Model w/ Short IV
<u>SITE CHARACTERISTICS</u>				
Trip Cost	-0.0378 (64.97)	-0.0324 (59.10)	-0.0206 (30.39)	-0.0287 (51.49)
<u>Site Capital</u> (a in equation 4)				
Site Capital (Mean)	--	3.460 (16.73)	5.281 (23.73)	6.058 (25.80)
Site Capital (Deviation)	--	0.089 (0.90)	0.653 (2.40)	0.151 (.087)
<u>Sites with Capital</u> (b_{capt} in equation 4)				
Length	0.002 (0.04)	-0.025 (0.67)	-0.084 (2.20)	-0.321 (8.25)
Narrow	-0.256 (3.02)	0.129 (1.42)	-0.294 (3.20)	-0.294 (3.23)
Wide	-0.836 (16.01)	-0.550 (10.14)	-0.614 (11.36)	-0.697 (12.96)
Park	0.556 (3.76)	0.503 (2.86)	0.649 (3.54)	0.632 (3.53)
High Rise	-0.476 (7.28)	-0.562 (7.86)	-0.731 (9.81)	-0.962 (13.09)
Private	-0.669 (11.18)	-0.369 (5.82)	0.121 (1.91)	0.500 (7.76)
Park Within	1.549 (14.27)	0.647 (5.68)	0.739 (6.39)	0.759 (6.57)
Boardwalk	0.612 (4.48)	0.532 (3.21)	0.747 (4.31)	0.538 (3.18)
Amusements	1.491 (26.99)	1.007 (17.64)	1.267 (22.14)	0.132 (1.83)
Surfing	0.818 (17.24)	0.574 (10.92)	1.050 (19.76)	0.930 (17.66)
Facilities	-0.308 (3.08)	-0.292 (2.50)	-0.392 (3.28)	-0.256 (2.19)
Parking	0.412 (3.13)	0.200 (1.24)	0.386 (2.29)	0.247 (1.50)
Atlantic City	1.590 (12.71)	0.375 (2.86)	0.604 (4.56)	-0.634 (4.59)
New Jersey	-1.351 (14.67)	0.011 (0.11)	0.136 (1.32)	2.282 (17.79)
<u>Sites without Capital</u> ($b_{no\ capt}$ in equation 4)				
Length	--	0.615 (4.27)	1.376 (9.28)	1.150 (7.66)
Narrow	--	0.723 (1.94)	1.622 (3.68)	1.869 (4.60)
Wide	--	0.582 (2.14)	0.517 (1.97)	0.608 (2.30)
Park	--	-2.641 (4.47)	-3.736 (6.64)	-3.905 (6.79)
High Rise	--	-1.169 (3.82)	-0.873 (2.91)	-1.062 (3.50)
Private	--	-1.429 (4.73)	-4.021 (17.31)	-3.603 (15.66)
Park Within	--	-0.187 (0.61)	-0.445 (1.16)	-0.432 (1.49)
Boardwalk	--	-0.157 (0.40)	-0.242 (0.65)	-0.492 (1.29)
Amusements	--	0.011 (0.04)	-0.913 (3.14)	-2.077 (6.82)
Surfing	--	-0.460 (1.86)	-2.717 (14.03)	-2.900 (14.74)
Facilities	--	1.378 (3.75)	0.972 (2.44)	1.020 (2.52)
Parking	--	-0.131 (0.33)	-1.284 (3.18)	-1.335 (3.25)
Atlantic City	--	1.272 (2.53)	1.637 (2.69)	1.320 (2.68)
New Jersey	--	-0.215 (0.94)	-2.585 (10.58)	-0.650 (2.42)

Individual Characteristics

Constant (Mean)	4.924 (70.29)	7.408 (35.62)	7.039 (58.65)	7.390 (55.79)
Constant (Deviaiton)	0.199 (1.20)	-0.059 (0.70)	-0.046 (0.53)	-0.040 (0.48)
Kids Under 10	-0.037 (1.13)	0.020 (0.70)	-0.062 (2.15)	0.027 (0.93)
Kides Between 10-16	-0.170 (5.70)	-0.152 (5.08)	-0.204 (6.72)	-0.269 (8.83)
Flexible Work Hours	-0.170 (3.71)	0.034 (0.71)	-0.034 (0.69)	-0.015 (0.31)
Part Time Work	0.126 (2.48)	0.0950 (1.75)	0.042 (0.76)	-0.104 (1.88)
Work at Home	0.895 (11.09)	0.919 (11.55)	0.828 (10.28)	0.633 (7.81)
Volunteer	-0.382 (6.04)	-0.111 (1.80)	-0.031 (0.50)	0.270 (4.25)
Student	-0.633 (12.77)	-0.493 (9.83)	-0.921 (17.53)	-0.561 (11.22)
Retired	0.422 (8.24)	0.333 (6.23)	0.472 (8.75)	0.181 (3.36)
Log Likelihood	-0.344361	-0.315677	-0.315488	-0.315524

Table 6: Instrumental Variable Regression Results

	DEFINITION	SHORT MODEL	LONG MODEL
Distance	Distance from home residence to beach in miles	-0.0027 (60.67)	0.0001 (0.47)
Log(age)	Log of age of respondent	0.1146 (66.21)	0.1072 (45.84)
Income	Household Income of respondent (1997$)	0.0008 (9.75)	0.0013 (12.32)
Vacation Home	1 if respondent owns a second home on a Mid-Atlantic Beach	0.1341(8.31)	0.1034 (7.017)
Trip Cost	See Table 2		-0.0011 (10.20)
Length	.	-	0.0409 (11.24)
Boardwalk	.	-	0.0343 (5.02)
Amusements	.	-	0.1840 (19.89)
Private	.	-	-0.0632 (9.84)
Park	.	-	0.0038 (0.29)
Wide	.	-	0.0143 (2.23)
Narrow	.	-	0.0048 (0.74)
Atlantic City	.	-	0.1804 (9.71)
Surfing	.	-	0.0127 (2.41)
High Rrise	.	-	0.0317 (4.56)
Park Within	.	-	0.0185 (2.41)
Facilities	.	-	-0.0116 (1.44)
Parking	.	-	0.0135 (1.69)
New Jersey	.	-	-0.3180 (37.45)
Kids Under 10	.	-	0.0099 (3.71)
Kids Between 10-16	.	-	-0.0052 (1.56)
Part Time	.	-	-0.0214 (3.01)
Retire	.	-	-0.0396 (6.28)
Flexible Work	.	-	0.0272 (4.71)
Student	.	-	0.0578 (6.07)
Volunteer	.	-	0.0522 (4.51)
Work at Home	See Table 2	-	-0.0269 (2.99)
R-SQUARED		0.138	0.293

Table 7: Implicit Prices for Site Characteristics

	MODEL 1: Baseline Model	MODEL 2: Exogenous Model	MODEL 3: Endogenous Model w/ Short IV	MODEL 4: Endogenous Model w/ Long IV
Site Capital				
Site Capital (Mean)	--	106.78	256.36	211.08
Site Capital (Deviation)	--	2.74	31.68	5.26
Sites with Capital				
Length	0.04	-0.77	-4.08	-11.19
Narrow	-6.76	3.97	-14.25	-10.24
Wide	-22.11	-16.96	-29.83	-24.29
Park	14.71	15.54	31.52	22.01
High Rise	-12.60	-17.36	-35.46	-33.52
Private	-17.70	-11.38	5.87	17.43
Park Within	40.98	19.98	35.85	26.45
Boardwalk	16.19	16.43	36.24	18.75
Amusements	39.44	31.08	61.48	4.58
Surfing	21.64	17.70	50.99	32.39
Facilities	-8.16	-9.02	-19.03	-8.92
Parking	10.89	6.18	18.73	8.60
Atlantic City	42.07	11.57	29.34	-22.09
New Jersey	-35.73	0.35	6.61	79.52
Sites without Capital				
Length	--	18.97	66.80	40.07
Narrow	--	22.32	78.72	65.12
Wide	--	17.97	25.09	21.18
Park	--	-81.51	-181.36	-136.07
High Rise	--	-36.08	-42.39	-37.01
Private	--	-44.10	-195.18	-125.55
Park Within	--	-5.76	-21.62	-15.03
Boardwalk	--	-4.83	-11.75	-17.13
Amusements	--	0.35	-44.33	-72.37
Surfing	--	-14.19	-131.90	-101.04
Facilities	--	42.54	47.20	35.55
Parking	--	-4.05	-62.33	-46.51
Atlantic City	--	39.25	79.44	46.00
New Jersey	--	-6.64	-125.50	-22.65

Table 8: Beach Closure Seasonal Welfare Loss Per Person (1997 Dollars)

	Baseline Model (1)	Site Capital Models (2) EXOGENOUS	(3) ENODGENOUS W/ SHORT LIST	(4) ENDOGENOUS W/ LONG LIST
Loss of Sites: Multiple Beaches				
All Delmarva: Cape Henlopen St. Park DE to Assateague Island VA	$443.81	$657.43	$1035.19	$735.56
All Delaware:	374.90	567.70	893.38	633.77
Northern Delaware Beaches: Cape Henlopen St. Park, North Shores, Henlopen Acres, Rehoboth Beach, Dewey Beach, and Indian Beach	255.72	383.53	619.58	437.92
Southern Delaware Beaches: Delaware Seashore St. Park, North Bethany Beaches, Bethany Beach, Sea Colony, Middlesex Beach, South Bethany Beach, Fenwick Island St. Park, and Fenwick Island	111.96	168.97	250.12	178.91
All New Jersey Beaches:	25.88	36.11	58.87	39.33
Loss of Sites: Most Popular Beaches				
Rehoboth, DE:	125.46	162.28	260.89	185.02
Ocean City, MD:	50.72	60.26	97.63	70.26
Cape Henlopen, DE:	55.99	84.83	148.45	104.44
Loss of Sites: Least Popular Beaches				
Ortley, NJ:	0.19	0.06	0.04	0.03
Chadwick, NJ:	0.05	0.03	0.03	0.02
Normandy, NJ:	0.03	0.04	0.04	0.03
Beach Erosion: All Beaches Reduced to Narrow				
All Delaware:	30.19	74.38	106.82	70.74
Northern Delaware Beaches:	24.97	22.56	62.53	44.77
Southern Delaware Beaches:	55.33	96.57	170.23	116.10
Site Capital:				
Site Capital (d)	--	664.25	940.35	735.84

References

Adamowicz, W. (1994). "Habit Formation and Variety Seeking in a Discrete Choice Model of Recreation Demand." Journal of Agricultural and Resource Economics **19**(1): 19-31.

Angrist, J. D. and A. B. Krueger (2001). "Intrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." Journal of Economic Perspectives **15**(4): 69-85.

Heckman, J. J. (1981). Heterogeneity and State Dependence. Studies in Labor Markets. S. Rosen. Chicago, IL, University of Chicago Press.

Imbens, G. W. and J. D. Angrist (1994). "Identification and Estimation of Local Average Treatment Effects." Econometrica **62**(2): 467-475.

Craig E. Landry and Haiyong Liu. 2007. "A Semi-Parametric Estimator for Revealed and Stated Preference Data: An Application to Recreational Beach Visitation" Manscript East Carolina University.

Lin, Pei-Chien, R. M. Adams, and R. P. Berrens (1996). "Welfare Effects of Fishery Policies: Native American Treaty Rights and Recreational Salmon Fishing". *Journal of Agricultural and Resource Economics* **21(2):** 213-276.

Loomis, J. (1995). "Four Models for Determining Environmental Quality Effects on Recreation Demand and Regional Economics". *Ecological Economics.* **12**: 55-68.

Massey, D. M. (2002). Heterogeneous Preferences in Random Utility Models of Recreation Demand. Department of Economics. Newark, DE, University of Delaware.

McConnell, K. E., I. E. Strand, and N. E. Bockstael. (2002) "Habit Formation and the Demand for Recreation: Issues and a Case Study." *Advances in Appl. Microecon.* 5(1990):217-35.

Moeltner, K. and J. Englin (2004). "Choice Behavior Under Time-Variant Quality: State Dependence Versus "Play-It-by-Ear" in Selecting Ski Resorts." Journal of Business and Economic Statistics **22**(2): 214-224.

Parsons, G. R. and D. M. Massey (2002). A Random Utility Model of Beach Recreation. The New Economics of Outdoor Recreation. N. Hanley, D. Shaw and R. E. Wright. Northampton, MA, Edward Elgar.

Parsons, G. R., D. M. Massey, et al. (2000). "Familiar and Favorite Sites in a Random Utility Model of Beach Recreation." Marine Resource Economics **14**: 299-315.

Phaneuf, D. J. and V. K. Smith (2004). Recreation Demand Models. Handbook of Environmental Economics. K. Maler and J. Vincent.

Pollak, R. A. (1970). "Habit Formation and Dynamic Demand Functions." The Journal of Political Economy **78**(4): 745-763.

Provencher, B. and R. Bishop (1997). "An Estimable Dynamic Model of Recreation Behavior with an Application to Great Lakes Angling." Journal of Environmental Economics and Management **33**(1): 107-127.

Rust, J. (1987). "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." Econometrica **55**(5).

Smith, M. D. (2005). "State Dependence and Heterogeneity in Fishing Location Choice." Journal of Environmental Economics and Management **50**(2): 319-340.

Swait, J., W. Adamowicz, et al. (2004). "Choice and temporal welfare impacts: incorporating history into discrete choice models." Journal of Environmental Economics and Management **47**(1): 94-116.