EXPLAINING PATTERNS OF CHOICE AMONG OUTDOOR RECREATION SITES IN METROPOLITAN AREAS

John F. Dwyer
Research Forester/Project Leader
USDA Forest Service
North Central Research Station
1033 University Place, Suite 360
Evanston IL 60201-3172

David B. Klenosky
Associate Professor
Purdue University
Department of Health & Kinesiology
Lambert Fieldhouse, Room 111B
West Lafayette, IN 47907-2046

Cherie L. LeBlanc
Social Scientist
USDA Forest Service
North Central Research Station
1033 University Place, Suite 360
Evanston IL 60201-3172

Abstract

This paper reports on a study of the outdoor recreation behavior of 618 (non-Hispanic) Whites, 647 African Americans, and 346 Hispanic Americans in Cook County, Illinois. Respondents were contacted for a phone survey using random digit dialing and a quota for each group. Binary logistic regression models were estimated to explain the use of 19 sites in and near the Chicago Metropolitan Area and one distant site. This paper presents the results of newly-expanded models that include 1) distance to 19 alternative destinations; and 2) demographic characteristics of origins. Suggestions are made for improving modeling and other efforts to explain the use of recreation sites in and near Metropolitan areas.

1.0 Introduction

In a previous paper (Dwyer and Klenosky 2004) we presented binary logistic regression models that explain the use of 19 sites in the Chicago Metropolitan Area and one distant site in southern Illinois (Fig. 1). These models were estimated from data on outdoor recreation behavior and individual characteristics of 618 (non-Hispanic) Whites, 647 African Americans, and 347 Hispanic Americans in Cook County, IL interviewed via telephone survey. The initial models helped explain the implications of demographic variables for participation at each of the 20 recreation sites. Explanatory variables in those models include estimated distance from the individual's home to the site, race/ethnicity (three categories), age (five categories), area of residence (five areas in Cook County), education (seven categories), annual household income (six categories), household size, and gender. Travel distance and section of the county where respondents lived were calculated from the residential zip codes that respondents provided. In this paper we extend our analysis beyond the characteristics of individuals and the distances they live from a particular site to include proximity to other sites and demographic characteristics of the zip code where they live. Among the characteristics of zip code origins that will be evaluated are: 1) racial/ethnic composition: (2) income: and (3) population density. With the other recreation sites included in the analysis, we will estimate how distance from the individual's residence to the other sites, as well as characteristics of origins, contribute to site choice models for each site. The results will help outdoor recreation site planners, managers, and researchers better understand patterns of use for sites in metropolitan areas and simultaneously appreciate the methodological and analytical challenges involved in such an endeavor.

2.0 The Impact of Distance to Alternative Sites

The decision to visit a site for recreation can be viewed as a choice among a set of alternative sites that the decision maker considers. These alternatives may vary in terms of their locations and the attributes that may attract or discourage visitation by a given individual. Information on the attributes and locations of sites that are seen as substitutes for or complements to a particular site could improve site choice models.

The 20 diverse sites in our study are difficult to characterize in a simple or straightforward way. Many of the sites are large and complex, offering a wide range of experiences, and we do not have a systematic set of attributes for all of them. We do know the location (zip code) of each site, and we calculated the straight line distances from the center of the zip code where each respondent lives to the center of the zip code where the site is located. Theoretically we could simultaneously enter distance to each of the sites into the site choice model and see which ones make a significant

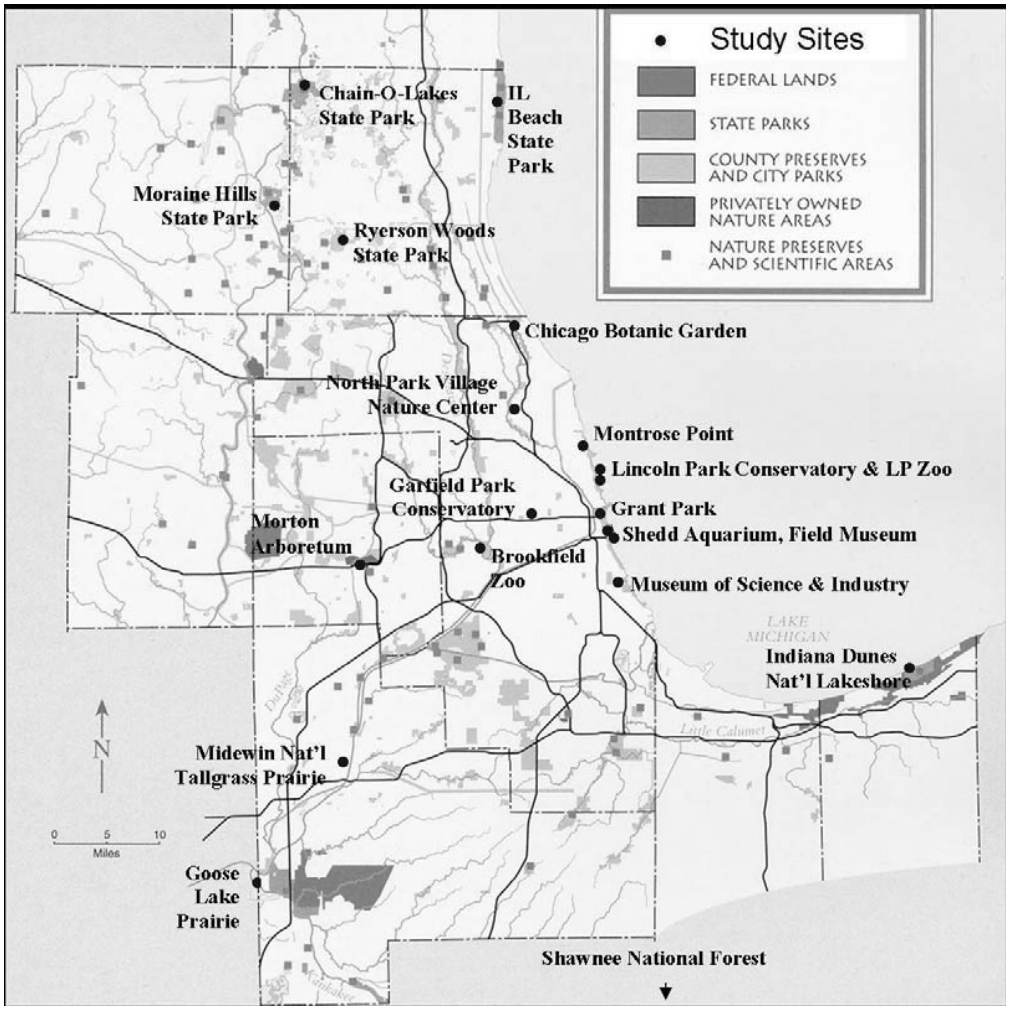

Figure 1.—Map of Chicago area study sites.

contribution to the model for predicting visitation to each site. As with our previous models, we expect that the coefficient for distance from an origin to the site being evaluated would be negative – the farther the individual is from the site, the less likely he or she is to visit it. If another site in the choice set was a substitute for the site being evaluated, we would expect the coefficient for distance to the substitute to be positive, indicating that the farther the individual is from the substitute site, the more likely it is that he or she would visit the site being evaluated. Alternatively, a site may be a complement to another site, in which case if a person visits one, he or she is likely to visit the other. With a complementary site, we expect a negative sign for distance to that site – the farther that the complementary site is from the individual, the less likely it is that he or she would use the site being evaluated. Given the diversity of the sites, relative crudeness of the measures of distance to sites, and clustering of sites in the Chicago area, it is difficult to anticipate which sites might be substitutes and which might be complements for a given site.

Initial analyses indicated numerous significant correlations among both the distances to the site being evaluated and among the distances to alternative sites. Correlations among distances to the various sites were particularly troublesome because of the concentration of all origins and many of the sites in Cook County, and the geographic clustering of many of the sites. Consequently, we adopted a simple approach where we tested the effect of each individual site as a substitute or a complement by entering the distance to that site into the model along with distance to the site being evaluated and all other variables in the basic site-choice model (respondent's

race/ethnicity, age, education, income, household size, and gender). Variables representing the section of Cook County where the respondent resided were dropped from the model because they were correlated with distance to some of the sites.

The results of our efforts to include single sites in the model one at a time were highly variable (Table 1). Models for 12 of the sites had no alternative sites with a significant coefficient for distance (i.e., suggesting no complement or substitute sites). This may be due, in part, to the inter-correlations among distances to the sites in our data set. Another dimension of the difficulty with this analysis is that distance from an origin (zip code) to two sites largely defines the section of Cook County where the individual lives. That section may have its own unique set of physical, biological, and social attributes which themselves might influence choice of recreation sites. It can be difficult to sort out the influence of these place-specific variables from the distance variables.

When distances to alternative sites in a model are correlated, it is not possible to sort out the influence of alternative sites on site choice. With our data, correlations among distances to sites are very strong for sites in relatively close proximity (e.g., 0.956** between distances to Grant Park and Lincoln Park Zoo) and moderate for sites distant from each other (e.g., 0.712** between Chicago Botanic Garden and Lincoln Park Zoo). This suggests that only distant pairs of sites could be tested for substitutability/complimentarity by entering them into the model.

Correlations between distance to sites and demographic variables are common as well, often a result of the unique distributions of demographic variables across Cook County. As can be seen in the data from the 2000 U. S. Census, the African American population is concentrated immediately west and south of downtown Chicago. The Asian population in concentrated in the north. The Hispanic population is found mostly to the west and southwest of downtown. The White population is concentrated to the north and the west. Median household income is highest just north of downtown Chicago and in the northern and western parts of the county. Population density is highest within Chicago and particularly to the north of the downtown area

Table 1

Beta Coefficients for Models Estimated to Predict Number of Visits from Each of 140 Zip Code to the 20 Recreation Sites (Using Aggregated Dataset Weighted by Number of Survey Respondents in the Zip Code)

Table 1.1

Linear Regression Models without Substitute Sites

	Grant Park	Museum Science Industry	Field Museum	Shedd Aquarium	Brookfield Zoo	Lincoln Park Zoo	North Park Village	Lincoln Park Cons.	Garfield Park Cons.	Montrose Point	Morton Arb.	Chicago Botanic Garden	Indiana Dunes	Illinois Beach SP	Chain-O-Lakes SP	Ryerson Woods SP	Moraine Hills SP	Shawnee Nat'l Forest	Goose Lake Prairie	Midewin Prairie
Total Population	0.646**	0.659**	0.669**	0.640**	0.703**	0.603**	0.154**	0.428**	0.474**	0.336**	0.364**	0.613**	0.569**	0.580**	0.496**	0.378**	0.429**	0.362**	0.311**	0.308**
Percent Black	0.256**	0.220**	0.254**	0.316**	0.074**	0.143**	- 0.263**	0.098**	0.555**	- 0.095**	- 0.278**	0.010	- 0.169**	- 0.048	- 0.516**	0.102**	- 0.022	- 0.299**	0.047	0.250**
Percent Hispanic	- 0.186**	- 0.110**	- 0.113**	0.079*	0.022	0.054	- 0.210**	0.066*	0.075*	0.026	- 0.266**	- 0.254**	- 0.287**	- 0.204**	- 0.410**	- 0.073	- 0.220**	- 0.252**	0.081*	0.149**
Per Capita Income	0.075**	0.115**	0.094**	0.149**	- 0.065*	0.100**	- 0.314**	0.092**	0.088**	0.073*	0.102**	0.073*	0.001	0.016	- 0.210**	0.111**	0.008	- 0.116**	- 0.044	0.019
Population per sq miile	0.094**	- 0.091**	- 0.023	- 0.092**	- 0.238**	0.215**	0.373**	0.295**	- 0.039	0.349**	0.042	0.066*	- 0.123**	- 0.056	- 0.121**	- 0.236**	- 0.214**	0.102**	0.049	- 0.061*
Distance to Site	- 0.025	- 0.232**	- 0.071**	- 0.067*	- 0.096**	0.056*	- 0.153**	- 0.021	- 0.101**	- 0.052	- 0.140**	- 0.175**	- 0.490**	0.058*	- 0.218**	- 0.079**	- 0.160**	- 0.408**	- 0.120**	0.009
R Square	0.481	0.527	0.469	0.490	0.463	0.537	0.322	0.461	0.577	0.416	0.150	0.334	0.383	0.194	0.304	0.086	0.109	0.177	0.163	0.196

Table 1.2

Linear Regression Models with Substitute Sites

	Museums		Zoos		Conservatories		Arboreta	
	Museum Science Industry	Field Museum	Brookfield Zoo	Lincoln Park Zoo	Lincoln Park Cons.	Garfield Park Cons.	Morton Arb.	Chicago Botanic Garden
Total Population	0.647**	0.647**	0.710**	0.563**	0.391**	0.481**	0.397**	0.586**
Percent Black	0.175**	0.096**	0.078**	0.131**	0.134**	0.549**	- 0.365**	- 0.026
Percent Hispanic	- 0.065*	- 0.080*	0.029	0.065*	0.118**	- 0.070*	- 0.275**	- 0.242**
Per Capita Income	0.151**	0.149**	- 0.047	0.035	- 0.001	0.106**	0.179**	0.023
Population per sq miile	0.026	0.105**	- 0.187**	0.087**	0.112**	0.004	0.196**	0.017
Distance to Site	- 0.452**	0.349**	- 0.141**	- 0.213**	- 0.917**	- 0.238**	- 0.247**	- 0.258**
Distance to Sub Site	0.293**	- 0.474**	0.089**	0.240**	0.798**	0.171*	0.331**	0.181**
R Square	0.540	0.509	0.464	0.558	0.492	0.578	0.205	0.357

* coefficient significant at 0.05 level, ** coefficient significant at 0.01 level

and along Lake Michigan. Significant correlations among these variables include percent African American with a negative correlation with median household income (-0.431**) and percent Hispanic American (-0.217**). Percent Hispanic American has a negative correlation with median household income (-0.337**), and a positive correlation with population per square mile (+0.298**). Population per square mile has a negative correlation with median household income (-0.406**).

When looking at correlations between distance to individual study sites and these same demographic variables, similar findings emerge. Some examples of these correlations include: Population per square mile has a strong negative correlation with distance from Montrose Point (-0.801**). Median income per household has a strong positive correlation with distance from the Museum of Science and Industry (+0.670**). Percent Asian has a moderately strong positive correlation with distance from sites in the west (Goose Lake Prairie (+0.454**) and Midewin National Tallgrass Prairie (+0.503**), and a moderate negative correlation with distance from sites in the north (Illinois Beach State Park (-0.472**), Ryerson Woods (-0.453**), Chain-O-Lakes State Park (-0.450**), and Moraine Hills State Park (-0.439**)). Percent Hispanic has a moderate negative correlation with distance from downtown sites such as Lincoln Park Zoo (-0.471) and Grant Park (-0.427**), and distance to sites on the near west side such as Garfield Park Conservatory (-0.512). Percent African American is positively correlated with distance from sites in the north such as Moraine Hills State Park (+0.535**) and Chain-O-Lakes State Park (+0.533**), and negatively correlated with sites to the south such as the Museum of Science and Industry (-0.559**) and Indiana Dunes National Lakeshore (-0.591**). As we discuss later, these correlations, which are an inherent characteristic of metropolitan areas, complicate efforts to model site visitation patterns in such areas.

Despite these differences due to intercorrelations, the results of the modeling effort do shed some light on possible substitution among some of the sites. The models for Lincoln Park Zoo and Lincoln Park Conservatory suggest that Garfield Park Conservatory, Morton Arboretum, Brookfield Zoo, and Goose Lake Prairie are substitutes for these two sites. Lincoln Park Zoo and Lincoln Park Conservatory are adjacent and visitors can easily visit both. Garfield Park Conservatory and Morton Arboretum both offer special displays of flowers and other plants, while Brookfield Zoo presents animals in a well-landscaped context. It is not difficult to see these sites as substitutes for Lincoln Park Zoo and Lincoln Park Conservatory in providing certain types of experiences. Goose Lake Prairie, the distance to which is also significant in the models for Lincoln Park Zoo and Lincoln Park Conservatory, offers a distinctly different wildland setting west of the Chicago area. It would be difficult to see it as a substitute for the two Lincoln Park sites because it is far away. The positive coefficient for distance to Goose Lake Prairie suggests that people in the Chicago area are more likely to visit the two Lincoln Park sites. This can be seen as reinforcing the negative coefficient for distance to the site. People who live a long way from Goose Lake Prairie also live closer to Lincoln Park Zoo and Lincoln Park Conservatory and are likely to visit those sites. However, with both intercorrelated distance variables in the model, distance to the site and to a possible substitute, it is impossible to interpret the two coefficients. And with the critical significance that a travel cost model places on the coefficient for distance to the site, use of such a model in a travel cost analysis could be misleading (McCullom et al. 1990).

In the model explaining use of the Chicago Botanic Garden, both Morton Arboretum and Brookfield Zoo emerge as substitute sites. This is not surprising in that Morton Arboretum and Brookfield Zoo are both places where people can go to see a wide range of unusual plants in attractive displays. Brookfield Zoo focuses on animals, but the animals are found in an attractive landscape. All three sites charge entry fees, offer educational programs, and are situated in an extended natural environment that includes nearby forest preserves.

In the model explaining use of Grant Park, there were a number of sites that emerged as possible substitutes, including North Park Village Nature Center, Garfield Park Conservatory, Morton Arboretum, Brookfield Zoo, Moraine Hills State Park, Chain-O-Lakes State Park, and Ryerson Woods. Grant Park is a highly-developed park with a spectacular fountain that is located on the Chicago Lakefront not far from the center of the city. It seems unlikely that any of the above sites would be considered substitutes by users, and what may be at work is that distances to these sites, all of which are peripheral to the City Center, are a proxy for living close to Grant Park and thereby being likely to visit.

Other site choice models where alternative sites emerged as substitutes and complements are Brookfield Zoo, Illinois Dunes National Lakeshore, Moraine Hills State Park, and Chain-O-Lakes State Park. With all of these sites, there were a number of alternative sites where the coefficients for distance are significant and positive, suggesting substitutes, or significant and negative, suggesting complements. Shawnee National Forest was a significant complement for all four of the sites. This makes sense in that a trip to the Shawnee in far southern Illinois would not be viewed as a substitute for a trip to a local site. Midewin National Tallgrass Prairie emerged as a significant complement for three sites (Brookfield Zoo, Illinois Dunes National Lakeshore, and Chain-O-Lakes State Park). This may be attributable to the unique character of Midewin, in that it is a large site that is fenced and closed to the public unless they are invited for special events. This is unlike any of the other sites. Moraine Hills State Park and Goose Lake Prairie also emerged as complements for Chain-O-Lake State Park (along with Midewin). Moraine Hills State Park is in the same general vicinity as Chain-O-Lakes State Park; and Goose Lake Prairie, while distant, offers somewhat similar experiences. With all four sites there are a number of sites where the distances to them are significant but many of them are clustered. There are similar coefficients for distance to sites that are clustered together.

3.0 The Impact of Origin Characteristics

The characteristics of the areas where people live can influence site choices by affecting an individual's desire to visit places. Important local influences may include the kinds of recreation opportunities available nearby, the accessibility of those opportunities to the individual, and the characteristics of the people who live in the area. In our analysis we used the 140 Cook County zip codes that were reported by respondents in our original phone survey as our basis for the study of origins. Using data about each of these zip codes from the 2000 U.S. Census, we selected the following variables to include in the analysis: population per square mile, mean per capita income, percent African American, and percent Hispanic American. We reasoned that population density may be an indicator of open space in the area and the other variables would describe the population living there. These local-area variables were added on to each observation based on the zip code the survey respondent reported. Thus all of the observations from a given zip code had the same values for these origin characteristics. The basic site choice model to which we added these characteristics included the race/ethnicity and income of individual respondents.

We then re-estimated all site choice models using the basic information that was in the original model (Dwyer and Klenosky 2004) plus the origin-based variables outlined above. These variables did not add significantly to the explanatory power of the models. In fact, the coefficients for the variables were significantly different from zero in only a few instances, largely reflecting the influence of ethnic populations nearby. The overall lack of significant coefficients may reflect, in part, correlations between the characteristics of origins, and (1) the characteristics of individuals sampled from those origins, and (2) the distance to the site being evaluated.

Travel cost models often use the characteristics of origins, usually counties, and their distances from sites to estimate the number of trips to sites. There is evidence from other studies that the characteristics of areas where people live influence their recreation preferences and behavior. Edwards (1981), for example, reported that African Americans living in predominately White residential settings, as opposed to non-white or mixed communities, had higher expectations of leisure needs. To test that finding, we divided our sample into areas (zip codes) where different racial/ethnic groups predominated (i.e., 66% or more of the population). Each of the 140 zip codes in the study area was classified as predominately White (79 zip codes), predominately African American (22), predominately Hispanic American (6), or mixed (33). The zip codes in each category were aggregated, and for those groups of zip codes, the percent of each racial/ethnic group that visited each site was calculated. The results were mixed, and interpretation was hampered by small sample sizes, such as the number of African Americans living in Hispanic neighborhoods that used a particular site.

African Americans who lived in predominately White zip codes had a higher rate of participation at 10 of the sites than African Americans living in predominately African American zip codes. However, with most of these sites, the predominately White neighborhoods were closer to these sites than were the predominately African American neighborhoods. Thus it is not clear to what extent the pattern of participation by neighborhood is due to ethnic composition or location with respect to a site. Similarly

Whites who lived in predominately African American neighborhoods were more likely to visit 11 sites than Whites who live in predominately White neighborhoods, at least partly because they lived in neighborhoods that were closer to the sites. Much, but not all, of the difference in participation across ethnic neighborhoods appears to be attributable to proximity to a site. Analysis of the site participation patterns by additional racial/ethnic groups across neighborhoods of different racial/ethnic compositions did not yield clear patterns, although this may be due, in part, to limited sample sizes by neighborhood and racial/ethnic group.

4.0 Discussion

Based on past successes in estimating simple travel cost models and site choice models for urban sites, we set out to build more complex, and hopefully more useful, models to explain urban site choices. We were particularly interested in building the availability of alternative sites and the characteristics of origins into the model. Our analysis was limited, in part, by our data, which came from a previous study. We had fairly crude data for locations (zip codes) of the alternative sites and residences of respondents, and no systematic information on site attributes. Travel distances between residences and sites were approximated by straight line distances between the centers of zip codes. Information at the origin level was limited to simple demographic characteristics, the values for which were often clustered in particular parts of Cook County. The 20 study sites are diverse, and frequently several were clustered near each other. Many of the sites and all of the origins were in Cook County, which resulted in numerous correlations among travel distances to sites. It was not clear what relationship to expect between participation at various sites since we did not have prior information about how individuals perceived sites or their interrelationships.

Distances to each of the 19 other sites were not particularly helpful in increasing the explanatory power of our site choice models. The analysis was complicated by correlations among the distance variables (including to the site being evaluated and possible complements and substitutes), as well as correlations with demographic variables. These correlations were significant because demographic characteristics tended to cluster in the landscape, resulting in significant correlations with distances to a number of the sites.

What did emerge as significant in the analysis was substitutability among some of the more similar sites such as Lincoln Park Zoo, Lincoln Park Conservatory, Garfield Park Conservatory, Chicago Botanic Garden, and Brookfield Zoo. These sites have relatively specific missions and provide distinctive outdoor environments that have some commonalities. This is in relatively sharp contrast to some of the other sites in the analysis which include city parks, county forest preserves, or state parks that provide a wide range of experiences and opportunities, some of which are unique.

In this study we approached site choice from a gross overall view looking for relationships between sites, and relatively few emerged. Would another approach be better? Are complements and substitutes the most useful framework for looking at the interrelationships among sites? Should the analysis be extended beyond sites used in a single year to those used over a multi-year period? Our data suggests that people experience a range of sites over time (Dwyer and Barro 2001), and some sites may enter an individual's life at various stages – such as taking children to a zoo, and then at a much later time taking grandchildren to that same zoo. Do earlier experiences at a site affect how people currently perceive natural resources, or are these perceptions based mostly on sites that are visited at present? Should we limit our analysis to groups of similar sites such as state parks, arboreta and botanic gardens, etc?

5.0 Summary and Conclusions

Our efforts to build more complex models of urban site choice were less than fully successful. We were able to identify complementary and substitute sites for some of the 20 study sites, but this resulted in limited improvements in our site choice models. Part of the problem could be solved with additional information – for example, by including more respondents in a study and gathering more information about sites and neighborhoods across a wider geographic area. But it is also clear that the complexity of urban neighborhoods and sites is a significant factor that is likely to become even more challenging over time. It may be useful to rethink the precise questions that need to be addressed concerning urban natural resource management and use before proceeding with additional research on modeling site choices in urban systems. Is it a matter of identifying the sites that substitute for or complement an individual

site? Or is it more a matter of identifying the clusters of sites that people visit over a period of time? And examine whether and how these clusters change over time? Would we better understand how management options influence urban site choice behavior if we presented possible management options such as changes in the landscape, educational programs, or number and character to sites to individuals and asked how they would respond? Or if we perhaps asked them how and why their site choice behavior has changed over time? Finally, whereas this research has focused primarily on the utility of incorporating information about the characteristics of individuals and where they live to examine site visitation behavior, future modeling efforts might include additional information about the characteristics of the sites themselves (i.e., observable/objective site characteristics and/or unobservable/perceived site characteristics). Prior research by Lin et al. (1988) conducted with site visitors in an urban setting suggested that accounting for the impact of such site characteristics may lead to a more complete and (albeit arguably) more useful site choice models.

6.0 Acknowledgment

This paper is based on research funded, in part, under a research joint venture agreement between the USDA Forest Service North Central Research Station and Purdue University.

7.0 Citations

Dwyer, J. F., & Barro, S. C. (2001). Linkages in the use of recreation environments across the urban to ex-urban spectrum by urban residents. In: Proceedings of the 2001 Northeastern Recreation Research Symposium. General Technical Report NE-289 (pp. 202-207). Newtown Square, PA: USDA Forest Service Northeastern Research Station.

Dwyer, J. F., & Klenosky, D. B. (in press). (2004). The implications of demographic change in metropolitan areas for the use of recreation sites. In: Proceedings of the 2003 Northeastern Research Symposium. General Technical Report NE-317. Newtown Square, PA: USDA Forest Service Northeastern Research Station.

Edwards, P. K. (1981). Race, residence and leisure style: Some policy implications. Journal of Leisure Science, 4(2), 95-111.

Lin, Y., Peterson, G. L., & Rogerson, P. A. (1988). A nested urban recreation site choice model. Journal of Leisure Science, 10(1), 1-15.

McCollum, D. W., Peterson, G. L., Arnold, J. R., Markstrom, D. L., & Hellerstein, D. M. (1990). The net economic value of recreation on the National Forests: Twelve types of primary activity trips across nine Forest Service Regions. Research Paper RM-289. Fort Collins, CO: USDA Forest Service.