Appalachian
STATE UNIVERSITY

Department of Economics Working Paper

Number 14-09 | September 2014

Title: Predictive Validity of Stated Preference Data: Evidence from Mountain Bike Park Visits Before and After Trail System Expansion

Kevin Atkinson
Appalachian State University

John C. Whitehead
Appalachian State University

Department of Economics
Appalachian State University
Boone, NC 28608
Phone: (828) 262-2148
Fax: (828) 262-6105
www.business.appstate.edu/economics

Predictive Validity of Stated Preference Data: Evidence from Mountain Bike Park Visits Before and After Trail System Expansion[1]

Kevin Atkinson and John C. Whitehead

Department of Economics

Appalachian State University

September 25, 2014

Abstract. This paper investigates the validity of stated preference data for use in recreation demand estimation. We use stated preference and revealed preference data from users of a mountain bike park collected before and after an expansion of the trail system. The ex-ante stated preference data elicited before the change exhibits hypothetical bias, but, it would provide useful information for demand prediction.

JEL: Q26

Keywords: hypothetical bias, recreation demand, revealed preference, stated preference

[1] A previous version of this paper was presented at the National Conference on Undergraduate Research, April 3-5, Lexington, KY. Funding for this study was provided by a Walker College of Business Dean's Club Research Grant.

Introduction

Economists estimate the nonmarket value of public goods with revealed preference and stated preference data (Whitehead et al. 2008). Revealed preference data is preferred because it may better reflect actual behavior. But, it may fail to provide the information desired about future behavior under different scenarios because the scope of revealed preference data is limited to observed historical variation. Stated preference data gathered with hypothetical survey questions can be used to forecast beyond historical variation but has uncertain accuracy because respondents may err in predicting their own future behavior or future economic conditions may differ from current conditions. The accuracy of stated preference data in predicting the future is the subject of ongoing controversy in the economics literature (Hausman 2012, Haab et al. 2013).

A number of studies have investigated the predictive validity of stated preference behavior data. Loomis (1993) compares the length of hypothetical lake visits to observed behavior and finds that the two measures are not different. Grijalva et al. (2002) compares pre-policy hypothetical behavior to behavior observed after the implementation of a policy change,using two rounds of surveys given to rock climbers who visit a state park in Texas. They find that climbers do not overestimate their behavior to a policy change. Whitehead (2005) and Whitehead, Noonan and Marquardt (forthcoming) find that survey respondents overstate their future behavior. But, using jointly estimated ex-ante revealed and stated preference data models, a hypothetical bias correction yields statistically equivalent predictions to the ex-post actual behavior. Whitehead, Groothuis and Weddell (2014) find that stated preference data accurately predicts actual behavior after adjusting for respondent certainty. The results of these studies

suggest that stated preference data be used to make better predictions about the future.

This paper aims to evaluate the predictive validity of stated preference data obtained through surveys of visitors at the Rocky Knob Mountain Bike Park in Boone, North Carolina. Rocky Knob Park opened in May 2011 with 2.6 miles of trail and expanded to almost 8 miles over the next two years. We use a panel survey to collect revealed and stated preference data. By comparing the ex-ante revealed preference and stated preference data before trail expansion with the ex-post revealed preference data after trail expansion, we evaluate the predictive validity of the both types of data.

Methods

Our data comes from two internet surveys of mountain bikers. Participants in the ex-ante survey either completed the survey from a link on the Rocky Knob Park website (http://rockyknob.wordpress.com/) or were recruited on-site during mountain biking season. Survey responses were collected over a two year period beginning in October 2011 and ending in October 2013. A convenience sample of 485 responses were obtained with 38% recruited on-site. Three hundred and two respondents provided enough information for analysis. The ex-ante survey covered past biking behavior and anticipated future use of Rocky Knob at different stages of completion.

Rocky Knob Park was completed in 2012 with 7.8 miles of trail. The ex-post survey was emailed in November 2013 to 145 ex-ante survey respondents who had indicated a willingness to participate in a follow-up survey. The response rate was 70% with 101 responses. We asked respondents for the number of mountain bike trips they had taken to Rocky Knob during the past

12 months. We eliminated 33 respondents who did not answer the hypothetical question, provided unusable answers due to faulty recall and those who took the ex-post survey within the 12 month period between surveys. For the empirical analysis, we further restrict the sample by deleting those respondents who did not take trips in the ex-ante (n=3) and ex-post (n=10) surveys. We speculate that those who took zero trips ex-post have left the market for reasons that they could not predict ex-ante (e.g., injury or relocation).

Each respondent answered questions about five recreation behavior scenarios: two revealed preference scenarios and three stated preference scenarios of recreation behavior. The first three scenarios are from the ex-ante survey. Scenario one is revealed preference trips in the past 12 months from the ex-ante survey when there were 3.8 miles of trail. Scenario two is stated preference trips for the next 12 months with an expected 6 miles of trail. Scenario three is stated preference trips for the next 12 months with an expected 8 miles of trail. Scenario four in the ex-post survey is revealed preference trips over the past 12 months with the 7.8 completed miles of trail. Scenario five is stated preference trips over the next 12 months with the 7.8 completed miles of trail. We use trips from scenarios one, two and four to investigate predictive validity.

Results

The mean ex-ante revealed preference trips with 3.8 miles of trail is 16 (Table 1). The mean ex-ante stated preference trips with 8 miles of trail is 33. The mean ex-post revealed preference trips with 7.8 miles of trail is 19. The revealed preference trips from the ex-ante survey and the revealed preference trips from the ex-post survey have a Pearson correlation of .645. In a paired t-test the difference between their means is -2.22 (standard error of 1.04). The

stated preference data with 8 miles of trail from the ex-ante survey and the revealed preference data from the ex-post survey with 7.8 miles has a Pearson correlation of .635. The difference between their means is 15.03 (standard error of 1.43).

We estimate linear and double-log ordinary least squares regression models where revealed preference trips from the ex-post survey is a function of revealed and stated preference trips from the ex-ante survey. The models are $x_2 = \alpha + \beta x_1$ and $lnx_2 = \alpha + \beta lnx_1$; where x is trips, subscript 1 indicates the ex-ante survey and subscript 2 indicates the ex-post survey. The null hypothesis of one to one correspondence would imply $\alpha = 0$ and $\beta = 1$ in both functional forms. The alternative hypothesis of hypothetical bias where ex-ante stated preference trips are greater than revealed preference trips, would imply $\alpha = 0$ and $\beta < 1$. When $\alpha = 0$ in the linear model the inverse of the slope measures the ratio of trip overstatement: $x_2/x_1 = 1/\beta$. When $\alpha = 0$ in the double-log model the inverse of the slope measures trip overstatement as an elasticity: $ln\, x_2/lnx_1 = 1/\beta$. If $\alpha \neq 0$ in both models there is a bias in the prediction of ex-post trips that is uncorrelated with ex-ante trips.

All of the models have a significant amount of explained variation in the dependent variable (Table 2). Considering first the linear regression of ex-post revealed preference trips on ex-ante revealed preference trips with different underlying trail miles (3.8 miles vs. 7.8 miles), the model has a statistically significant constant. Counterfactually, for respondents with zero ex-ante trips the model predicts seven trips after trail expansion. The coefficient on ex-anted revealed preference trips has a 95% confidence interval of [0.55, 0.99]. This model indicates that using the ex-ante revealed preference data to predict trips would under predict trips for less avid users (those with less than 30 trips) and over predict trips for more avid users. Considering next

the linear regression of ex-post revealed preference trips on ex-ante stated preference trips with about the same underlying trail miles (8 miles vs. 7.8 miles), the model has a statistically insignificant constant. The coefficient on ex-anted stated preference trips has a 95% confidence interval of [0.36, 0.65] and implies trip overstatement of 1.98 over the entire range of ex-ante trips.

The double-log models have improved fit with about 50% more variation explained. Considering the double-log regression of ex-post revealed preference trips on ex-ante revealed preference trips, the model has a statistically significant constant. The coefficient on ex-ante revealed preference trips has a 95% confidence interval of [0.65, 0.95]. Considering next the double-log regression of ex-post revealed preference trips on ex-ante stated preference trips, the model has a statistically insignificant constant. The coefficient on ex-anted stated preference trips has a 95% confidence interval of [0.65, 0.98] and implies trip overstatement of 23% over the entire range of ex-ante trips.

One caveat to these results is that the ex-post survey was conducted after an unusually rainy mountain biking season. The average monthly rainfall for the sample was five inches for the ex-ante stated preference trip question and seven inches for the ex-ante revealed preference trip question (a 34% increase). This is a potentially significant intervening variable that we are not able to adequately take into account given the small sample size. When we conduct the double-log regression test with ex-post revealed preference trips adjusted for rainfall the slope is not significantly different from one. However, this is an ad-hoc adjustment to the trips and should only be considered speculative.

Conclusions

In our models that attempt to predict ex-post revealed preference trips with data from an ex-ante survey we find that both revealed and stated preference data has limitations. The revealed preference models have a constant that is not significantly different from zero (i.e., a demand shift), indicating a downward bias of ex-post trips that is uncorrelated with ex-ante trips. The stated preference models have slopes that indicate an overstatement of future trips. We find a hypothetical bias ratio of about 23% in our preferred double-log model. Similar hypothetical bias results are found by Whitehead, Noonan and Marquardt (forthcoming). Additional studies such as these could be used to support a meta-analysis of hypothetical bias in stated preference behavior studies and results could be used to calibrate stated preference studies that do not have the benefit of ex-post surveys.

Table 1. Mountain Bike Park Trips				
Variable	Mean	Std Dev	Minimum	Maximum
Ex-Ante Revealed Preference Trips	15.64	22.00	1	90
Ex-Ante Stated Preference Trips	32.65	34.87	1	150
Ex-Post Revealed Preference Trips	19.07	25.99	1	100
Cases	55			

Table 2. Determinants of Ex-Post Revealed Preference Trips								
	Linear				Double Log			
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Constant	7.43	3.38	2.54	3.57	0.65	0.18	-0.09	0.27
Ex-Ante Revealed Preference Trips	0.75	0.13			0.80	0.08		
Ex-Ante Stated Preference Trips			0.51	0.08			0.81	0.08
R2	0.40		0.40		0.67		0.64	
F-Value	47.14		44.58		109.21		92.6	
Cases	55		55		55		55	

References

Grijalva, Therese C., Robert P. Berrens, Alok K. Bohara, Paul M. Jakus, and W. Douglass Shaw. "Valuing the loss of rock climbing access in wilderness areas: a national-level, random-utility model." Land Economics 78, no. 1 (2002): 103-120.

Haab, Timothy C., Matthew G. Interis, Daniel R. Petrolia, and John C. Whitehead. "From Hopeless to Curious? Thoughts on Hausman's "Dubious to Hopeless" Critique of Contingent Valuation." Applied Economic Perspectives and Policy 35, no. 4 (2013): 593-612.

Hausman, Jerry. "Contingent valuation: from dubious to hopeless." Journal of Economic Perspectives 26, no. 4 (2012): 43-56.

Loomis, John B. "An investigation into the reliability of intended visitation behavior." Environmental and Resource Economics 3, no. 2 (1993): 183-191.

Whitehead, John C., Melissa Weddell and Peter A. Groothuis, Mitigating Hypothetical Bias in Stated Preference Data: Evidence from Sports Tourism, Appalachian State University Department of Economics Working Paper Number 14-06, November 2014.

Whitehead, John C. "Environmental Risk and Averting Behavior: Predictive Validity of Jointly Estimated, Revealed and Stated Behavior Data." Environmental and Resource Economics, 32, 2005. 301-16.

Whitehead, John C., Douglas Noonan and Elizabeth Marquardt. “Criterion and Predictive Validity of Revealed and Stated Preference Data: The Case of Music Concert Demand.” Economics and Business Letters, forthcoming.