Using a Control Function to Resolve the Travel Cost Endogeneity Problem in Recreation Demand Models

Richard T. Melstrom*
Department of Economics and Finance
Department of Environmental Studies
Salisbury University

Frank Lupi
Department of Agricultural, Food and Resource Economics
Department of Fisheries and Wildlife
Michigan State University

Selected Poster prepared for presentation at the Agricultural & Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington, DC, August 4-6, 2013

*Corresponding author. Contact: rtmelstrom@salisbury.edu

Using a Control Function to Resolve the Travel Cost Endogeneity Problem in Recreation Demand Models

Max Melstrom[1], Frank Lupi[2]

1 Department of Economics and Finance, Department of Environmental Studies, Salisbury University

2 Department of Agricultural Food and Resource Economics, Department of Fisheries and Wildlife, Michigan State University

Background

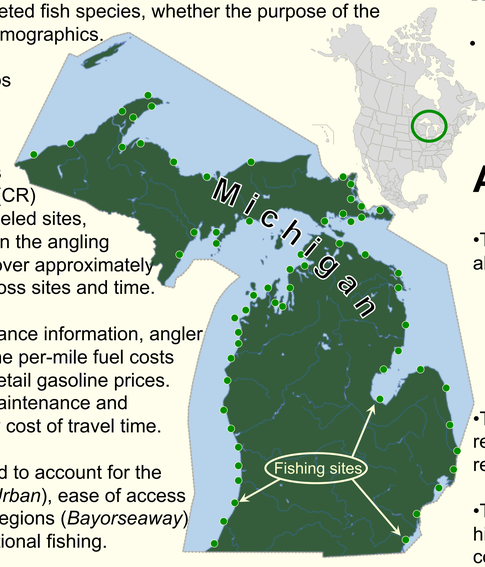
- Endogeneity arises when the observed, explanatory variables are correlated with omitted variables.
- Standard two-stage least-squares (2SLS) approaches to correct for endogenous variables cannot be used in nonlinear models. So random utility maximization (RUM) models, which include many recreation demand models, require alterantive ways to correct for endogeneity.
- In recreation demand cost endogeneity is a particularly serious issue because biased estimates of the travel cost effect will lead to invalid measures of site value (Parsons 1991).

Objectives

- Test and correct for travel cost endogeneity in a RUM model of recreation demand.
- Use a novel method to correct for endogeneity in recreation demand models: a control function.
- Compare the results of the control function approach with the results from using alterantive-specific constants (ASCs), which has been used in prior research to correct for endogeneity in recreation demand models (Murdock 2006).

Application: The Demand for Great Lakes Recreational Fishing

- Our application uses data on Michigan Great Lakes recreational anglers collected from a 2008-2009 survey. Information from the survey included the location and date of anglers' most recent trips, targeted fish species, whether the purpose of the trip was primarily to fish and angler demographics.
- The destinations from the reported trips were then matched to the fishing sites creeled by the Michigan DNR.
- Michigan DNR monthly creel data was used to construct per-hour catch rate (CR) estimates for six fish species at 66 creeled sites, to describe the quality of fishing sites in the angling demand model. Trips in were spread over approximately two years, so observed CR varied across sites and time.
- Travel costs were calculated from distance information, angler characteristics and gasoline prices. The per-mile fuel costs were derived from Michigan monthly retail gasoline prices. Travel cost is the sum of fuel costs, maintenance and depreciation costs and the opportunity cost of travel time.
- Several dummy variables were created to account for the effects of the availability of services (*Urban*), ease of access (*Highway*) and fish species in limited regions (*Bayorseaway*) in the demand for Great Lakes recreational fishing.



Site Choice Model

- The RUM model follows a conditional logit framework, with the utility of visiting a site is:

$$U_{ikt}=\rho(y_i-p_{ikt})+\boldsymbol{\beta_1 q_k}+\boldsymbol{\beta_2 q_{kt}}+\varepsilon_{ikt}$$

where y_i is income of angler i, p_{ikt} travel cost to site k at time t, $\boldsymbol{q_k}$ a vector of quality measures that vary across sites and $\boldsymbol{q_{kt}}$ a vector of quality measures that vary across sites and time.

- Travel cost endogeneity arises when p_{ikt} is correlated with ε_{ikt}. This can arise, for example, if anglers are willing to drive farther in order to fish remote sites, inducing a positive correlation between remoteness and travel cost, but the researcher lacks a measure of remoteness.
- Two alternatives that distinguish between trips targeting salmonids versus those targeting warm water species are assigned to each site, yielding a total of 132 alternatives in the model.

Control Function

- Employing a control function is similar to standard instrumental variables methods involving a two-step estimation procedure, such as 2SLS.
- Travel cost endogeneity can be eliminated by conditioning the site choice model on the part of p_{ikt} correlated with ε_{ikt}. This factor can be estimated by regressing travel cost on the observables and an instrument and then calculating the residuals. Inserting the residuals, which is referred to as the control function ("*CF*"), into the model yields

$$U_{ikt}=\rho(y_i-p_{ikt})+\boldsymbol{\beta_1 q_k}+\boldsymbol{\beta_2 q_{kt}}+\varepsilon_{ikt}+\varphi CF+\omega_{ikt}$$

- Evidence for endogeneity is found if φ is estimated significantly different from zero. With a valid *CF*, $cov(p_{ikt},\omega_{ikt})=0$, and the parameters will be estimated consistently (Petrin and Train 2010).

Instrumenting for Travel Cost

- As an instrument, we use monthly crude oil prices interacted with the mileage distance to fishing sites. This instrument satisfies the condition that it be correlated with travel cost. It is not possible to test the exogeneity restriction, although unobserved characteristics, such as scenic beauty, are not likely to be correlated with (current) oil prices.

Alternative Specific Constants

•The ASCs method uses a vector of $K-1$ site-specific fixed effects, with K the number of choice alternatives. Denoting the ASCs vector as $\boldsymbol{\alpha_k}$, then angler utility is transformed into

$$U_{ikt}=\rho(y_i-p_{ikt})+\alpha_k+\boldsymbol{\beta_2 q_{kt}}+\zeta_{ikt}$$

where $\boldsymbol{\alpha_k}$ absorbs the observed effects $\boldsymbol{\beta_1 q_k}$ and unobserved site effects in ε_{ikt}. Assuming that $cov(p_{ikt},\zeta_{ikt})=0$ the remaining parameters can be estimated consistently.

•The $\boldsymbol{\beta_1}$ are not estimable in RUM model, but they can still be recovered. This is done by regressing the estimated ASCs ($\boldsymbol{\alpha_k}$) on the site-specific characteristics ($\boldsymbol{q_k}$). This auxiliary regression can be estimated with OLS or, if $\boldsymbol{q_k}$ is endogenous, with 2SLS (Murdock 2006).

•The ASCs method has limitations. For sites with few visits $\boldsymbol{\alpha_k}$ is unlikely to be well identified, hindering identification of $\boldsymbol{\beta_1}$ in the auxiliary regression. In fact, it is not possible to estimate the constants for sites with no visits, so unvisited sites are dropped from the model with ASCs.

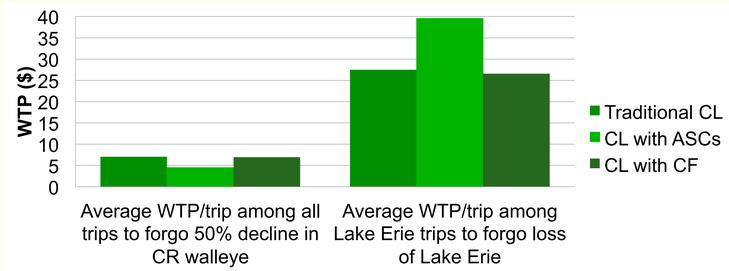
Results

- The site choice model is estimated as a conditional logit (CL). The first model, a traditional CL, ignores any travel cost endogeneity.

	Traditional CL		CL with ASCs[a]		CL with CF[b]	
Variable	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Travel cost	*-0.0348	-46.10	*-0.0359	-40.97	*-0.0357	-28.92
CR chinook	*11.759	14.10	*7.996	4.09	*12.018	13.43
CR coho	*6.546	3.34	6.802	2.05	*6.289	3.61
CR lake trout	0.327	0.62	-0.538	-0.45	0.338	0.80
CR steelhead	6.657	2.36	-4.835	-1.00	6.229	2.31
CR walleye	*3.927	12.62	*1.915	4.26	*3.954	13.07
CR yellow perch	*0.318	6.79	0.167	1.88	*0.303	7.15
Urban	*0.716	9.36	-0.469	-1.69	*0.727	10.10
Highway	*1.801	6.88	-0.298	-0.67	*1.799	6.60
Bayorseaway	*-1.091	-9.31	-0.157	-0.47	*-1.134	-9.08
Bayorseaway Warmwater*	*1.748	15.24	*1.260	2.94	*1.753	14.71
CF					*0.011	3.26
Trips	2233		2233		2233	
Log likelihood	-4184.530		-3645.443		-4174.000	

[a]The ASCs estimates are withheld for brevity. The estimates for *Urban*, *Highway* and the *Bayorseaway* variables are calculated from OLS of the ASCs.
[b]T-statistics derived from standard errors calculated from bootstrapping with 250 replications.
*Significant at the 1 percent level.

- Welfare implications:



Conclusions

- The control function estimate reveals statistically significant travel cost endogeneity.
- The travel cost coefficient in the model with ASCs is nearly equivalent to its counterpart in the model with the control function. However, the model with ASCs yields weaker results compared with the other models.
- Comparing the travel cost coefficients, the bias from travel cost endogeneity is small, so ignoring the endogeneity will not meaningfully affect estimates.

References

- Murdock, J. (2006). Handling Unobserved Site Characteristics in Random Utility Models of Recreation Demand. *Journal of Environmental Economics and Management* 51: 1-25.
- Parsons, G.R. (1991). A Note on Choice of Residential Location in Travel Cost Demand Models. *Land Economics* 67: 360-364.
- Petrin, A. and K. Train (2010). A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research* 47: 3-13.