Recreational trip timing and duration prediction: A research note

Atakelty Hailu[a] and Lei Gao[a*]

[a]School of Agricultural and Resource Economics, The University of Western Australia, Crawley, WA 6009, Australia
*E-mail address: dr.leigao@gmail.com

30 November 2010
Working Paper 1002
School of Agricultural and Resource Economics
http://www.are.uwa.edu.au

THE UNIVERSITY OF WESTERN AUSTRALIA
Achieve International Excellence

Citation: Hailu, A. and Gao, L. (2010) *Recreational trip timing and duration prediction: A research note*, Working Paper 1002, School of Agricultural and Resource Economics, University of Western Australia, Crawley, Australia.
Contact authors for a revised version of the paper.

Abstract:

This paper presents models that predict two recreational fishing trip parameters: the length of a trip and the timing of a trip within a year. A discrete choice (logit) model linking the choice of trip timing to calendar events, the demographic characteristics of anglers as well as the nature of the trip is econometrically estimated. A Tobit model is used to evaluate the relationship between fishing trip length and personal and trip characteristics. The results indicate that timing choice and trip length can be explained well in terms of observable personal and trip variables. Knowledge of these relationships is a useful input to tourism/recreational fishing management as well as to the development of tourism/fishing activity simulation models.

Keywords: recreational fishing, trip timing, length of recreational trips, tourism simulation, environmental impact management

1. Introduction

Increasingly, models are being used to simulate management outcomes and the effects of policy measures (Little et al., 2009, Kramer, 2008, McClanahan, 1995, McDonald et al., 2008). By systematically tracing the complex relationships between resource use and biophysical components, models allow us to better evaluate management and policy measures. For example, a model estimating the impact of fishing on biophysical stocks is likely to include sub-models linking fish species or groups to each other as well as fishing activities. However, while models of destination choice are usually modeled using empirical data, the timing and duration of recreation are rarely adequately modeled. Instead, *ad hoc* approaches are used to determine these parameters.

Trip timing has received little attention in literature. Some studies (Shailes et al., 2001, Eugenio-Martin and Campos-Soria, 2010) indirectly touch on the determinants of tourism trip timing, including the effects of congestion, climate, and demographic variables. There are relatively more studies focusing on tourists' length of stay. Recently, some researchers (Alegre and Pou, 2006, Gokovali et al., 2007, Martínez-Garcia and Raya, 2008) have used econometric models and identify different influences, including nationality, age, income, employment status, visitation rate, education, type of accommodation available, daily spending, and stage in the family life cycle. These studies focus at a length of stay at a particular destination.

In this paper, we used empirical data from multiple sites to develop a logit model of trip timing and a Tobit model of trip duration. The two models provide a general way of simulating recreational timing and duration that are superior to simpler approaches, e.g. those based on histograms or empirical frequencies. Unlike our approach, trip timing or length prediction methods that are based on observed frequencies do not relate the variables of interest to personal/trip characteristics and, are, therefore, difficult to extrapolate into other environments or periods. The

models presented here have been used as components in an integrated economic and ecosystem model of recreational fishing for a marine environment that includes trip demand and site choice models in (Gao and Hailu, 2010). The components of the simulation model are shown in Figure 1.

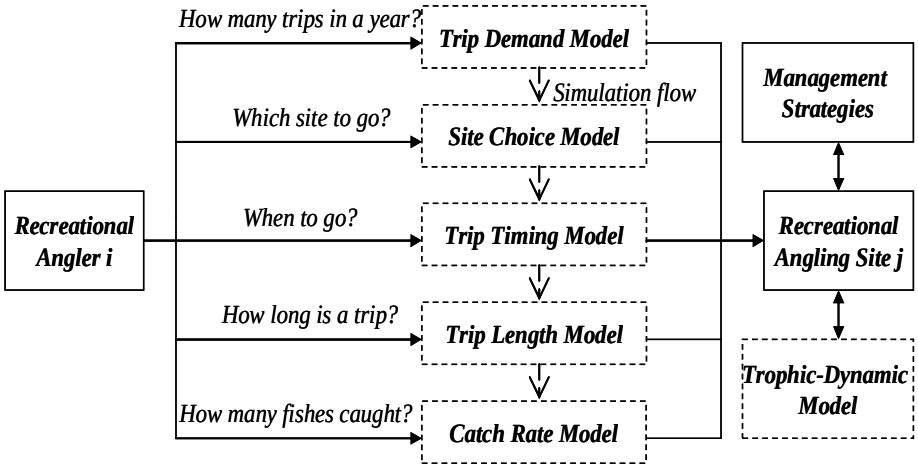

Figure 1. A set of econometric models that underpin agent-based simulation model of recreational fishing in a marine environment. Source: (Gao and Hailu, 2010)

The paper is organized as follows. The logit model for trip timing is presented in Section 2. This is followed by a presentation of the Tobit model of trip length in Section 3. The paper concludes in Section 4.

2. Trip timing model

The trip timing model focuses on the probability that an individual/household starts a trip on a given date. A logit model is used. This probability is hypothesized to be a function of three sets of factors:

1) the characteristics of the day (e.g. weekend, holiday, etc.),
2) the characteristics of the person (e.g. employment status), and
3) the nature of the trip (e.g. direction of the trip).

For example, an employed person will be inclined to choose a weekend or holiday to start a trip. This will particularly be the case for a fishing trip that involves travel for a weekend or a day. For longer trips, an employed person is likely to start a holiday at the end of the week. The demographic characteristics of the person are also important, e.g. retired people have more flexibility with recreational trips than employed or working people. Finally, the nature of the trip (e.g. whether it is going to north or south regions) will affect the choice of the timing because of weather effects.

The probability p_{ij} that a person *i* starts a trip on day *j* among all possible sets of days *d* is given by the following logit formula:

$$p_{ij} = \frac{\exp(\sum_k \beta_k x_{kj} + \sum_l \sum_m \beta_{lm} x_{li} x_{mj})}{\sum_d \exp(\sum_k \beta_k x_{kd} + \sum_l \sum_m \beta_{lm} x_{li} x_{md})} \tag{1}$$

where x_{kj} (or x_{mj}) is the *k*-th (or *m*-th) characteristics of day *j*, x_{li} is the *l*-th characteristics of angler *i*, and β_k and β_{lm} are the coefficients to be estimated.

In our case, we include the following variables:

$$\begin{aligned}\sum \beta_k x_{kij} + \sum_l \sum_m \beta_{lm} x_{li} x_{mj} = {} & \beta_0 + \beta_1 \cdot Weekend_j + \beta_2 \cdot PublicHoliday_j + \beta_3 \cdot SchoolHoliday_j \\ & + \beta_4 \cdot Employ_i \cdot PublicHoliday_j + \beta_5 \cdot Employ_i \cdot SchoolHoliday_j \\ & + \beta_6 \cdot Employ_i \cdot Weekend_j + \beta_7 \cdot KidProxy_i \cdot SchoolHoliday_j \\ & + \beta_8 \cdot Retire_i \cdot Weekend_j + \beta_9 \cdot South_i \cdot Jan_j + \beta_{10} \cdot South_i \cdot Feb_j + \beta_{11} \cdot South_i \cdot Mar_j \\ & + \beta_{12} \cdot North_i \cdot Apr_j + \beta_{13} \cdot North_i \cdot May_j + \beta_{14} \cdot North_i \cdot Jun_j + \beta_{15} \cdot North_i \cdot Jul_j\end{aligned} \tag{2}$$

These variables are defined in Table 1. The key variables include whether the day is a weekend, a public holiday, or a school holiday. However, the likelihood with which a person initiates a trip on a non-working day also depends on other factors, such as whether they are employed or not, or whether they have children or not in the case of school holidays. The direction of a trip and the month of the year together are an important influence. In the state where the data is collected, Western Australia, the northern half is warmer and thus attractive for recreational trips in the cooler months (like May to July). The reverse is true for trips heading south, where it is cooler in the south.

The reference point for the definition of a trip as south or north is the individual's origin or place of residence.

Table 1. Definition of variables in recreational fishing trip date choice.

Variables	**Description**
$Employ_i$	equals 1 if angler *i* is employed, and equals 0 otherwise;
$Retire_i$	equals 1 if angler *i* is retired, and equals 0 otherwise;
$KidProxy_i$	equals 1 if two or more persons in a household go fishing with angler *i*, and equals 0 otherwise;
$South_i$	equals 1 if angler *i* heads south to go fishing, and equals 0 otherwise;
$North_i$	equals 1 if angler *i* heads north to go fishing, and equals 0 otherwise;
$Distance_i$	distance (kilometers) between home location of angler *i* and the fishing site;
$Weekend_j$	equals 1 if day *j* is a weekend, and equals 0 otherwise;
$PublicHoliday_j$	equals 1 if day *j* is a public holiday, and equals 0 otherwise;
$SchoolHoliday_j$	equals 1 if day *j* is a public holiday, and equals 0 otherwise;
Jan_j	equals 1 if day *j* is in January, and equals 0 otherwise;
Feb_j	equals 1 if day *j* is in February, and equals 0 otherwise;
Mar_j	equals 1 if day *j* is in March, and equals 0 otherwise;
Apr_j	equals 1 if day *j* is in April, and equals 0 otherwise;
May_j	equals 1 if day *j* is in May, and equals 0 otherwise;
Jun_j	equals 1 if day *j* is in June, and equals 0 otherwise;
Jul_j	equals 1 if day *j* is in July, and equals 0 otherwise;

The sample data used for the estimation is drawn from the Australian National Survey of Recreation Fishing conducted in 2000/2001 (Henry and Lyle, 2003). A total of 3135 observations for 778 individuals from Western Australia were used in the analysis (Burton et al., 2008). For the trip direction variables, we divided the state into three zones (north, south and centre). These zones are used as indications of the effect of weather on the timing of trips. In the summer, the south is cooler and hence it is more likely that a trip to this zone will occur. Similarly in the winter the reverse will be true and trips to the north will be undertaken. We use 30 degrees south latitude (-30) and 33 degrees south latitude (-33) to classify recreational fishing destinations (48 fishing sites in total) as well as information on 17 angler home region locations to classify trips into the three categories, i.e. south, north, and center. To the reader get a sense of weather differences, we present in Table 2 the mean temperature data for three representative cities: Albany (south), Perth (central), and Geraldton (north).

Table 2. Mean temperature data (^{o}C) for three representative cities, Western Australia.

Location (Region)	Mean Temperature	Month												
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
ALBANY	Maximum	24.3	24.8	24.0	22.0	19.1	16.7	15.8	16.2	17.5	18.7	20.7	22.7	20.2
(South)	Minimum	13.7	14.3	13.6	11.7	10.0	8.2	7.4	7.6	8.3	9.2	10.9	12.4	10.6
PERTH	Maximum	31.9	32.1	29.9	26.1	22.2	19.2	18.1	18.7	20.3	22.9	26.3	29.1	24.7
(Center)	Minimum	17.3	17.6	16.2	13.4	10.9	9.0	8.1	8.3	9.3	10.5	13.3	15.2	12.4
GERALDTON	Maximum	31.6	32.8	31.2	28.2	24.5	21.2	19.6	20.0	22.0	24.7	27.5	29.4	26.1
(North)	Minimum	18.0	19.1	18.0	15.6	13.3	11.1	9.5	9.1	9.4	11.0	14.0	16.3	13.7

Note: The data is statistics result from year 1981 to 2010 from the Bureau of Meteorology, Australia (http://www.bom.gov.au/)

The maximum likelihood coefficient estimates for equation (2) are shown in Table 3. Almost all the variables are statistically significant, except the $North_i \cdot Apr_j$ interaction variable. April is the borderline month, between hot and cool season in the north, and it is not surprising that this variable is insignificant in the model.

Table 3. Coefficient estimates of trip timing model.

Variables[a]	Estimated coefficient	Std. Err.	z
Constant	-6.32503***	0.03137	-201.612
$Weekend_j$	0.49828***	0.07353	6.776
$PublicHoliday_j$	0.58434***	0.12476	4.684
$SchoolHoliday_j$	0.31548***	0.05603	5.631
$Employ_i \cdot Weekend_j$	0.36363***	0.08128	4.474
$Employ_i \cdot PublicHoliday_j$	0.40508**	0.15587	2.599
$Employ_i \cdot SchoolHoliday_j$	-0.20619**	0.06381	-3.231
$Retire_i \cdot Weekend_j$	-0.53059***	0.09943	-5.337
$KidProxy_i \cdot SchoolHoliday_j$	0.29356***	0.06550	4.482
$South_i \cdot Jan_j$	0.40984***	0.09087	4.510
$South_i \cdot Feb_j$	0.33794**	0.10931	3.092
$South_i \cdot Mar_j$	0.44416***	0.10080	4.406
$North_i \cdot Apr_j$	0.17936	0.10145	1.768
$North_i \cdot May_j$	0.33822**	0.10753	3.145
$North_i \cdot Jun_j$	0.23434*	0.11411	2.054
$North_i \cdot Jul_j$	0.28494**	0.10119	2.816

Note: log likelihood: -21326.31; Chi-square: 618.3177; p-value: 0.0. Three asterisks (***) indicate significance at the 0.1% level, while ** and * indicate significance at the 1% and 5% levels, respectively.

[a] variable definitions are provided in Table 1.

The coefficient estimates indicate that recreational anglers are inclined to select a holiday or non-working day (such as weekend, public holiday, and school holiday) to start a trip. If they are

heading south, they are likely to pick trip times in January, February, and March. If it is a trip to the north, on the other hand, they will tend to have it in May, June, and July; August to October dummies are not significant in our model. Further, employed anglers are more likely to select a weekend or a public holiday for a trip while retirees are more likely to select a weekday to go fishing. As expected, an angler with children is more likely to undertake a trip during school holidays. But it should also be noted that being employed makes one less likely to go fishing during school holidays, all else being equal. This suggests that working people without children prefer the quieter recreational periods outside the crowded school holidays.

3. Trip Length Model

A Tobit model is used to fit a model predicting the length of a fishing trip taken by an individual. Tobit (Amemiya, 1984) is an econometric model that is used to describe the relationship between a limited dependent variable (e.g. non-negative dependent variable) y_k ($k = 1,2,\cdots,n$) and an observed set of influences or explantors x_k. The model supposes that there is a latent (i.e. unobservable) variable y_k^*. Formally, the latent variable y_k^* is related to the explanatory variables as follows:

$$y_k^* = \sum \beta_k x_k + u_k, u_k \sim N(0,\sigma^2) \tag{3}$$

The latent variable is related to the observed variable y_k as follows:

$$y_i = \begin{cases} y_k^*, & y_k^* > 0 \\ 0, & y_k^* \le 0 \end{cases} \tag{4}$$

Here, β_k is the unknown vector of parameters that we want to estimate, x_k is a known vector of regression variables for the k-th observation, and u_k is assumed to be independently distributed with a symmetric error term.

Trip length in days (*lengthOfTrip*) is assumed to be a function of personal characteristics and the characteristics of the period during which the trip is taken:

$$\begin{aligned} lengthOfTrip &= \sum_k \beta_k x_{kj} + \sum_l \sum_m \beta_{lm} x_{li} x_{mj} \\ &= \beta_0 + \beta_1 \cdot Employ_i + \beta_2 \cdot Retire_i + \beta_3 \cdot KidProxy_i \\ &+ \beta_4 \cdot Weekend_j + \beta_5 \cdot PublicHoliday_j + \beta_6 \cdot SchoolHoliday_j \\ &+ \beta_7 \cdot South_i \cdot Jan_j + \beta_8 \cdot South_i \cdot Feb_j + \beta_9 \cdot South_i \cdot Mar_j \\ &+ \beta_{10} \cdot North_i \cdot Apr_j + \beta_{11} \cdot North_i \cdot May_j + \beta_{12} \cdot North_i \cdot Jun_j \\ &+ \beta_{13} \cdot North_i \cdot Jul_j + \beta_{14} \cdot South_i + \beta_{15} \cdot North_i + \beta_{16} \cdot Distance_i \end{aligned} \tag{5}$$

where x_{kj} (or x_{mj}) is the *k*-th (or *m*-th) characteristics of the trip start day *j*, x_{li} is the *l*-th characteristics of angler *i*, and β_k and β_{lm} are the coefficients to be estimated. The variables are defined in Table 1.

We use the same 3135 observations described above. A maximum likelihood method is used to obtain the coefficient estimates presented in Table 4.

Table 4. Coefficient estimates of the trip length Tobit model.

Variables[a]	**Estimated coefficient**	**Std. Err.**	**z**	**p**
Constant	1.110485***	0.0507	21.886	3.52E-106
$Employ_i$	-0.054103	0.0402	-1.346	0.178
$Retire_i$	0.016726	0.0463	0.361	0.718
$KidProxy_i$	0.048794	0.041	1.189	0.235
$Weekend_j$	-0.105353***	0.0309	-3.411	0.000648
$PublicHoliday_j$	0.138226*	0.0619	2.232	0.0256
$SchoolHoliday_j$	0.052224	0.0326	1.604	0.109
$South_i \cdot Jan_j$	0.115964	0.0789	1.47	0.141
$South_i \cdot Feb_j$	0.217101**	0.0932	2.33	0.0198
$South_i \cdot Mar_j$	0.089395	0.0862	1.037	0.3
$North_i \cdot Apr_j$	-0.076393	0.0869	-0.879	0.379
$North_i \cdot Mar_j$	-0.064605	0.0921	-0.702	0.483
$North_i \cdot Jun_j$	-0.182671	0.0978	-1.868	0.0618
$North_i \cdot Jul_j$	-0.109620	0.0912	-1.202	0.229
$South_i$	-0.019860	0.0413	-0.481	0.631
$North_i$	0.060258	0.0421	1.43	0.153
$Distance_i$	0.000358***	0.0000488	7.334	2.23E-13

Note: Number of observations used is 3135; log likelihood (model): -3804.2; log likelihood (intercept only): -3854.8; Chi-square: 101.16 on 16 degrees of freedom; p-value: 0.00. Three asterisks (***) indicate significance at the 1% level, while ** and * indicate significance at the 5% and 10% levels, respectively.

[a] Variable definitions are given in Table 1.

The results indicate that trip length is significantly affected by timing, distance of destination site and an interaction between the direction of trip and timing. The significant variables are: $Weekend_j$, $PublicHoliday_j$, $South_i \cdot Feb_j$, and $Distance_i$. The distance from a recreational angler's home location to recreational site is positively correlated with the trip length. Also, holidays (including weekends) included in a trip time also affect the length of the trip. Trips initiated over a weekend tend to be shorter, all else being the same. Recreational anglers who head south in February are likely to go on longer trips those travelling in other months. Other attributes of the person (e.g. employed, retired, or whether they have children) and other interactions between direction of trip and the timing of the trip seem to be statistically insignificant. However, there is some evidence that the length is likely to be longer if the angler is unemployed or retired, or is accompanied by children.

4. Conclusion

Models predicting trip timing and trip length are useful. There has been little systematic research on trip timing and little research on duration modeling involving multiple destinations. Potential areas of use include tourism promotion, congestion management as well as environmental impact management. These models are also useful for research purposes as they provide a key input into simulation models. Currently, models simulating recreational behaviour utilize *ad hoc* approaches. Even if these models utilize empirical models that relate site choice to personal and site characteristics (e.g. random utility models of fishing site choice), the timing and length choice predicitons for trips are rarely based on empirically sound grounds.

This paper has developed and econometrically estimated two models. A trip timing model for predicting the probability that a persons/household initiates a recreational trip on a given day was estimated as a logit model. The results show that a large number of variables (demographic and trip

related) can be used to explain trip timing. A Tobit model was econometrically estimated to determine influences on trip length and the results show that demographic variable as well as calendar events are important in predicting trip durations.

Acknowledgements

We thank J. Durkin for comments on the manuscript. The research reported here is part of a bigger project funded through the Ningaloo Collaboration Cluster, CSIRO Wealth from Oceans Flagship Program.

References

Alegre, J. & Pou, L. (2006), 'The length of stay in the demand for tourism', *Tourism Management,* Vol 27, No 6, pp 1343-1355.

Amemiya, T. (1984), 'Tobit models: A survey', *Journal of Econometrics,* Vol 24, No 1-2, pp 3-61.

Burton, M. P., Raguragavan, J. & Hailu, A. (2008), 'State-wide recreational fishing site choice study', *School of Agricultural and Resource Economics, University of Western Australia, Technical Report of Ningaloo Cluster Project 4,* Vol, No pp.

Eugenio-Martin, J. L. & Campos-Soria, J. A. (2010), 'Climate in the region of origin and destination choice in outbound tourism demand', *Tourism Management,* Vol 31, No 6, pp 744-753.

Gao, L. & Hailu, A. (2010), Integrating recreational fishing behaviour within a reef ecosystem as a platform for evaluating management strategies, in *The 24th IEEE International Conference on Advanced Information Networking and Applications*, Perth, Western Australia.

Gokovali, U., Bahar, O. & Kozak, M. (2007), 'Determinants of length of stay: A practical use of survival analysis', *Tourism Management,* Vol 28, No 3, pp 736-746.

Henry, G. W. & Lyle, J. M. (2003). *The National Recreational and Indigenous Fishing Survey,* Canberra, Australia, Australian Government Department of Agriculture, Fisheries and Forestry.

Kramer, D. B. (2008), 'Adaptive Harvesting in a Multiple-Species Coral-Reef Food Web', *Ecology and Society,* Vol 13, No 1, pp 17. [online] http://www.ecologyandsociety.org/vol13/iss1/art17/.

Little, L. R., Punt, A. E., Mapstone, B. D., Begg, G. A., Goldman, B. & Williams, A. J. (2009), 'An agent-based model for simulating trading of multi-species fisheries quota', *Ecological Modelling,* Vol 220, No 23, pp 3404-3412.

Martínez-Garcia, E. & Raya, J. M. (2008), 'Length of stay for low-cost tourism', *Tourism Management,* Vol 29, No 6, pp 1064-1075.

Mcclanahan, T. R. (1995), 'A coral reef ecosystem-fisheries model: impacts of fishing intensity and catch selection on reef structure and processes', *Ecological Modelling,* Vol 80, No 1, pp 1-19.

Mcdonald, A. D., Little, L. R., Gray, R., Fulton, E., Sainsbury, K. J. & Lyne, V. D. (2008), 'An agent-based modelling approach to evaluation of multiple-use management strategies for coastal marine ecosystems', *Mathematics and Computers in Simulation,* Vol 78, No 2-3, pp 401-411.

Shailes, A., Senior, M. L. & Andrew, B. P. (2001), 'Tourists' travel behaviour in response to congestion: the case of car trips to Cornwall, United Kingdom', *Journal of Transport Geography*, Vol 9, No 1, pp 49-60.