Using Spatial Microsimulation Modelling Techniques and Geographic Information Systems to Estimate the Demand for Outdoor Recreation in Ireland*

John E. Cullinan†, Cathal O'Donoghue and Stephen Hynes

Rural Economy Research Centre (RERC), Teagasc, Ireland
and
Department of Economics,
National University of Ireland, Galway

Abstract

This paper discusses how spatial microsimulation modelling techniques can be used in conjunction with geographical information systems to model the demand for non-priced, open-access outdoor recreation activities in Ireland. By integrating data from a number of sources, this paper proposes a methodological approach for predicting the number of visitors to outdoor recreation sites.

In particular, we present a general framework for modelling the demand for outdoor recreation using the 'Simulation Model for the Irish Local Economy' (SMILE). We discuss how a GIS-based spatial microsimulation model can be used to simulate a spatially distributed population of recreationalists. We also describe how behavioural models explaining the demand for trips among recreationalists can be combined with the simulated population to predict the overall demand for trips.

This paper also presents an illustrative example of the potential of spatial microsimulation models for estimating the demand for outdoor recreation, using a synthetic population of recreationalists and applying benefit function transfer procedures. Initially, a synthetic population of recreationalists is developed, based on a simple matching exercise between survey data for a group of whitewater kayakers and population distribution data. A 'spatial interaction modelling' exercise is then undertaken by modelling the demand of the synthetic population for recreation at different sites across Ireland using a transferable distance-decay function.

Keywords: Spatial Microsimulation, Spatial Interaction Analysis, Geographic Information Systems, Benefit Function Transfer, Visitor Arrival Function Transfer, Distance-Decay Function, Non-Priced Open-Access Recreation, Whitewater Kayaking.

* Draft – not for quotation.

† Address for correspondence: Rural Economy Research Centre, Teagasc, Athenry, Co. Galway, Ireland. Email: jcullinan@rerc.teagasc.ie.

1. Background and Motivation

1.1 Background

This paper discusses how spatial microsimulation modelling techniques can be used in conjunction with geographical information systems (GIS) to model the demand for non-priced, open-access outdoor recreation activities in Ireland. By integrating data from a number of sources, this paper proposes a methodological approach for predicting the number of visitors to outdoor recreation sites.

In particular, we present a general framework for modelling the demand for outdoor recreation using the 'Simulation Model for the Irish Local Economy' (SMILE)[1]. We discuss how a GIS-based spatial microsimulation model can be used to simulate a population of recreationalists at the level of small-area electoral division (ED), using a combination of survey data and small-area census of population data. We also describe how behavioural models explaining the demand for trips among recreationalists can be combined with the simulated population to predict the overall demand for trips.

This paper also presents an illustrative example of the potential of spatial microsimulation models for estimating the demand for outdoor recreation, using a synthetic population of recreationalists and applying benefit function transfer[2] (BFT) procedures. Initially, a synthetic population of recreationalists is developed, based on a simple matching exercise between survey data for a group of whitewater kayakers and population distribution data. A 'spatial interaction modelling' exercise (see Brainard *et al.*, 1999) is then undertaken by modelling the demand of the synthetic population for recreation at different sites across Ireland using a transferable distance-decay function.

Overall this paper considers how spatial microsimulation can be used in conjunction with GIS to predict the demand for recreation. The research represents work-in-progress and is one element of ongoing doctoral research being undertaken by one of the authors (Cullinan[3]). This paper therefore should be thought of as a staging post for methodological developments in an on-going research project.

---

[1] SMILE is a spatial microsimulation models that is currently being developed jointly by Teagasc, NUI Galway and the University of Leeds. SMILE uses survey and census data for a given year to develop spatially disaggregated population microdata estimates for each ED in Ireland and is used primarily to analyse spatial issues in Ireland (see Ballas *et al.*, 2005a).

[2] Benefit transfer (BT) methods involve the transferral of existing estimates of non-market values, such as willingness to pay, from a study site **s** to a different policy site **p.** When the transfer is conducted by means of an estimated function or equation, the process is known as benefit function transfer.

[3] The research topic under consideration by Cullinan is "The Use of Spatial Microsimulation Modelling Techniques in Estimating the Demand for Outdoor Recreation Activities in Ireland."

1.2 Motivation

The principal objective of SMILE is to evaluate the spatial impact of changes in society and the economy. While work to-date has focussed on the development of the modelling framework of SMILE, the motivation of this research project is to extend SMILE to focus on the potential for, and economic impact of, diversifying from agriculture to agri-tourism, in particular outdoor pursuit based activities. Specifically, the aim of this research is to develop a methodology for linking recreation demand models to SMILE in order to produce an agri-tourism demand module within SMILE. This can then be combined with environmental information to create indicators of potential agri-tourism hotspots in Ireland. Thus the initial motivation for this research centres on developing spatial estimates of visitor numbers to outdoor recreation sites in Ireland using SMILE and modelling the impact on the local economy.

A second motivation for this research comes from the environmental valuation literature, and the benefit transfer literature in particular. Benefit function transfer is an increasingly important issue within the area of environmental resource valuation as it provides an alternative to conducting individual studies at different sites. Thus, a second goal of this research is to investigate the potential for developing transferable models that estimate visitor numbers to recreation sites. While benefit function transfer has previously been used in Ireland to estimate measures of willingness to pay (WTP) at various 'policy' sites based on estimates of WTP at a particular 'study' site, little work has been done in considering transferable models that estimate visitor numbers. In other words, previous research has concerned itself with the transferral of value estimates as opposed to visitor estimates. Since the total estimated value of a recreation site is a function of both variables, this research addresses an important element in the environmental valuation literature, the development of 'visitor arrival function transfer' (VAFT) models.

A further contribution of this research is its consideration of the potential benefits of geographic information systems in conjunction with spatial microsimulation and BFT. GIS can help to address the issue of 'spatial complexity' in BFT by helping to consistently estimate distance, travel times and costs, population distributions, visitor outset origins, recreation location points as well as other variables commonly used in transfers. According to Bateman *et al.* (2004), "it is the potential for benefit transfer analyses offered by GIS methods which present the most exciting opportunities in this field."

In the following sections we start by presenting a brief discussion of spatial microsimulation and discuss some previous applications. Next we introduce SMILE and discuss how we propose to use it as a framework for considering the demand for outdoor recreation in Ireland as well as other relevant issues. We then present an illustrative example as well as preliminary results and potential outputs from such an approach. The final section concludes and discusses planned future work.

2. Spatial Microsimulation Models[4]

2.1 Introduction to Spatial Microsimulation Models

Microsimulation can be defined as "a methodology that is concerned with the creation of large-scale simulated population microdata sets"[5] in order to describe economic and social events by modelling the behaviour of individual agents. Microsimulation models are typically large-scale datasets of the attributes of micro units (individuals, households or firms) and are useful in evaluating economic, social and tax policy.

There are three main types of microsimulation models, namely static, dynamic and spatial. Static models examine micro units at one point in time and have been used extensively to examine the differential impact of budget changes. Dynamic microsimulation models project the population in a base year forward through time by simulating transitions such as fertility and mortality at the individual level. Finally, spatial microsimulation models have also been developed and link individuals, households or firms with a location. They are generated, in general terms, by merging spatially disaggregated census data and survey data in order to simulate a population of individuals and households which have characteristics that are as close as possible to the true overall population. They can then be used to explore spatial relationships and to analyse the spatial implications of policy scenarios.

While microsimulation modelling has traditionally been used to consider tax-benefit policy, it is increasingly being used to analyse the spatial impact of policies and regional development issues (see for example Holm *et al.*, 1996; Ballas and Clarke, 2001; King, 2003; Ballas *et al.* 2006). This increase in spatial microsimulation modelling is mainly due to the availability of spatial data from national censuses and new methods to generate micro data from aggregated tabulations (Williamson, 2003). According to Ballas *et al.* (2006), "one of the major advantages of microsimulation is that it can be a substitute for conducting detailed surveys to produce survey data ….. at the small area level." Crucially, "the power of spatial microsimulation modelling frameworks lies in their ability to estimate policy-relevant variables at the small area level for which published data do not exist.[6]"

O'Donoghue *et al.* (2004) set out four additional advantages of spatial microsimulation models. First, they allow data from different sources to be linked if datasets contain at least one attribute in common. In addition, survey data (such as the European Community Household Panel Survey) can be linked to cross-sectional data (such as the Census of Population). A second advantage is that models are flexible in terms of their spatial scale so that data can be re-aggregated or disaggregated. For example, outputs from SMILE can be aggregated to counties, regions or the country as a whole. A third advantage is that spatial microsimulation models store data efficiently as lists while a fourth advantage is that such models allow for updating and projecting.

---

[4] This section is based on O'Donoghue *et al.* (2004) and Ballas *et al.* (2005b).
[5] Ballas *et al.* (2006).
[6] *Ibid*.

2.2 Previous Applications of Spatial Microsimulation Models

There are notable examples of spatial microsimulation models in Britain, Sweden, the Netherlands, Australia, the US and elsewhere. In Britain for example, SYNTHESIS, a spatial microsimulation model developed at the University of Leeds, estimated a spatially disaggregated micro dataset for the Leeds metropolitan district and used this dataset to generate incomes for individuals (Birkin and Clarke, 1988, 1989). Ballas and Clarke (2000, 2001a,b) at the University of Leeds have subsequently developed SimLeeds, a spatial microsimulation model for the Leeds labour market. This model has been used to explore the potential spatial impact of a factory closure on Leeds. Ballas and Clarke (2000) estimated the areas that would likely suffer the most job losses and subsequently have the largest multiplier effects on their local economy. SimYork was developed as a spatial microsimulation model for York (Ballas *et al.*, 2001) and as a pilot study for extending SimYork to the whole of Britain – see Ballas *et al.* (2005) for a discussion of SimBritain. A recent study by Clarke *et al.* (2006) considers the use of GIS and spatial microsimulation in the analysis of health inequalities.

Spatial microsimulation models have also been developed in other countries. In Sweden, the Total Population Simulation Model (TOPSIM) and the System for Visualizing Economic and Regional Influences Governing the Environment (SVERIGE) have been developed. TOPSIM is a national microsimulation model with the location of individuals defined by their municipality of residence. It is a dynamic model and contains demographic and labour force characteristics (see Holm *et al.*, 1996). SVERIGE is the first national level spatial microsimulation model based on a longitudinal database of socio-economic information on every resident of Sweden between 1985 and 1995. The database also contains geographic information in the form of co-ordinates accurate to 100 meters, for each resident (Vencatasawmy *et al.*, 1999). SVERIGE is designed to generate geographically detailed reports for policy makers and regional scientists.

At the University of Utrecht in the Netherlands, event-history analysis, microsimulation and GIS are combined to study how firms and individuals are linked through space. Daily mobility, such as commuting, and lifetime mobility, such as residential relocation or migration, have been analysed at the micro level for individuals. Firms were analysed in a similar fashion in terms of moving inputs and outputs on a daily basis and locating the firm over its lifetime. This analysis is designed to be of use for regional planning particularly in the context of increased awareness of environmental concerns related to commuting and transporting goods (Hooimeijer, 1996).

In Australia, NATSEM (National Centre for Social and Economic Modelling), at the University of Canberra, has developed a spatial microsimulation model to examine issues such as poverty and ageing in a spatial context (Harding, 2002). In particular, a regional microsimulation model has been developed in conjunction with Centrelink, the agency responsible for administering social benefit payments, to project regional demographics and likely use patterns for Centrelink services (King, McLellan and Lloyd, 2002). Recent work by NATSEM includes spatial microsimulation modelling of care needs, costs and the capacity for self-provision for older Australians to 2020 – see NATSEM (2004).

The above represents only a selection of spatial microsimulation models that have being used to analyse the implications of policy at a spatial level. Other examples of applications of such models, though by no means an exhaustive list, include using spatial microsimulation approaches to model households and location choices in metropolitan Manila (Tiglao and Tsutsumi, 2005) and modelling crime using spatial microsimulation (Kongmuang, 2004).

3. Using SMILE to Model the Demand For Recreation

3.1 Introduction to SMILE

SMILE is a static, dynamic population, spatial microsimulation model, constructed using the 2002 Census of Population Small Area Population Statistics (SAPS)[7]. SMILE is an object-oriented model, built in Java and C++. It contains two types of process. First, the static process creates the base population and assigns census attributes to individuals. SMILE uses survey and census data for a given year to develop spatially disaggregated population microdata estimates for each electoral division[8] (ED) in Ireland. Specifically, European Community Household Panel Survey (ECHP) data is used as the population microdata, while Small Area Population Statistics (SAPS) data from the Irish Census of Population are used as the geographically referenced small area population data. The second process, known as the dynamic process, ages the population by simulating life cycle characteristics such as demographics, labour market outcomes and migration patterns. The approach discussed in this paper does not utilise the dynamic process in considering the demand for recreation.

SMILE is designed to analyse the relationships among regions and localities and to project the spatial implications of economic development and policy change in rural areas. Potential analyses using the model include spatial concentrations of social exclusion, spatial projections of the demand for housing, analysis of commuting patterns, the impact of declining population in remote areas, regional policy formation and other spatial related policy. It can also be used to help model the demand for outdoor recreation activities in Ireland as discussed below.

[7] See O'Donoghue *et al.* (2004) for a more detailed discussion of SMILE.
[8] There are a number of different spatial categorisations within Ireland from nation through province, county and small area electoral division.

3.2 Proposed Methodological Approach to Model the Demand for Recreation

While the main objective of SMILE is to evaluate the spatial impact of changes in society and the economy in Ireland, the aim of this ongoing research project is to develop a methodology for linking recreation demand models to SMILE in order to produce an agri-tourism demand module. <u>In particular we are interested in combining behavioural models that predict visitor numbers to outdoor recreation sites in Ireland to datasets within SMILE and undertaking spatial interaction modelling of the total demand for outdoor recreation activities at different sites</u>. These estimates can then be used in conjunction with SMILE to model the impact on the local economy and be combined with environmental information to create indicators of potential agri-tourism hotspots in Ireland. Thus the research aims to develop spatial estimates of visitor numbers to outdoor recreation sites in Ireland using SMILE and to subsequently model the impact on the local economy.

The overall objective is to use SMILE as a mechanism for linking spatial data from a number of sources by merging datasets (for example at ED level) and by undertaking statistical matching exercises, in order to construct a database of consistent, spatially referenced datasets. These merged datasets, which will form the spatial microsimulation model, can then be used as inputs in spatial and economic models of individual behaviour.

In the case of modelling the demand for outdoor recreation, it is proposed to use SMILE in conjunction with econometric models, benefit function transfer procedures, geographical information systems analysis, network analysis and other modelling techniques to undertake the spatial interaction analysis. There are three main steps involved as set out in Table 1 and these are discussed in more detail below.

Table 1: Proposed Methodological Approach

1.	Generate population microdata estimates for each ED in Ireland
2.	Link spatial datasets to SMILE and populate the model with information about recreationalists
3.	Link behavioural models to SMILE *via* benefit function transfer procedures

3.2.1 Generate spatially disaggregated population microdata estimates

The first element of the methodological approach begins by generating population microdata estimates for each ED in Ireland. As discussed above, spatial microsimulation techniques involve the merging of survey data such as the ECHP with census or other geographical area data to simulate a population of individuals. In the words of Ballas *et al.* (2006), spatial microsimulation models aim to "simulate virtual populations in given geographical areas, so that the characteristics of these populations are as close as possible to their 'real-world' counterparts." To this end, SMILE uses ECHP survey data and SAPS data to create a small area population microdata set containing data on a range of different personal and socio-economic characteristics for each individual. Such data can then easily be used as an input in a GIS to produce maps of the simulated information[9].

In terms of the procedures involved in this stage, spatial microsimulation implies the "reweighting of an existing microdata sample (which is only available at coarse levels of geography), so that it would fit small area population statistics tables"[10]. So, for example, we can take a microdata set such as the ECHP and reweight it to 'populate' the small area electoral divisions. Since the ECHP includes detailed information for a sample of individuals, in the reweighting process we aim to sample from this dataset in order to select a set of individuals that "best matches the population"[11] described in the SAPS dataset by ED. This then provides us with a simulated population of individuals with associated personal and socio-economic characteristics. It is not proposed to discuss the statistical matching techniques required to undertake the reweighting process in this paper – see O'Donoghue *et al.* (2004) for a discussion of the procedures used in SMILE.

3.2.2 Link spatial datasets to SMILE and populate the model with information about recreationalists

In this research project the main aim is to use SMILE as a mechanism for linking spatial data from a number of sources that can be used as inputs in behavioural models that explain the demand for recreation, amongst other things. There are two ways in which we propose to link data: (i) by merging relevant spatial datasets; and (ii) by undertaking data fusion by statistical matching.

In relation to the merging of spatial datasets, the Rural Economy Research Centre is currently involved in establishing a spatial data archive to act as a repository for historical, current and future spatial data, including census data, farm level data, environmental data, business directory data, infrastructure data, public policy data, and other relevant spatial datasets. The aim is to develop a consistent, spatially referenced database to allow analysis that produces outputs that can be represented using geographical information systems. The merged datasets, which form the spatial microsimulation model, will also be used as inputs in spatial and economic models of individual behaviour, including in relation to recreation demand.

It is also intended to incorporate information about recreationalists into SMILE by undertaking data fusion by statistical matching. In principle this will require survey data for a group of recreationalists in relation to their residential locations as well as some personal and socio-economic characteristics. Data fusion techniques will then be utilised to populate SMILE with recreationalists to create a simulated population of recreationalists at a spatial level (either at county or ED level).

[9] "One of the major advantages of frameworks based on GIS spatial microsimulation is the ability to create thematic maps of the simulated information and to explore whether there is any spatial pattern" - Ballas *et al.* (2006).
[10] Ballas *et al.* (2006).
[11] *Ibid*.

In relation to data fusion procedures, these techniques rely on merging datasets that have a subset of variables that are common to both, in order to create a dataset with the necessary set of variables – see Redway (2005) for a full discussion of the techniques involved. It is planned to use these techniques to match recreationalists from survey data to the SMILE dataset, to essentially populate the SMILE microsimulation dataset with characteristics of recreationalists and to benefit from the broad range of variables contained within SMILE.

3.2.3 Link behavioural models to SMILE via benefit function transfer procedures

The third step in our approach involves developing econometric models that explain individual behaviour in relation to recreation demand and linking these models to the SMILE spatial microsimulation model. For example, having simulated a population of recreationalists at ED level, we might seek to develop a model that predicts the total number of trips made to a particular recreation site as a function of where an individual recreationalist lives. Other factors are also likely to impact on the demand for recreation activities, such as the specific attributes of different recreation sites and the personal and socio-economic characteristics of recreationalists, and these variables should also be included in the behavioural model.

The process by which such behavioural models will be linked to the datasets within SMILE involves the use of benefit function transfer techniques. In general terms, BFT entails "the transfer of a benefit or demand function from a study site, [which] adapt the function to fit the specifics of the policy site such as socioeconomic characteristics, extent of market and environmental impact, and other measurable characteristics that systematically differ between the study site(s) and the policy site. The adapted function is then used to forecast a benefit measure for the policy site."[12] In the case of the BFT proposed for this approach we intend to use the transfer function across different individuals, as well as different sites.

Essentially the aim will be to combine the behavioural models with the simulated recreationalists within SMILE to undertake a spatial interaction analysis of the demand for trips by each individual. This demand may relate to an individual's overall demand for trips or to their demand for trips at different recreation sites. Within this context, there are be scope to utilise geographical information systems to consistently estimate visitor outset origins, recreation location points, travel distances, as well as other variables required for the BFT. Using GIS can also facilitate the implementation of network analyses. A detailed example of a spatial interaction analysis incorporating these elements is set out in the next section, including an illustrative example of the potential outputs from such an approach.

[12] Rosenberger and Loomis (2001).

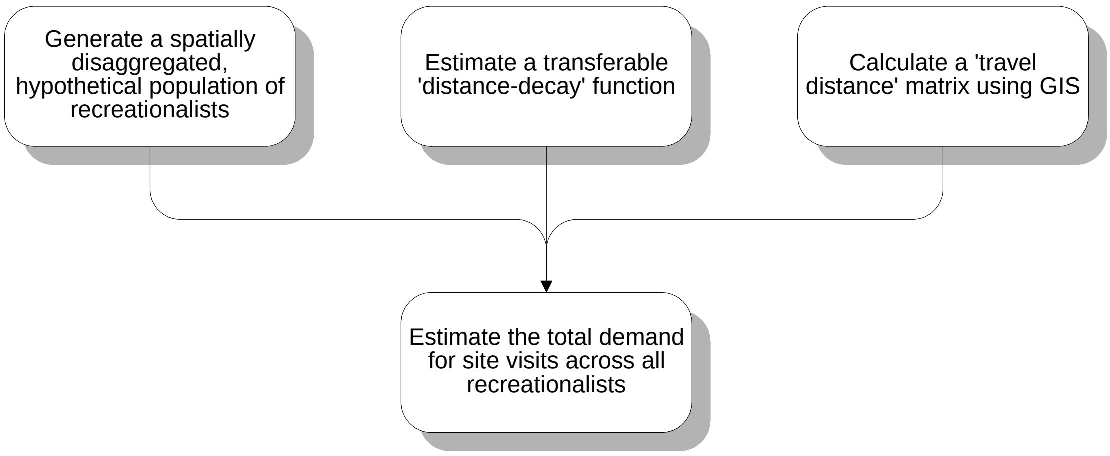
4. Illustrative Example and Preliminary Results

4.1 Introduction

The overall approach outlined above includes three principal elements for modelling the demand for outdoor recreation in Ireland. The first element involves creating population microdata estimates for each ED in Ireland and work is almost complete in relation to this element – see O'Donoghue *et al.* (2004) for details. The second element involves linking spatial datasets to SMILE and populating the dataset with information about recreationalists and work is ongoing in relation to this element. Finally the third element involves combining the spatial microsimulation model with behavioural models to estimate the demand for recreation. In this section we set an illustrative example of this third element, specifically the development of transferable models that explain the behaviour of recreationalists and the development of a simple 'spatial interaction model' of the demand for recreation. The results presented below are preliminary and should be considered as work-in-progress.

There are three main steps involved in our initial spatial interaction model, as presented in Figure 1. First, we develop a spatially distributed, synthetic/ hypothetical population of whitewater kayakers. Second, we estimate a transferable, individual-specific, distance-decay function[13] and third, we estimate the distance that each whitewater kayaker in the synthetic population must travel to visit each of a number of kayaking put-in points. These estimated distances are then used in the estimated transferable distance-decay function to predict the demand for trips by each individual in the synthetic population of kayakers to each site. Summing these estimates over all individuals generates estimates of total trips per site.

Figure 1: Initial Spatial Interaction Model



[13] A distance-decay function can be defined as the mathematical representation of the effect of distance on the accessibility and number of interactions between locations. The functions estimated in this paper are individual-specific, since I model the relationship between the number of trips made by an individual kayaker to a whitewater kayaking site and the distance (s)he must travel to that site.

The analysis is based on two principal datasets. First we utilise data on a group of whitewater kayakers from Hynes *et al.* (2005) in both generating the synthetic population of kayakers and for estimating a distance-decay function that models the demand for kayaking trips. In generating the spatially distributed synthetic population of kayakers we also use the Small Area Population Statistics (SAPS) dataset from the 2002 Census of Population in Ireland.

4.2 Generate a Spatially Distributed, Synthetic Population of Kayakers

The previous section set out a discussion on incorporating details about recreationalists within SMILE and subsequently using the simulated, spatially distributed, population in modelling the demand for recreation activities. Since work is ongoing and not yet complete in relation to this element of the overall approach, we present instead an alternative approach that is being used initially to help develop the overall methodology. It should be stressed that this represents a very simplistic substitute for the on-going work in relation to populating SMILE with characteristics of recreationalists. It essentially involves matching survey data on kayakers from Hynes *et al.* (2005) with SAPS data on the spatial distribution of the total population by electoral division (ED) in Ireland. In the first instance the proportion of kayakers in each county was estimated using survey data[14] and then multiplied by an estimate of the total number of kayakers in Ireland to develop estimates of the total number of kayakers by county[15]. The next step involved allocating the kayakers in each county by ED within that county based on the spatial distribution of the total population in Ireland in 2002 by electoral division using SAPS data. Specifically, the proportion of a county's population in each ED was used to allocate kayakers, hence effectively matching the proportion of a county's kayakers in each ED to the proportion of a county's population in each ED.

The estimated number of kayakers per ED is presented in Map 1 and suggests a concentration of kayakers in and around the principal population centres of Dublin, Cork, Limerick and Galway. Kayakers are concentrated, as expected, in the principal urban areas, where populations are greatest and university and non-university kayaking clubs are located[16].

It is worth stressing again that work is ongoing in relation to populating SMILE with recreationalists' characteristics and that this part of the approach will be significantly improved upon in future work.

---

[14] This spatial distribution by county is taken to be representative of the whitewater kayaking population as a whole in Ireland. This is a very strong assumption to make and given the relatively small sample size involved (279 individuals) is likely to be somewhat inappropriate. It is intended in the future to gather more extensive and precise data on the residential locations of whitewater kayakers to improve this element of the approach and we are awaiting more detailed data from the Irish Canoe Union (ICU) in relation to kayaker residential locations. Since however the main focus of this section is on the application of a VAFT using a distance-decay function and the usefulness of GIS in this context, we feel that overall the assumption is useful in helping to develop the methodology in this paper.

[15] An estimate of the total number of kayakers in Ireland of 5,000 from the ICU was used.

[16] In fact, the correlation between the estimated number of kayakers per county and the actual number of kayaking clubs per county is 0.82.

4.3 Estimate a Transferable Distance-Decay Function

The next step in the approach involves estimating a transferable distance-decay function. In particular we wish to estimate the relationship between the number of trips made per kayaker and a set of explanatory variables, including distance travelled. Since the dependent variable is a positive integer count variable, we estimate a series of count data models. Such models are typically estimated using either a Poisson or negative binomial model (see Cameron and Trivedi, 1998). The basic model can be represented by Equation [1]:

$$T = f(D, S, R) \qquad [1]$$

where T, the dependent variable, is the number of trips taken by individual *i* to site *j*. Included in the regressors, **D** denotes a vector of distance related explanatory variables, **S** denotes a vector of site attribute variables and **R** is a vector of recreationalists' characteristics. The full list of explanatory variables considered in the econometric modelling is set out in Table 2. A series of reduced form models and full models were estimated and models were compared on the basis of tests of over-dispersion, heterogeneity and fixed effects in order to choose the most appropriate transferable model (see Annex 1 for details).

4.4 Calculate Travel Distances Using a Geographical Information System

In order to predict the total number of trips to different rivers using the estimated distance-decay function, we need to know, *inter alia*, the distance that each kayaker in our synthetic population must travel to each recreation site. This can be estimated using a GIS. As a first step, the spatial (Irish grid) coordinates for the centroid[17] of each ED was calculated using ESRI's ArcGIS 9.0 XTools Pro extension, which allows centroid coordinates to be calculated for a georeferenced polygon. These coordinates take the form of precise easting and northing (x,y) grid references calibrated for the Irish national grid. The grid coordinates for the eleven separate kayaking put-in points in the Hynes survey were then identified and mapped in the GIS as a separate point feature class layer. Furthermore, coordinates for a selection of additional selected kayaking rivers were also obtained and mapped in the GIS in order to test out-of-sample predictive properties. Map 2 presents these three layers of grid coordinates.

ArcGIS can be also used to calculate the Euclidean distances between specified grid coordinates both between and within layers. Specifically, the calculation of distances between assumed visitor outset points (i.e. ED centroids) and kayaking put-in points were calculated using the Hawth's Tools extension within ArcGIS. Initially we calculated the straight-line 'Euclidean' distance between each of the 3,441 ED centroids in the State and the eleven kayaking put-in points resulting in a 3,441 x 11 matrix of Euclidean distances. These distances can then be used as an input in the estimated transferable distance-decay function to estimate the number of trips taken by a kayaker in a particular ED to each of the different kayaking sites.

[17] A centroid of an ED is defined as its geometric centre. It can be thought of as the point within an ED on which it would balance when placed on a needle, assuming that the ED was a smooth flat surface.

However, as pointed out by Brainard *et al.* (1999), "the inadequacies of Euclidean geometry for describing access are several, not least because it assumes a homogenous road network in all directions". As an initial method to deal with this issue, the estimated distances were uprated using a 'distance adjustment factor', estimated with ArcGIS and using data on the road network in Ireland, in order to better approximate the distance travelled by kayakers to different sites[18].

4.5 Preliminary Results

The final element in the approach involves estimating the total number of trips made by the synthetic population of kayakers to each of the eleven rivers considered by combining the above three elements. First we take the spatially disaggregated, synthetic population of kayakers, which provides information on each of the estimated 5,000 kayakers in Ireland, including the spatial information necessary to determine the distance a kayaker must travel to a particular kayaking site. This distance is a key input in the transferable distance-decay function. Using the distance-decay function, the estimated number of trips for individual *i* to a particular whitewater kayaking river *j* is then estimated as $\hat{T}_j = \sum_{i=1}^{N} \hat{\lambda}_{i,j}$, where $\hat{\lambda}_{i,j}$ is the estimated number of trips by individual *i* to river *j* - see the footnotes to Tale 3 and Table 4 in Annex 1 for further details on estimating $\hat{\lambda}_{i,j}$ for the reduced form and full models respectively.

In terms of the outputs of the approach, the total estimated annual number of trips is presented in Table 5 for each of the eleven recreation sites considered in the Hynes *et al*. (2005) survey, while Table 6 presents estimates of trips to a selection of five others kayaking rivers in Ireland.

5. Conclusions and Future Work

This paper discusses how spatial microsimulation modelling techniques can be used in conjunction with geographical information systems to model the demand for non-priced, open-access outdoor recreation activities in Ireland. By integrating data from a number of sources, this paper proposes a methodological approach for predicting the number of visitors to outdoor recreation sites and sets out some preliminary results.

[18] Specifically, the distance adjustment factor was calculated by comparing the Euclidean distance and the actual road distance between a number of cities and towns in Ireland and then taking the simple average of the ratio between the two measures. This represents an initial basic method to address the problem identified by Brainard *et al.* (1999). A more sophisticated approach to dealing with this issue incorporating a 'road network analysis' is currently being undertaken which involves estimating the travel distance between any two points in Ireland via the road network.

A number of elements of the approach have yet to be completed and work is ongoing in relation to them. This includes the linking of spatial datasets to SMILE as well as populating the spatial microsimulation model with information about recreationalists. This represents a key are of ongoing future work. We also intend to refine the methodology in Section 4 by incorporating more extensive data on recreationalists' residential locations as well as undertaking a 'road network analysis' in future modelling.

Other areas for future research include developing tests of the 'convergent validity' of the estimates of total trips as well as the development of a more sophisticated two-stage econometric model. This would involve modelling the total numbers of trips taken by each recreationalist in the first stage and, in the second stage, allocating these trips across recreation sites. We also intend to use these methodologies to assess the impact of exogenous changes in site attributes and to apply them to other outdoor recreation activities such as recreation walking and forest park visits. Overall the long-term aim is to develop methodologies using SMILE to examine the potential for diversifying from agriculture to agri-tourism, in particular outdoor pursuit based activities.

References

Ballas, D., Clarke, G., Dorling, D., Rigby, J. and B. Wheeler (2006), *Using Geographical Information Systems and Spatial Microsimulation for the Analysis of Health Inequalities,* Health Informatics Journal, Vol. 12, No. 1, 65-79.

Ballas, D., Clarke G.P. and E. Wiemers (2005a), *Building a Dynamic Spatial Microsimulation Model for Ireland*, Population, Space and Place (In Press).

Ballas, D., Clarke, G., Dorling, D., Eyre, H., Thomas, B. and D. Rossiter (2005b), *Sim Britain: A Spatial Microsimulation Approach to Population Dynamics*, Population, Space and Place, 11, 13-34.

Ballas, D. and G.P. Clarke (2001a), *Modelling the Local Impacts of National Social Policies: A Spatial Microsimulation Approach*, Environment and Planning C: Government and Policy, 19, pp.587-606.

Ballas, D. and G.P. Clarke (2001b), *Towards Local Implications of Major Job Transformations in the City: A Spatial Microsimulation Approach*, Geographical Analysis 33, 291-311.

Ballas, D. and G.P. Clarke (2000), *GIS and Microsimulation for Local Labour Market Policy Analysis*, Computers, Environment and Urban Systems, 24, 305-330

Bateman, I.J., Jones, A.P., Lovett, A.A., Lake I. and B.H. Day (2004), *Applying Geographical Information Systems (GIS) to Environmental and Resource Economics*, Paper presented at the School of Economics, University of New South Wales, Sydney, Australia, August 6th, 2004.

Birkin, M. and M. Clarke (1988), *SYNTHESIS – A Synthetic Spatial Information System for Urban and Regional Analysis: Methods and Examples*, Environment and Planning A, 20, 1645-1671.

Birkin, M. and M. Clarke (1989), *The Generation of Individual and Household Incomes at the Small Area Level using Synthesis*, Regional Studies, 23, 535-548.

Brainard, J.S., Lovett, A.A. and I.J. Bateman (1999), *Integrating Geographical Information Systems into Travel Cost Analysis and Benefit Transfer*, International Journal of Geographical Information Systems, 13(3): 227-246.

Cameron, A.C. and P.K. Trivedi (1998), *Regression Analysis of Count Data*, Econometric Society Monographs No. 30.

Cullinan, J.E. (2006), *An Overview of Benefit Transfer in Environmental Valuation*, Rural Economy Research Centre Working Paper Series, forthcoming.

Harding, A. (2002), *Using Microsimulation Models in the Policy Process in an Ageing Society*, NATSEM Online Conference Paper - CP2002_001.

Holm, E., Lindgren, U., Makila, K., and G. Malmberg (1996), *Simulating an Entire Nation*, In G.P. Clarke (Ed.), Microsimulation for Urban and Regional Policy Analysis, Pion, London.

Hooimeijer, P. (1996), *A Life-Course Approach to Urban Dynamics: State of the Art in and Research Design for the Netherlands,* In G.P. Clarke (Ed.), Microsimulation for Urban and Regional Policy Analysis. Pion, London.

Hynes S., Hanley, N. and E. Garvey (2005), *Up the Proverbial Creek Without a Paddle, Accounting for Variables Participant Skills Levels in Recreational Demand Modelling*, Working Paper.

King, A. (2003), *The SYNAGI Model – Synthetic Australian Geo-Demographic Information*, Paper presented to International Microsimulation Conference on Population, Ageing and Health: Modelling Our Future, Canberra.

King A., McLellan, J. and R. Lloyd (2002), *Regional Microsimulation for Improved Service Delivery in Australia: Centrelink's CuSP Model*, NATSEM Online Conference Paper - CP2002_009.

Kongmuang (2004), *Modelling Social Determinants of Crime in Leeds*, The 2nd National Crime Mapping Conference, 9-10 March, 2004.

NATSEM (2004), *Spatial Microsimulation Modelling of Care Needs, Costs and the Capacity for Self-Provision: Detailed Regional Projections for Older Australians to 2020*, Conference Paper, Australian Population Association Conference, Canberra, September 2004.

O'Donoghue, C., Lennon, J., Ballas, D. and G. Clarke (2004), *Location Choice Decisions in Ireland*, RERC Working Paper.

Redway, H. (2005), *Data Fusion by Statistical Matching*, Mimeo: Department of Work and Pensions.

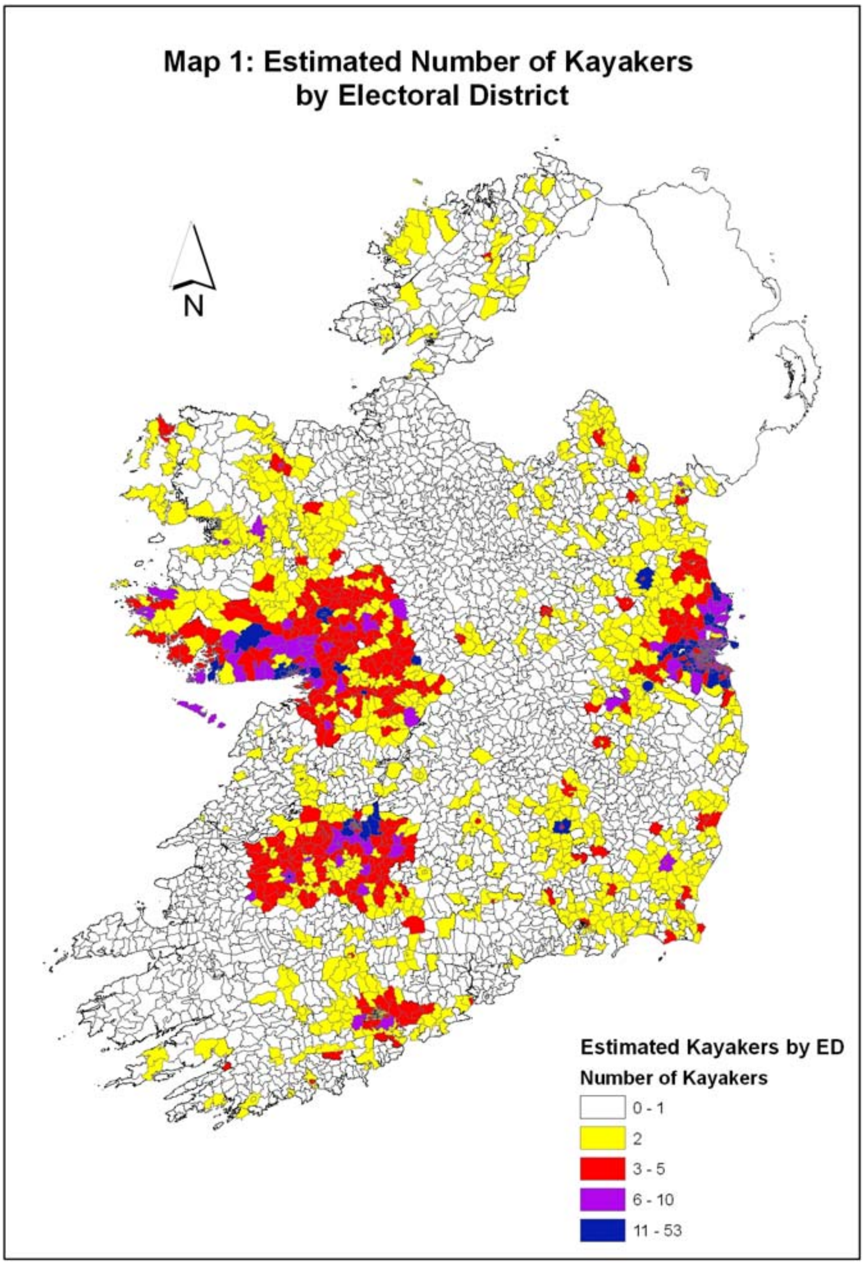
Rosenberger, R.S. and J.B. Loomis (2001), *Benefit Transfer of Outdoor Recreation Use Values: A Technical Document Supporting the Forest Service Strategic Plan (2000 Revision)*, Gen. Tech. Report RMRS-GTR-72, Fort Collins, CO: USDA Forest Service, Rocky Mountain Research Station.

Tiglao N. and M. Tsutsumi (2005), *Modeling Households and Location Choices in Metro Manila Using Spatial Microsimulation Approach*, Proceedings of the Eastern Asia Society for Transportation Studies, Vol. 5, pp. 2179 - 2194, 2005.

Vencatasawmy, C.P., Holm, E. and T. Rephann (1999), *Building a Spatial Microsimulation Model*, Paper presented at the 11th Theoretical and Quantitative Geography European Colloquium, Durham, UK.

Williamson P. (2002), *Synthetic Microdata*, Ch. 17 in Rees, Martin and Williamson [Eds] The Census Data System, Wiley, Chichester, 231-241.

Annex 1: Maps and Tables



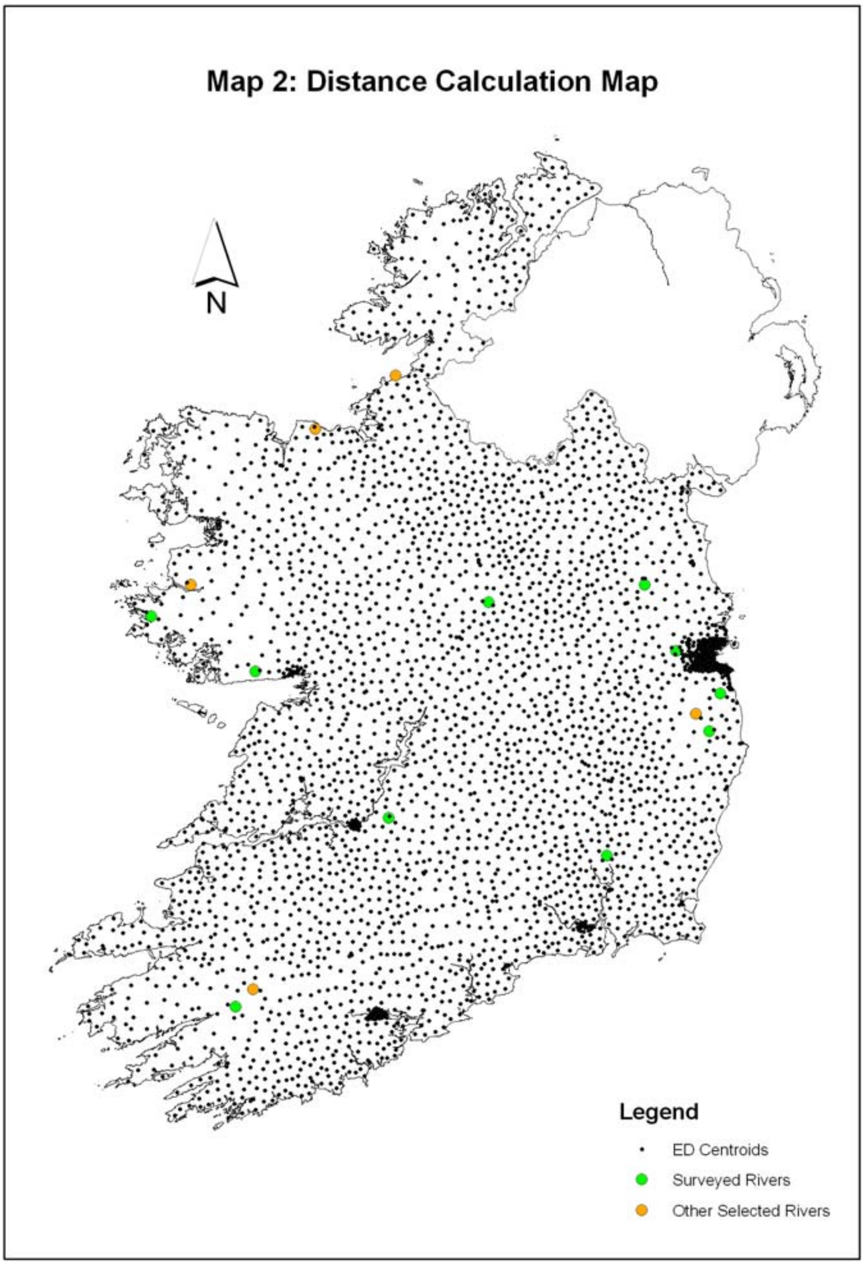

Map 2: Distance Calculation Map
N
Legend
ED Centroids
Surveyed Rivers
Other Selected Rivers

Table 2: Explanatory Variables

D	**Distance Related Variables**
$Dist_j$	Distance to river *j*
$DistSq_j$	Distance squared
S	**Site Attribute Variables**
$Parking_j$	Quality of parking at river *j*
$Crowding_j$	Extent of crowding at river *j*
$Stars_j$	Star rating of river *j*
$Water_j$	Quality of water at river *j*
$Scenery_j$	Quality of scenery at river *j*
$Info_j$	Information available for river *j*
$Other_j$	Availability of substitute sites for river *j*
R	**Variables Relating to Characteristics of Recreationalists**
$Income_i$	Income of kayaker *i*
$IncSq_i$	Income squared
Age_i	Age of kayaker *i*
$YrsPadd_i$	Number of years kayaker *i* has kayaked
$YrsPaddSq_i$	Number of years kayaked squared
$Male_i$	Dummy variable for gender of kayaker
$Married_i$	Dummy variable indicating married status
$Prof2_i$	Dummy variable indicating a level 2 proficiency kayaker
$Prof3_i$	Dummy variable indicating a level 3 proficiency kayaker
$Import_i$	Variable relating to importance of kayaking as a hobby
$ObFree_i$	Number of obligation free days available per year

Estimation – Reduced Form Model[19]

Table 3: Model Estimates – Reduced Form Model[20, 21]

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Dist	-0.007	-0.022	-0.01	-0.01	-0.023	-0.023
	(9.43)**	(8.84)**	(11.81)**	(11.90)**	(13.22)**	(12.67)**
DistSq	-3.7E-05	4.71E-05	-2.3E-05	-2.3E-05	6.4E-05	6.35E-05
	(8.14)**	(3.99)**	(4.96)**	(4.85)**	(7.22)**	(6.88)**
Parking	0.28	0.172	0.312	0.313	0.111	0.118
	(24.93)**	(4.93)**	(24.32)**	(24.31)**	(3.55)**	(3.74)**
Crowding	-0.308	-0.266	-0.294	-0.293	-0.138	-0.126
	(29.55)**	(7.75)**	(23.66)**	(23.46)**	(4.69)**	(4.26)**
Stars	0.054	0.037	0.163	0.166	-0.006	0.022
	(3.70)**	-0.73	(9.76)**	(9.94)**	-0.17	-0.58
Water	-0.41	-0.389	-0.517	-0.521	-0.228	-0.231
	(37.08)**	(9.48)**	(37.61)**	(37.62)**	(6.90)**	(6.89)**
Scenery	0.173	0.083	0.166	0.165	-0.094	-0.101
	(14.10)**	(2.07)*	(11.45)**	(11.34)**	(2.72)**	(2.86)**
Info	0.392	0.344	0.586	0.594	0.328	0.341
	(33.89)**	(9.22)**	(40.67)**	(40.89)**	(9.84)**	(10.15)**
Other	0.234	0.197	0.211	0.21	0.027	0.02
	(27.78)**	(7.03)**	(21.69)**	(21.48)**	-1.05	-0.76
Constant	1.048	2.241	0.614	-	0.606	0.505
	(13.36)**	(8.85)**	(5.57)**	-	(2.68)**	(2.20)*
Ln r	-	-	-	-	0.633**	-
Ln s	-	-	-	-	2.048**	-
Observations	3068	3068	3068	3068	3068	3068
Groups	-	-	279	279	279	279

Absolute value of z statistics in parentheses
* significant at 5%; ** significant at 1%

[19] The reduced form model is given by $T = f(D, S)$, so that the numbers of trips taken is modelled as a function of distance travelled and site attributes. Estimation methods include Poisson and negative binomial maximum likelihood estimation of both pooled multi-site and panel data models in Stata. Post-estimation, we predict $\lambda_{i,j}$, the conditional mean number of trips by individual *i* to site *j*, by in effect 'plugging in' the relevant values for **D** and **S**: $\hat{\lambda}_{i,j} = \exp(\hat{\beta}_0 + \sum_{k=1}^{2} \hat{\beta}_k D_{i,k} + \sum_{l=1}^{7} \hat{\beta}_{l+2} S_j)$.

[20] Model 1 is a pooled multi-site Poisson regression model, while Model 2 represents a pooled multi-site negative binomial model. Models 3 and 4 respectively are Poisson RE and FE panel models while Models 5 and 6 are negative binomial RE and FE panel models. A series of models were estimated and tests for over-dispersion, heterogeneity and fixed effects were all undertaken resulting in a RE panel model (Model 5) as the preferred model.

[21] All models are found to be significant using likelihood ratio (LR) and Wald tests where appropriate.

Estimation – Full Model[22]

Table 4: Model Estimates – Full Model[23]

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Dist	-0.008	-0.024	-0.01	-0.01	-0.023	-0.023
	(10.94)**	(9.78)**	(11.84)**	(11.90)**	(13.38)**	(13.22)**
DistSq	-3E-05	5.38E-05	-2.3E-05	-2.3E-05	6.43E-05	6.52E-05
	(6.63)**	(4.64)**	(4.93)**	(4.85)**	(7.28)**	(7.21)**
Parking	0.264	0.135	0.311	0.313	0.122	0.119
	(23.79)**	(3.83)**	(24.28)**	(24.31)**	(3.98)**	(3.88)**
Crowding	-0.299	-0.249	-0.295	-0.293	-0.128	-0.111
	(28.89)**	(7.30)**	(23.74)**	(23.46)**	(4.46)**	(3.89)**
Stars	0.09	0.002	0.163	0.166	0.018	0.045
	(6.07)**	-0.03	(9.79)**	(9.94)**	-0.48	-1.2
Water	-0.429	-0.383	-0.516	-0.521	-0.219	-0.227
	(37.59)**	(9.82)**	(37.55)**	(37.62)**	(6.79)**	(6.98)**
Scenery	0.164	0.058	0.163	0.165	-0.099	-0.103
	(13.16)**	-1.45	(11.30)**	(11.34)**	(2.91)**	(3.02)**
Info	0.356	0.329	0.585	0.594	0.323	0.344
	(30.59)**	(8.96)**	(40.58)**	(40.89)**	(9.78)**	(10.39)**
Other	0.203	0.156	0.21	0.21	0.038	0.027
	(24.27)**	(5.50)**	(21.62)**	(21.48)**	-1.51	-1.09
Income	1.73E-05	2.16E-05	2.39E-05	-	1.4E-05	4.36E-06
	(10.98)**	(3.95)**	(2.69)**	-	(2.66)**	-0.64
IncSq	0	0	0	-	0	0
	(7.76)**	(2.97)**	-1.62	-	(2.98)**	-1.49
Age	-0.004	-0.032	-0.014	-	-0.015	-0.016
	-1.53	(4.08)**	-1.1	-	-1.84	-1.57
YrsPadd	0.035	0.039	0.034	-	-0.039	-0.054
	(15.23)**	(4.44)**	(2.40)*	-	(4.53)**	(5.08)**
YrsPaddSq	0	0	0	-	0	0
	(15.21)**	(4.68)**	(2.47)*	-	(4.50)**	(4.96)**
Male	-0.248	-0.222	-0.188	-	-0.151	-0.173
	(8.34)**	(2.40)*	-1.28	-	-1.64	-1.39
Married	-0.415	-0.193	-0.252	-	0.162	0.167
	(9.56)**	-1.24	-1.01	-	-1.11	-0.91
Prof2	0.588	0.58	0.544	-	0.534	-0.102
	(9.94)**	(4.17)**	(2.56)*	-	(3.63)**	-0.43
Prof3	0.836	0.967	0.784	-	1.16	0.664
	(13.74)**	(6.32)**	(3.30)**	-	(7.34)**	(2.66)**
Import	-0.148	-0.074	-0.127	-	0.069	0.184
	(10.45)**	(1.99)*	(2.15)*	-	-1.86	(3.59)**
ObFree	0.004	0.002	0.003	-	-0.001	-0.002
	(27.41)**	(3.41)**	(3.74)**	-	-1.47	(3.32)**
Constant	-0.024	2.081	-0.51	-	0.157	0.965
	-0.22	(6.21)**	-1.41	-	-0.49	(2.41)*
Ln r	-	-	-	-	0.693**	-
Ln s	-	-	-	-	2.143**	-
Observations	3068	3068	3068	3068	3068	3068
Groups	-	-	279	279	279	279

Absolute value of z statistics in parentheses
* significant at 5%; ** significant at 1%

[22] In the full model, $\hat{\lambda}_{i,j} = \exp(\hat{\beta}_0 + \sum_{k=1}^{2} \hat{\beta}_k D_{i,k} + \sum_{l=1}^{7} \hat{\beta}_{l+2} S_j + \sum_{m=1}^{11} \hat{\beta}_{m+9} R_{i,m})$.

[23] Again all models are found to be significant using the appropriate LR and Wald tests.

Estimates of Total Trips

Table 5: Estimates of Total Trips – Preferred Reduced Form and Full Models

	Reduced Form Model	**Full Model**
Liffey	61,377	60,425
Clifden	16,241	16,346
Curragower	29,138	28,802
Boyne	23,009	22,257
Roughty	6,998	7,290
Clare Glens	10,028	10,271
Annamoe	22,215	22,940
Barrow	15,522	15,534
Dargle	18,372	19,180
Inny	15,306	15,217
Boluisce	12,544	12,758
Total Trips	230,750	231,021

Table 6: Estimates of Total Trips to Other Rivers – Preferred Reduced Form and Full Models

	Reduced Form Model	**Full Model**
Inchavore	10,899	11,000
Middle Flesk	19,176	20,072
Bun Doracha	13,623	13,503
Easky	4,670	4,913
Duff	4,092	4,043
Total Trips	52,460	53,531