A combinatorial optimisation approach to non-market environmental benefit aggregation

Stephen Hynes

Nick Hanley

Cathal O'Donoghue

Online at http://www.economics.stir.ac.uk

A combinatorial optimisation approach to non-market environmental benefit aggregation

Stephen Hynes[1][†], Nick Hanley[2], Cathal O'Donoghue[1]

1. Environmental Modelling Unit, RERC, Teagasc, Ireland

2. Department of Economics, University of Stirling, Scotland

Abstract

This paper considers the use of spatial microsimulation in the aggregation of regional environmental benefit values. The developed spatial microsimulation model uses simulated annealing to match the Irish Census of Agriculture data to a Contingent Valuation Survey that contains information on Irish farmers' willingness to pay (WTP) to have the corncrake restored as a common sight in the Irish countryside. We then use this matched farm survey and Census information to produce regional and national total WTP figures, and compare these to figures derived using more standard approaches to calculating aggregate environment benefit values. The main advantage of the spatial microsimulation approach for environmental benefit value aggregation is that it allows one to account for the heterogeneity in the target population. Results indicate that the microsimulation modelling approach provides aggregate WTP estimates of a similar magnitude as those produced using the usual sample mean WTP aggregation at the national level, but yields regional aggregate values which are significantly different.

Keywords: Environmental benefit value, aggregation spatial microsimulation, willingness to pay, corncrake conservation.

JEL codes: Q1, Q2, C8.

† Corresponding author: Environmental Modelling Unit, Rural Economy Research Centre, Athenry, Galway, Ireland. Tel: +353(0) 91 845269, Fax: +353(0) 91 844296, email: stephen.hynes@teagasc.ie

1. Introduction

Space is of a heterogeneous nature, with the consequence that environmental costs and benefits have geographical discriminating, distributive impacts (Nijkamp, 2002). This statement is particularly true when one considers the aggregation of welfare estimates derived from non-market environmental valuation techniques. If one attempts to aggregate a nationally representative random sample's average welfare estimates to a particular region of interest it is generally not possible to take account of the heterogeneous characteristics in that region's population and one may therefore under or overestimate the aggregate regional benefit from an environmental policy change. In this context, we employ the principle of synthetic estimation (Williamson et al., 1998), whereby a spatial microsimulation model allows us to take into account the spatial heterogeneity of the target population in the aggregation process.

Consider, as an example, a common Contingent Valuation Method (CVM) survey situation where we have a random sample of 1500 urban households in Ireland where each was asked their willingness to pay (WTP) for improvement in the quality of their local water supply due to the implementation of new filtering technology. This sample of households is representative of the entire national population, but is not representative of each census small area jurisdiction in the country. Now suppose we are particularly interested in the aggregate WTP of residents in one small area jurisdiction, Galway city, for a cleaner water supply. We know from census records that households in Galway city display higher incomes and higher education levels than the average urban dweller in the state. This means that using the average WTP in our sample to aggregate up to the target population may underestimate the aggregate WTP of Galway city for a clean water supply.

Using the spatial microsimulation methodology we can instead select individuals from the CVM survey in order to define a subset synthetic dataset which is more representative of people living in

Galway city. The appropriate number of household records for areas within Galway city are taken at random from the CVM survey; the characteristics of income bracket and education level of these randomly selected (and perhaps duplicated) synthetic households for Galway city are then tabulated and the resulting synthetic income and education level tables compared with the real Census tables for the city, and the error (mismatch) is calculated. Assuming that there is a significant error between the synthetic and real tables, some of the previously-selected households are swapped for an alternative, equivalent number of household records randomly chosen from the CVM survey, and the error recalculated; if the error has substantially increased, that swap is rejected, otherwise the swapped records are retained; swapping continues until a 'best fit' between the synthesized data and the real Census tables is reached.

On completion of the matching process we now have a synthetic population of individuals from the CVM sample representing the target Galway population by income and education levels and we can then use these synthetic households original WTP values to get a truer aggregate estimate for WTP in Galway city that would be possible by multiplying the original samples average WTP by the population number in the jurisdiction of interest (since we are accounting for the fact that Galway city displays different characteristics from the average urban area in the country)[1]. The process outlined above is a combinational optimization technique referred to as simulating annealing and when employed to statistically match census and sample data the process is often referred to as spatial microsimulation.

---

[1] It is important to realize that a single household from the CVM survey may appear multiple times in the simulated synthetic population of a single borough of Galway city and could potentially appear in numerous boroughs in the simulated population of Galway city. This happens for the simple reason that we are only using the sample of 1500 records to produce a simulated population for the Galway city boroughs that is much larger.

The case study used in this paper considers a CVM study that asks Irish farmers their WTP to conserve an endangered farmland bird, the Corncrake (*Crex crex*)[2]. Corncrakes depend on the maintenance of suitable farm habitats, and have been in rapid decline due to changes in farming methods over the last 25 years. The CVM survey (which quantifies the WTP of farmers for a corncrake conservation program) and Census of Agriculture data is then used to produce small area population environmental benefit micro data estimates for the year 2005 using the spatial microsimulation framework outlined above. We contend that the microsimulation method could provide two potential benefits in terms of non-market valuation. Firstly, using the spatial microsimulation modeling framework allows us to efficiently and accurately expand a sample's individual welfare estimates to a particular region of interest or the entire population. Secondly, the methodology allows us to take into account the spatial heterogeneity of the target population in the aggregation process. These claims are investigated in this paper.

In the next section, we briefly review the topic of aggregation of environmental benefit values. Section 3 then describes the design of our WTP survey and discusses the datasets used in the microsimulation process. In section 4 we discuss the spatial microsimulation approach used to aggregate the WTP values for corncrake conservation at both a national and regional level of jurisdiction. This section also reviews the payment card elicitation format used in the CVM study and the associated use of a generalized Tobit interval modeling approach. Model results and the aggregated WTP estimates are presented in section 5. Finally, section 6 concludes with some recommendations for further research.

---

[2] The Corncrake is the only Irish breeding bird which is currently threatened with global extinction. It has been listed on the 2005 International Union for the Conservation of Nature (IUCN) Red List of Threatened, due to population and range declines of more than 50% in the last 25 years. According to the All Ireland Action Plan for endangered species in Ireland (NPWS, 2005) the main targets for the corncrake are an increase in the population in the three core areas in the Republic of Ireland to 150 in Donegal, 50 in West Connacht and 60 in the Shannon Callows by 2010 and to re-establish breeding populations in other parts of its former range, by 2015.

2.   The Aggregation of Environmental Benefit Values

Aggregating environmental benefit values is the process whereby sample mean values of WTP or other welfare measures are converted to a total value figure for the population (Hanley et al., 2003). Bateman et al. (2006) point out that because the methods for measuring non-market benefit values are based on analyses of individual behavior, there is a problem in knowing how changes in a resource will affect aggregate values, since aggregation will depend on both the benefits per person and the population of beneficiaries (the extent of the market). Indeed, Smith (1993) and Bateman (2000, 2006) argue that the extent of the market may be more important in determining aggregate values than any changes related to the precision of the estimates of per-person values. Using the spatial microsimulation methodology we demonstrate how it is possible to define the extent of the market by statistically matching respondents in a valuation survey to census data.

A number of other issues in the literature regarding the aggregation of environmental value estimates can also be resolved by using the spatial microsimulation approach developed in this paper. At a very basic level, Loomis (1987) states that the problem of generalizing results from a sample to the population relates to low response rates and small sample sizes. By statistically matching our sample of farmers and their associated WTP values with associated farm characteristics obtained from a census, we can generate a representative population with individual WTP values for the entire farming population. Our spatial microsimulation approach also alleviates the problem highlighted by Morrison (2000) in relation to how representative the sample of respondents, in any public good valuation study, is of the actual socioeconomic and demographic characteristics of the population in question.

Another concern relates to spatial representativeness in the aggregation process (Bateman et al., 2006). This issue is something that is also relevant for our study. Different parts of Ireland are represented by different types of farmers. The western seaboard for example is predominately

represented by relatively small extensively operated, livestock farmers while the south east of the country is populated by larger, more intensive dairy and tillage farm holdings. In any aggregation process for this particular population it is vital that these spatial differences in farm size and type is taken into account, especially if we wish to examine regional variations in the total benefit value (to the farming community) of the corncrake conservation program which forms the empirical focus of this paper.

Another issue that needs consideration when dealing with environmental benefit aggregation is aggregation error. Aggregation errors arise when estimates from a sample are aggregated to represent the total population value for a particular public good. These errors are inversely related to the degree of correspondence between the sample and the population (Rosenberger and Stanley, 2006). Calculating the extent of the error when aggregating up to the total population is a difficult prospect, but ensuring the correspondence of socio-economic characteristics between the sample and population (as is done in the spatial microsimulation approach developed here) should increase the accuracy of the aggregation. In the discussion at the end of the paper we also propose a follow-up field study that would facilitate a valid test of aggregation error where our estimates of total WTP at the Electoral Division (ED) level of geographical area are compared to the actual total WTP of farmers in the corresponding EDs, as revealed by the interviewing of all farmers in each of the EDs.

The present paper conducts what is, to our knowledge, the first systematic aggregation of contingent valuation data that accounts for heterogeneity in the target population using spatial microsimulation techniques. We contrast the aggregation approach used with a number of alternative procedures which are common in the literature. Our results show that the choice of aggregation approach can have a major impact upon derived estimates of total benefits at a

regional level, especially when the target population displays a large amount of heterogeneity across space.

3.   Data and WTP Survey format

In this section we describe the data used in this paper and the format of the willingness to pay questions. *The National Farm Survey (NFS)* is collected as part of the Farm Accountancy Data Network of the European Union. In the 2005 NFS, additional questions were asked in terms of farmers' willingness to pay towards the restoration of the corncrake in the Irish countryside. A pilot sample was used to inform general survey design and to gauge the likely range of farmers' willingness to pay in order to inform the bid design of the main survey. In carrying out the main survey each interviewee was told about the current population of the corncrake and how its numbers have fallen over the last 20 years. The farmers were also informed that …*"Bird Watch Ireland has operated an intensive Corncrake Conservation Project in Ireland since 1991, with the support of the Department of Environment, Heritage and Local Government and the Royal Society for the Protection of Birds"*. The farmers were then informed that *..."As the population of corncrakes increases and spreads across the country, their management and maintenance will impose additional costs on the funding bodies, local authorities and local landowners (restrictions in land use) compared to the status quo of no restoration program. This cost would have to be paid for by the general public so it is important to find out how much if anything YOU would be willing to pay to have the corncrake restored as a common sight in the Irish countryside"*.

The farmers interviewed were then asked if they were willing to pay something towards the restoration of the corncrake into the Irish countryside and the maintenance of a sustainable population of corncrakes into the future. The farmers were instructed to bear in mind their total annual budget, the amount they might allocate to wildlife conservation and finally how much of this they could afford to spend on this restoration program. Also, they were told to bear in mind

7

that paying too much for this restoration program may mean that they could not afford other worthwhile wildlife conservation schemes. Respondents answering "No" to this question were then asked which of several statements best described why they were not willing to pay anything. Those who answered the question in the affirmative were then presented with a payment card showing the bid amounts of €10, €20, €30, €40, €50 and €60 and were asked: "of these bid amounts which would be the maximum you would be willing to pay (€) each year into a conservation fund to aid in the restoration of this bird and bring the singing male population back up to a sustainable population of 900 birds".

The Payment Card Method (Cameron and Huppert, 1989) was chosen given the data collection method being used. Fifteen separate recorders collect the NFS on the individual farms annually. Given that the farmers are asked over 300 questions in these surveys, it was necessary to choose a simple approach to the WTP questions on the survey to avoid question-answering fatigue on the part of the respondents. As with any of the response formats in a CVM study, the use of the payment card method has advantages and disadvantages. Some of its best documented advantages are that it can provide a context to the bids and avoids "yea-saying" where some respondents answer yes to any single bid amount presented to them (Blamey et al. 1999). It can also help avoid starting point bias and may reduce the number of outliers in the sample (Boyle et al. 1996). Finally, the payment card method may also reduce the problem of respondents saying that they would pay high bid amounts that exceed their true values (Boyle et al. 1997). Some of the method's most documented disadvantages are that it can be subject to biases associated with the range of bids used on the card, and that some respondents will choose the first or last number in a sequence. It has also been pointed out that the method may lack incentive compatibility. Boyle (2003, p141) notes that "the literature does not support the choice of a single-response format (dichotomous choice) and it does not exclude the use of payment-card and multiple-bounded questions".

8

A total of 1117 surveys were collected. 47 of these were unusable due to the fact that the recorder did not collect any information on the WTP questions in the NFS. A total of 453 individuals responded that they would be willing to pay something towards a corncrake conservation program. However, 46 of these said they were not willing to pay even the lowest bid value presented to them on the payment card (€10). Of the remaining €0 WTP responses, 33 were treated as protest bids due to the fact that the respondents stated that they were not willing to pay anything because either they felt the payment vehicle was not appropriate or they could not give a legitimate reason why they were WTP €0 (also see footnote 4). These observations were excluded from the analysis. The total final number of usable responses was 928.

The other dataset used in this paper is the Census of Agriculture. The objective of the Census is to identify every operational farm in the country and collect data on agricultural activities undertaken on them (CSO, 2002). The census classifies farms by physical size, type and geographical location. In order to estimate the aggregate environmental value of a biodiversity conservation program to Irish farmers, we statistically match the CV responses within the NFS with the Census of Agriculture to create an attribute-rich synthetic dataset with information on the willingness to pay of every farmer in Ireland, including the Electoral Division (ED) where they are located.

4.   Methodology

Microsimulation models have been increasingly adopted to study the impacts of social and economic policies (Ballas et al., 2005). However, this is the first application of the spatial microsimulation methodology to the valuation of public goods. In the context of the research presented here we generate synthetic farm population microdata sets at the Electoral Division (ED) level for Ireland. We have a combinatorial optimisation problem where we try to find the set of NFS farms that best match the Census of Agriculture Small Area Populations (SAPs) tables of

farm size, farm system and soil type. These SAPs tables simply indicate the number of farms by size, system and soil type in each ED. These variables are also believed *a priori* to be useful in predicting farmers' WTP for corncrake conservation.

To formalise our combinatorial optimisation problem consider a pair (*R, E*), where *R* is the finite set of farm configurations (set of NFS farm records representing the number of farms in an ED by size, system and soil type) and *E* is an error function ($E : R \rightarrow \mathbf{R}$), which assigns a real number to each farm configuration. *E* is defined such that the lower the value of *E*, the better the corresponding configuration of NFS farms represents the Census SAPs tables. The problem then is to find the configuration of farms for which *E* takes its minimum value, i.e. an optimum configuration i_o satisfying:

$$E_{opt} = E(i_o) = \min_{i \in R} E(i),$$

where E_{opt} denotes the minimum error between the actual census tables of size, system and soil type and the simulated tables constructed using the configuration of NFS farms and the objective is to find the set of NFS farm records that produces this minimum error value in each ED. In order to solve this combinatorial optimisation problem we employ what is defined as an approximation algorithm which yields an approximate solution in an acceptable amount of computation time (Wu and Wang, 1998). The simulating annealing (SA) algorithm is one such general approximation algorithm. SA is used to locate a good approximation to the global optimum of a given function in a large search space using randomization techniques. The SA algorithm used in this paper was adopted from the algorithm employed by Ballas et al. (2005), where the authors generated a synthetic urban population in Leeds, UK to analysis urban planning issues[3] and the mathematical model of the algorithm is described fully in Laarhoven and Aarts (1987, chapter 2).

---

[3] We implement the SA algorithm in Java, an object-oriented programming language, which has been accepted as the most suitable type of programming language for spatial microsimulation modelling (Ballas and Clarke, 2000; Wu and Wang, 1998).

The simulated annealing process selects a set of farms from the 928 records of the NFS that best fits the Census SAP tables of (i) Farm Size in hectares; (ii) Farm System and (iii) Soil Class for every ED in the country. The simulated annealing procedure can be described as follows. We initially choose a configuration (*i*) of NFS farms to represent the SAP tables for a single ED. Given configuration *i*, another configuration *j* can be obtained by randomly selecting a number of records in configuration *i* and replaced them with ones chosen at random from the universe of NFS records. The number of records to be replaced is defined as *T*. In the first iteration T equals half the number of farms in an ED. Letting $\Delta E_{ij} = E(j) - E(i)$, then the probability that configuration *j* will be the next configuration of farms in a predefined sequence of configurations (the java program sets the number of iterations) is given by 1, if $\Delta E_{ij} < 0$ and by $\exp\left(-\Delta E_{ij} / T\right)$ if $\Delta E_{ij} > 0$. The acceptance of a new configuration is decided by drawing random numbers from a uniform distribution on [0, 1] and comparing these with $\exp\left(-\Delta E_{ij} / T\right)$. This process continues, with T being lowered at each step, either until the maximum number of iterations has been hit or the error falls within the desired setting[4].

The spatial microsimulation process is complete when the selection of farms from the NFS can simultaneously reproduce the census SAPS tables for the number of farms by size, system and soil type contained in the Census of Agriculture with less than 5% of a difference between the original SAPS tables and those generated from the NFS selection. Once this point is reached the programme stores the simulated set of NFS farm records for that ED and repeats the process to find the set of NFS farms that best fits the Census SAP tables for the next ED and so on. Matching the NFS and the SAPS data creates synthetic demographic, socio-economic and farm level variables, such as marital status, age, fertiliser usage, livestock units per farm, etc and most

---

[4] The static model also employs a restart method. When a restart occurs the simulated annealing process begins again with a new sample of records. The restart is used so that more farm combinations can be explored. The restart method is applied if the model fails to find a satisfactory solution within the maximum permitted iterations.

importantly from our research, predicted WTP values *for each farmer in the population*. The simulating annealing process conduced for this research produces 145,057 individual farm records.

The Payment Card Estimation Procedure

On completion of the microsimulation model it was important to be able to compare alternative environmental benefits aggregation methods, such as simple mean WTP aggregation in the NFS sample (or summation in the microsimulated farm population) and a value function approach. We use the generalized Tobit model in this comparison as the value function approach. As previously mentioned, the elicitation format chosen in this study was the Payment Card Method where each farmer was shown a payment card listing various euro amounts and asked to indicate the maximum amount they were willing to pay. Following Cameron and Huppert (1989), the response is interpreted not as an exact statement of willingness to pay but rather as an indication that the WTP lies somewhere between the chosen value and the next larger value above it on the payment card. Table 1 displays the distribution of the usable responses in the farm survey across the intervals. The price range used in this study was based on the responses to a pilot study (which utilised the open-ended elicitation format (see Haab and McConnell, 2002)).

The WTP responses were treated in a parametric model, where the WTP value chosen by each farmer[5] was specified as: WTP = $\mu + \varepsilon$. It is assumed that $\varepsilon \sim N(0, \sigma^2 I)$. This is a generalised Tobit model and is estimated via maximum likelihood procedures. Daniels and Rospabé (2005) provide a log-likelihood function adjusted to make provision for point, left-censored, right-censored (top WTP category with only a lower bound) and interval data. For farmers $j \in C$, we observe WTP_j, i.e. point data and for farmers $j \in L$, WTP_j are left censored. Farmers $j \in R$ are right censored; we know only that the unobserved WTP_j is greater than or equal to WTP_{Rj}. Finally

---

5 Farmers were initially focused on because any corncrake conservation program will only succeed if it has the support of Irish farmers, given that they manage the permanent grassland which is home to the corncrake. It would however be interesting to extend the CVM survey to the general population to calculate the aggregate WTP for the entire population of Ireland.

farmers $j \in I$ are intervals; we know only that the unobserved WTP_j is in the interval $[WTP_{1j}, WTP_{2j}]$. The log likelihood is then given by:

$$\ln L = -\frac{1}{2}\sum_{j \in C} w_j \left\{ \left(\frac{WTP_j - x\beta}{\sigma} \right) + \log 2\pi\sigma^2 \right\} + \sum_{j \in L} w_j \log \Phi \left\{ \left(\frac{WTP_{Lj} - x\beta}{\sigma} \right) \right\}$$

$$+ \sum_{j \in R} w_j \log \left\{ 1 - \Phi \left(\frac{WTP_{Rj} - x\beta}{\sigma} \right) \right\} + \sum_{j \in I} w_j \log \left\{ \Phi \left(\frac{WTP_{2j} - x\beta}{\sigma} \right) - \Phi \left(\frac{WTP_{1j} - x\beta}{\sigma} \right) \right\}$$

Where $\Phi()$ is the standard cumulative normal and w_j is the weight of the jth farmer. Table 1 presents the distribution of WTP by censorship type. Of the 928 usable responses, a total of 538 zero WTP values were treated as $j \in C$ [6]. 4 WTP values were considered right censored at €60 while the remaining 386 were treated as interval observations.

5.   Results

As Ballas et al. (2001) point out; one of the biggest drawbacks of microsimulation models is the difficulty in validating the model outputs. This is due to the fact that microsimulation models estimate distributions of variables which were previously unknown. The model used in this paper uses three different statistics to assess (internally) the models goodness-of-fit: total absolute error, relative error and z-scores (Kelly, 2004). As well as these internal validation measures we can also validate the synthetic microdata estimates produced by the spatial microsimulation model by re-aggregating the model results up to levels at which observed data sets exist (Irish Central Statistics Office (CSO) figures) and compare the estimated distributions with the observed. Table 2 presents a comparison of other summary statistics for both the NFS and our microsimulated population. The validation of our model results using the internal validation techniques and the reaggregation

---

[6] Those individuals who said they were not willing to pay anything for the conservation program and gave the reason that they 1.didn't like this bird or 2. felt the government should pay from existing revenues or 3. that the bird would be a nuisance to production or 4 couldn't simply afford to pay were considered as a point data observations of €0. Those 46 individuals who said they were not willing to pay anything for the conservation program and gave the reason that the price was to high were considered as interval data observations of between €0 and €10.

process are not reproduced here but for the interested reader they are discussed fully in Hynes et al. (2008).

The main goal of the microsimulation exercise carried out in this paper was to be able to analyze the aggregate value of the proposed corncrake conservation project to the Irish farming community at different levels of aggregation across space. As can be seen from figures 1 and 2 this can be done at a number of different levels including ED, county and regions (defined at the EU "NUTS III" level). Of the 3440 EDs in the country 2850 contain farms; the average number of farms in each of these EDs being 53 (min 10, max 320).

It is very evident from the map (figure 2), produced with our microsimualted farm population estimates, that farmers in EDs found in the west, south west and border areas of the country seem to be willing to pay higher amounts on average into a conservation fund to have the corncrake restored into the Irish countryside. This is a very interesting finding given that the remaining singing male population of corncrakes in Ireland is largely restricted to four areas, Co. Fermanagh (which is on the border on the Northern Ireland side), Donegal, west Connacht, and the Shannon Callows (Schäffer and Green, 2001). The WTP values in our microsimulated population of farmers are thus positively correlated spatially with areas where the corncrake can still be found and highlights what Bateman (2006) refers to as an 'ownership' dimension to aggregate benefit values.

In order to compare our microsimulation estimate of aggregate WTP to other more traditional approaches of aggregation used in the literature we calculate the aggregate environmental value of the corncrake conservation program in 4 alternative ways. These are:

1. The simple multiplication of the average value of the stated (maximum) WTP in the NFS sample by the number of farms in the country or county ($\sum_{i=1}^{n} \overline{WTP}_{NFS}$).

2. Aggregation using the CVM interval regression model outlined in section 3 (the value function approach) where the estimated average value of WTP in the NFS sample is multiplied by the number of farms in the country ($\sum_{i=1}^{n} \overline{\widehat{WTP}}_{NFS}$).

3. The summation of the stated (maximum) WTP for each farm i in the simulated farm population ($\sum_{i=1}^{n} WTP_{SIMi}$) and

4. Aggregation using a CVM interval regression model but in this case applied to the simulated farm population where the estimated values of WTP in the synthetic population for each farm i are summed to calculate total WTP at the desired level of aggregation ($\sum_{i=1}^{n} \widehat{WTP}_{SIMi}$).

The parametric regression results of the value function approach calculated using the 2005 NFS sample (weighted using the individual farm population weights provided in the NFS) and the microsimulated farm population are presented in Table 3. It can be seen from the results that WTP increased significantly with the income generated on the farm. The "REPS farm" variable indicates that farmers participating in the Rural Environment Protection scheme (REPS)[7] are willing to pay (significantly) higher amounts than those farmers not participating in the scheme. Given the environmental education component involved in the uptake of this scheme and the fact that farmers participating in an agri-environmental scheme are more likely to favour a biodiversity conservation program, this finding is not surprising. The Organic Nitrogen Production per hectare variable is an indicator of how intensive the farming enterprise is. As expected for both models, farms with higher rates of organic nitrogen per hectare are willing to pay significantly less for a corncrake conservation program.

---

7 The Rural Environment Protection Scheme (REPS) was introduced in Ireland under EU Council Regulation 2078/92 in order to encourage farmers to carry out their activities in a more extensive and environmentally friendly manner. Approximately 43,000 farmers were actively participating in the scheme in 2005.

The one major difference between the coefficients in the models is the fact that the "size" coefficient is negative in the NFS model and positive in the microsimulated model. It is however also found to be insignificant in the NFS model. It may be that a higher number of the larger farms with higher WTP values in the NFS are matched to each ED in the simulating annealing process than the larger farms with lower WTP in the NFS sample when the microsimulated farm population is generated. The Log Likelihood χ^2 statistic shows that, taken jointly, the coefficients in both the NFS and spatial microsimulated models are significant at the 1% level. The interval based models produce average WTP values that are significantly higher that the average stated maximum WTP values in the sample and microsimulated farm population[8].

The NFS generalized Tobit model produces a lower value for the average WTP per farm than the value generated from the microsimulation generalized Tobit model. It is also worth noting from table 4 that the associated confidence intervals are non-overlapping. In contrast to this, the difference between the $\overline{WTP}_{NFS}$ and $\overline{WTP}_{SIMi}$ (average WTP column, rows 1 and 2 in table 4) would appear not to be statistically different. In relation to the aggregation of the WTP values it can be seen from table 4 that at the national level of aggregation[9], the figures are similar in magnitude and not significantly different when comparing $\sum_{i=1}^{n} WTP_{SIMi}$ and the simple mean WTP aggregation approach for the NFS ($\sum_{i=1}^{n} \overline{WTP}_{NFS}$). In terms of the aggregation of the estimated WTP values however, $\sum_{i=1}^{n} \hat{W}TP_{SIMi}$ produces an estimate that is significantly smaller in terms of non-overlapping confidence intervals to $\sum_{i=1}^{n} \overline{\hat{WTP}}_{NFS}$. In absolute terms, the difference is 20%.

---

8 In all cases we have made the assumption that the preferences of farmers are stable across space. If it was assumed that this is not the case and that there are additional area-specific effects the WTP function could be amended using NUTSIII regional dummies that are available in the NFS.
9 For the national aggregation, n, the total number of farms is equal to 145,057 (CSO, 2002)

Similar to the water quality example discussed in the introduction we can use our microsimulated farm population to examine the aggregate WTP of particular regions within the country while taking into account the regional variation in farming activity. To this end we choose to examine 3 extensive farming counties and 3 intensive farming counties as defined in the Census Atlas of Agriculture in the Republic of Ireland (Lafferty et al. 1999). The results of this regional aggregation are displayed and compared to the other aggregation options in table 5. In terms of the regional aggregation of the estimated WTP values, $\sum_{i=1}^{n} \hat{WTP}_{SIMi}$ produces regional aggregate estimates that are significantly lower for the *intensive* farming counties (on average 7%) compared to $\sum_{i=1}^{n} \overline{\hat{WTP}}_{NFS}$. On the other hand, for the *extensive* farm counties, $\sum_{i=1}^{n} \hat{WTP}_{SIMi}$ produces regional aggregate estimates that are significantly higher (on average 12%) compared to $\sum_{i=1}^{n} \overline{\hat{WTP}}_{NFS}$. A similar pattern is evident when comparing the regional estimated values of $\sum_{i=1}^{n} \overline{\hat{WTP}}_{NFS}$ and $\sum_{i=1}^{n} WTP_{SIMi}$. By not recognizing the differences in farming activity across Irish rural space through the use of the microsimulated population, we would be substantially overestimating the regional aggregation of WTP for intensively farmed areas and underestimating the regional aggregation of WTP for extensively farmed areas.

6.   Discussion and Conclusions

As discussed in section 2, there are numerous examples in the literature where reliance upon sample means will fail to yield an accurate measure of aggregate WTP. As an alternative, we propose a new approach based upon the estimation of a spatial microsimulation model which takes into account the impact of variation in the characteristics of the relevant aggregation population and allows for the calculation of regional as well as total environmental benefit values. The comparison between aggregate estimates in table 4 (for national total WTP) and table 5 (for county

total WTP) demonstrated that the microsimulation modeling approach provides similar estimates of aggregate environmental value as the simple sample mean WTP aggregation approach at the national level but resulted in regional values (where there is substantial heterogeneity in the target farm populations) which were significantly different when assessing total WTP values at the county level. The results ultimately demonstrate that researchers failing to take account of the regional heterogeneity in their study population may be introducing biases in their attempts to estimate regional environmental aggregate benefit values.

We would speculate that there are a number of benefits of using spatial microsimulation methods to create synthetic microdata for use in aggregating welfare estimates. Firstly, using the spatial microsimulation modeling framework allows us to efficiently and accurately aggregate a sample's individual welfare estimates to a particular region of interest or the entire population. Secondly, the methodology allows us to take into account the spatial heterogeneity of the target population in the aggregation process. Finally, it should also be noted that the creation of the microsimulated population of farmers with their accompanying WTP values is not a technique that is unique to the datasets in this study but is an approach that should in theory be able to be replicated and used with large sample data sets in other CVM studies and indeed with revealed preference techniques such as the recreation travel cost method and the hedonic price valuation technique. Also, a number of SA algorithms are now available to download free off the internet[10] and as Hynes et al. (2008) point out, once the matching datasets are structured in a manner that can be read by the programming language being employed, the synthetic data can be produced without to much "reinvented the wheel" on the part of the researcher.

In terms of future research it would be interesting to investigate whether the habitat type representing the main breeding ground of the corncrake was positively associated with the average

---

10 See for example http://webscripts.softpedia.com/script/Miscellaneous/General-simulated-annealing-algorithm-34475.html.

WTP values per farm for corncrake conservation in each ED. This could be accomplished by combining our simulated spatial WTP values with land cover data in a GIS framework. One further area for further research in terms of using microsimulation techniques in public good valuation studies would be a "ground truthing" exercise to examine whether the WTP estimates in our microsimulated population are statistically equivalent to what one would find if the WTP questionnaire was to be conducted in each ED. This would facilitate a valid test for aggregation error for our estimate of total WTP in our microsimulated population. To this end it would be worth while to pick a number of EDs around the country and survey the farmers in them to see how close the actual WTP values of the farmers in the chosen EDs are to the estimates for those farmers in the corresponding simulated ED populations. Given that our microsimulated population is constrained to statistically match the census of agriculture tables, and the fact that our aggregation method meets the criteria discussed in the literature for the production of more reliable aggregate welfare estimates, (Bateman et al. 2000; Rosenberger and Stanley, 2006)[11] we would expect the aggregation error to be relatively small.

It has been claimed previously by Van Pelt (1993) that environmental policy has to be regional specific in light of distributional issues and site and population specific attributes. The results of this paper continue to support that viewpoint. Similarly, Nijkamp (2002) contended that progress at the interface of regional and environmental economics is contingent on the availability of proper spatial information systems and models. We believe that we have offered a new perspective for analyzing this linkage between space and the environment where land use and heterogeneous spatial behaviour are shown to be closely connected to alternative regional aggregate environmental benefit values.

---

11 The criteria that may affect the accuracy of aggregation include the quality of the population sample data, the methods used in modeling and interpreting the sample data, the analysts' judgments regarding research design and implementation and the closeness between the sample and the relevant population.

References

Ballas, D. and Clarke, G. 2000. "GIS and microsimulation for local labour market analysis." *Computers, Environment and Urban Systems* 24: 305-330

Ballas, D., Clarke, G. and Dorling, D. 2005. *SimBritain: A Spatial Microsimulation Approach to Population Dynamics*, Rowntree Publishers, York.

Bateman, I., Day, B., Georgiou, S. and Lake, I. 2006. "The Aggregation of Environmental Benefit Values: Welfare Measures, Distance Decay and Total WTP." *Ecological Economics*, 60: 450–460

Bateman, I., Jones, A. Nishikawa, N. and Brouwer, R. 2000. "Benefits transfer in theory and practice: A review and some new studies." available from http://www.uea.ac.uk/~e089/, CSERGE and School of Environmental Sciences, University of East Anglia.

Blamey, R., Bennett, J. and Morrison M. 1999. "Yea-saying in contingent valuation surveys." *Land Economics* 75, 126-141.

Boyle, K., Johnson, F. and McCollum, D. 1997. "Anchoring and adjustment in single-bounded, contingent-valuation questions." *American Journal of Agricultural Economics*, 79, 5: 1495-1500

Boyle, K., Johnson, F., McCollum, D., Desvousges, W., Dunford, R. and Hudson, S. 1996. "Valuing public goods: Discrete versus continuous contingent-valuation responses." *Land Economics*, 72, 3: 381-396

Boyle, K. 2003. "Contingent valuation in practice", in Champ, P., Boyle, K., and Brown, T. *A Primer on Nonmarket Valuation*. Kluwer Academic Publishers.

Cameron, T., and Huppert, D. 1989. "OLS versus ML estimation of non-market resource values with payment card interval data." *Journal of Environmental Economics and Management*, 17, 230-246.

Central Statistics Office (CSO), 2002. *The census of agriculture - main results*. Central Statistics Office, Cork.

Daniels, R. and S. Rospabe, 2005. "Estimating an Earnings Function from Coarsened Data by an Interval Censored Regression Procedure". Development Policy Research Unit Working Paper 05/91.

Goffe, W., Ferrier, G. and Rogers, J. (1992). "Global optimization of statistical functions: Preliminary results", in Amman, H., Belsley, D. and Pau, L. *Computational Economics and Econometrics*, 19–32.

Haab, T. and McConnell, K. 2002. *Valuing Environmental and Natural Resources*, Edward Elgar

Hanley, N., Schläpfer, F., Spurgeon, J. 2003. "Aggregating the benefits of environmental improvements: Distance-decay functions for use and non-use values." *Journal of Environmental Management* 68, 297–304.

Hynes, S., Morrissey, K. and O'Donoghue, C. 2008. "Building a Static Farm Level Spatial Microsimulation Model for Rural Development and Agricultural Policy Analysis in Ireland", *International Journal of Environmental Technology and Management*, forthcoming.

Kelly, D. (2004), *SMILE Static Simulator Software User Manual Teagasc*, Teagasc Publication, October 2004

Laarhoven, van, P. and Aarts, E. 1987. *Simulating Annealing: Theory and Applications*. Kluwer Academic Publishers, London.

Lafferty, S., Commins, P. and Walsh, J. 1999. *A census atlas of agriculture in the Republic of Ireland.* Dublin: Teagasc Publication.

Loomis, J. 1987. "Expanding contingent value sample estimates to aggregate benefit estimates: Current practices and proposed solutions." *Land Economics*, 63:4, 396-402

Morrison, M. 2000. "Aggregation Biases in Stated Preference Studies." *Australian Economic Papers,* 39 (2), 215–230.

National Parks & Wildlife Service (NPWS), Department of the Environment, Heritage and Local Government Ireland, 2005. *All-Ireland Species Action Plans for the Irish Hare, the Corncrake, the Pollan and Irish Lady's Tresses* (www.npws.ie)

Nijkamp, P. 2002. "Space in Environmental Economics" in J. van den Bergh (ed.), *Handbook of Environmental and Resource Economics*, Edward Elgar Publishers, Cheltenham, UK.

Pelt, van, M., 1993. *Ecological Sustainability and Project Appraisal, Aldershot*, UK: Avebury.

Rosenberger, R. and Stanley T. 2006. "Measurement, generalization, and publication: Sources of error in benefit transfers and their management." *Ecological Economics* 12 (2):372-378.

Schäffer, N. and R. Green, 2001. "The Global Status of the Corncrake." *RSPB Conservation Review* 13, 18-24.

Smith, V. 1993. "Nonmarket Valuation of Environmental Resources: An Interpretive Appraisal." *Land Economics,* 69 (1), 1–26.

Williamson, P., Birkin, M. and Rees, P. 1998. "The Estimation of Population Microdata by Using Data from Small Area Statistics and Samples of Anonymised Records." *Environment and Planning A,* 30, 785-816.

Wu L. and Wang, Y. 1998. "An Introduction to Simulated Annealing Algorithms for the Computation of Economic Equilibrium." Computational Economics, 12, 151-16

Figures

Figure 1. Average WTP for a Corncrake Conservation Programme per Farm per ED

Average WTP per Electoral Division
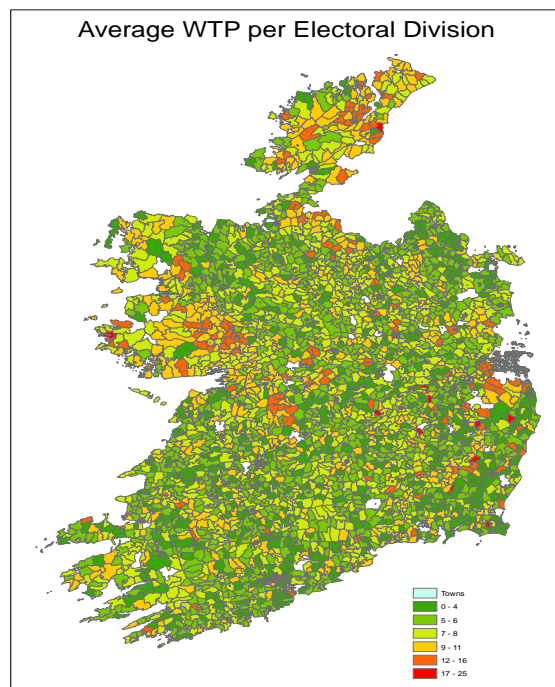
Figure 2. Total Environmental Value of a Corncrake Conservation Programme at Alternative levels of Spatial Aggregation

Total WTP per NUTSIII Region

Total WTP per County
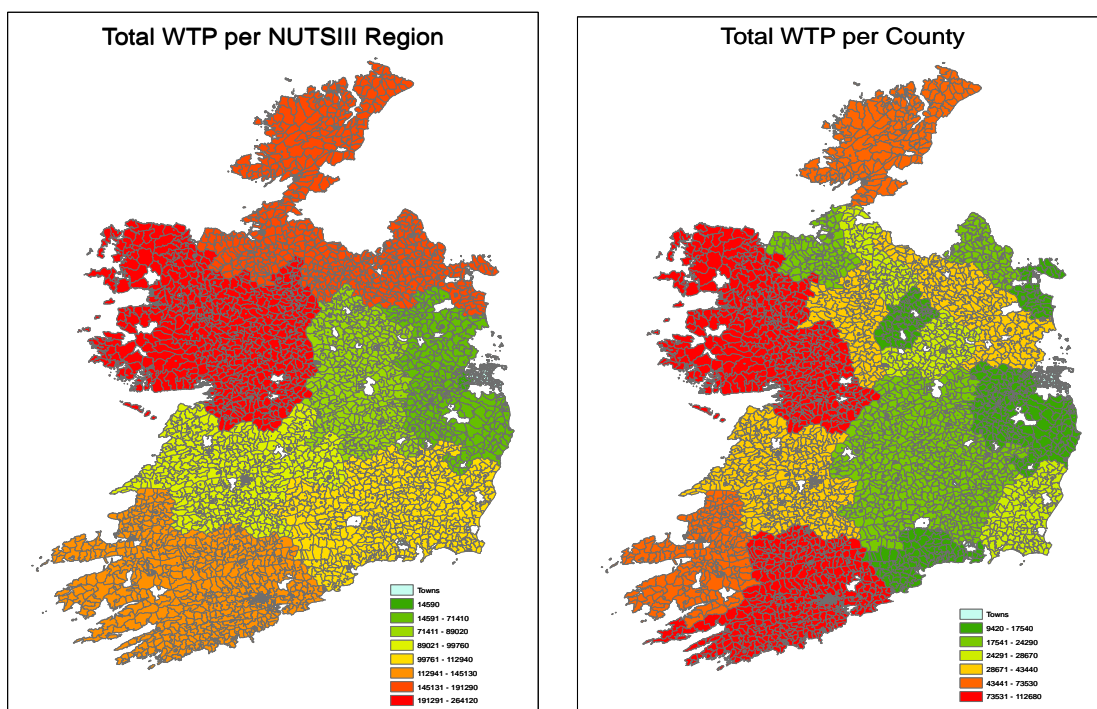
Tables

Table 1. Summary Statistics of the WTP Intervals in NFS Sample

Interval	Frequency	Percent	Cumulative Percent
0 (point value	538	57.97	57.97
0-10	46	4.96	62.93
10 - 20	102	10.99	73.92
20 - 30	129	13.9	87.82
30 - 40	40	4.31	92.13
40 - 50	20	2.16	94.29
50 - 60	49	5.28	99.57
60+	4	0.43	100
Total	928	100	

Table 2. Summary Statistics of the NFS and the Microsimulated Farm Population

	National Farm Survey Sample 1,177 Observations		Microsimulated Farm Population 145,057 Observations	
Variable	*Mean*	*Standard Deviation*	*Mean*	*Standard Deviation*
Size of Farm (acre)	37.28	32.93	30.93	26.27
Crop Pasture (acre)	83.17	71.22	72.53	61.13
Gross margin (€)	38980.89	40937.45	35039.79	37645.17
Farm income (€)	22456.92	24618.09	20026.95	22417.42
Grossoutput (€)	55465.31	59268.50	50421.83	54912.12
REPS payment (€)	2386.04	3393.09	1892.79	2959.51
Age (years)	53.95	12.71	54.34	12.83

Table 3. Interval Regression of WTP for Corncrake Conservation for NFS Sample and for the Microsimulated farm population

Variable	NFS Model	Microsimulation Model
Size of Farm (acres)	-0.0246 (-0.03)	0.0300 (0.004)***
Family Farm Income (€/1000)	0.0736 (0.03)**	0.0491 (0.003)***
Age of Farm Operator	0.05 (0.05)	0.109 (0.003)***
Organic Nitrogen Production (kg/hectare)	-0.0394 (-0.02)***	-0.0215 (-0.001)***
REPS farm^	2.112 (1.26)*	3.660 (-0.07)***
Total crops and pasture (acreage)	-0.00694 (-0.02)	-0.0362 (-0.002)***
Constant	10.31 (-3.30)***	3.752 (0.19)***
Log of the estimated standard error	2.723 (-0.001)***	2.598(-0.001)***
Log likelihood	-274362	-464498
Likelihood Ratio χ^2 (6) test	18	6253
Left Censored Observations	0	0
Right Censored Observations	4	532
Uncensored Observations	538	92772
Interval Observations	386	51753

Standard error in parentheses.* significant at 10%; ** significant at 5%; *** significant at 1%. ^REPS farm indicates that the farmer participates in the Rural Environment Protection scheme (REPS)

Table 4. WTP estimates for the 4 alternative estimation methods

Method of Analysis	Average WTP Per Farm (€)	Total environmental value of a corncrake conservation program (€)
NFS Max stated WTP	9.07 (8.19, 9.94)	985,200 (889,613; 1,079,701)
Microsimulation Farm Population Max stated WTP	6.79 (6.73, 6.85)	984,937 (976,234; 993,640)
Payment Card Interval Regression for NFS sample	10.55 (10.41, 10.69)	1,530,351 (1,510,043; 1,550,659)
Payment Card Interval Regression for microsimulated farm pop.	8.36 (8.35, 8.37)	1,213,307 (1,211,856; 1,214,758)

95% confidence Intervals in brackets

Table 5. Total WTP estimates per County for the 4 alternative aggregation methods

County	$\sum_{i=1}^{n} \overline{WTP}_{NFS}$	$\sum_{i=1}^{n} \overline{\hat{WTP}}_{NFS}$	$\sum_{i=1}^{n} WTP_{SIMi}$	$\sum_{i=1}^{n} \hat{W}TP_{SIMi}$
Extensive Farming Counties				
Donegal	59031	68405	73556	77137
Galway	92736	107463	112637	121266
Mayo	102341	118592	107758	131001
Intensive Farming Counties				
Cork	106702	123647	94338	115591
Tipperary S.	26443	30642	23455	28405
Waterford	19146	22186	17530	20510

Note that *n* in this case represents the total number of farms in each *county*. All values in €.