Enhancing Britain's rivers: an interdisciplinary analysis of selected issues arising from implementation of the Water Framework Directive

Submitted by Danyel Ian Hampson to the University of Exeter

as a thesis for the degree of

Doctor of Philosophy in Economics

in December 2016

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University

Signature: ……………………………………………………

Abstract

The Water Framework Directive requires reduced environmental impacts from human activities and for the assessment of the non-market benefits of pollution remediation schemes. This policy shift has exacerbated the research problems surrounding the physical, social and economic consequences of the relationship between land use and water quality. This research seeks to quantify the major socio-economic and environmental benefits for people which may arise as riverine pollution is reduced. To achieve these aims this research integrates primary data analyses combining choice experiment techniques with geographical information system based analyses of secondary data concerning the spatial distributions of riverine pollution.

Current knowledge on the microbial quality of river water, measured by faecal indicator organism (FIO) concentrations and assessed at catchment scale, is inadequate. This research develops generic regression models to predict base- and high-flow faecal coliform (FC) and enterococci (EN) concentrations, using land cover and population (human and livestock) variables. The resulting models are then used both to predict FIO concentrations in unmonitored watercourses and to evaluate the likely impacts of different land use scenarios, enabling insights into the optimal locations and cost-effective mix of implementation strategies.

Valuation experiments frequently conflate respondents' preferences for different aspects of water quality. This analysis uses stated preference techniques to disaggregate the values of recreation and ecological attributes of water quality, thereby allowing decision makers to better understand the consequences of adopting alternative investment strategies which favour either ecological, recreational or a mix of benefits. The results reveal heterogeneous preferences across society; specifically, latent class analysis identifies three distinct groups, holding significantly different preferences for water quality.

From a methodological perspective this research greatly enhances the ongoing synthesis of geographic and economic social sciences and addresses important policy questions which are of interest to a variety of stakeholders, including government departments and the water industry.

Contents

List of Tables

List of Figures

List of equations

Author's declaration: a statement of the author's (and contributing authors') contributions

Danyel Hampson's PhD studentship was created in part to facilitate collaboration between Catchment hydrology, Resources, Economics and Management (ChREAM) project partners. Due to the highly interdisciplinary nature of the research presented within this thesis sections of this thesis are based upon joint research. The approximate ratios of authorship are shown in Table 1.

Table 1: approximate contributions of the research within this thesis

	Authorship (percentage)		Approximate word count
	Danyel Hampson	Other contributors	
Chapter 1	55	45	14,100
Chapter 2	90	10	15,300
Chapter 3	95	5	36,500

Due to the close collaboration between ChREAM contributors, several of the datasets used within the context of this research had already been generated for other aspects of ChREAM land use and hydrological modelling. Authorship of the datasets used to generate the generic models (reported in chapter one) and used to model riverine faecal indicator organism (FIO) concentrations in the Humber River Basin District (RBD) (reported in chapter two) are indicated in Table 2.

Dr Paulette Posen developed the methodologies for calculating land use within hydrological response units (HRUs) and for interpolating Agricultural Survey data to HRUs and, subsequently, generated livestock population and land use datasets originally for use within ChREAM land use and hydrological modelling. These datasets were reapplied within this research. Humber catchment boundaries were generated by the Centre for Ecology and Hydrology (CEH) for ChREAM land use and hydrological modelling within the Humber RBD and used within this research.

Table 2: authorship of the datasets used within the FIO modelling

	Original author				
	Danyel Hampson	**Paulette Posen**	**Centre for Ecology and Hydrology**	**Centre for Research into Environment and Health**	**Total**
CREH catchment datasets (Underpinning the Generation of the Generic FIO Models)					
Land Use Profiles	9	6			15
Human Population Profiles	15				15
Livestock Population Profiles	15				15
Standard Percentage Runoff (SPR) Profiles	15				15
Soil Temperature	15				15
Catchment and HRU Boundaries	3			12	15
Historic FIO enumeration				15	15
Historic Hydrological Run-off				15	15
Humber catchment datasets (Used in the Transfer, Scenario and QMRA modelling)					
Land Use Profiles	1	17			18
Human Population Profiles	18				18
Livestock Population Profiles	1	17			18
Catchment and HRU Boundaries			18		18

Centre for Research into Environment and Health (CREH) datasets, assembled from 15 process based catchment studies, comprise digitised catchment boundary data, an empirical database of enumerated FIO concentrations and accurate runoff and discharge data for catchments. Datasets relating to historic FIO enumeration and historic hydrological runoff for CREH catchments are commercially sensitive and were not released from CREH premises for statistical analysis by Danyel Hampson. Datasets were also subject to the restrictions outlined within the *Agreement between Professor Dave Kay (CREH) and Professor Ian Bateman (UEA) concerning collaborative work under the ChREAM project* (see Appendix I)[1].

Professor David Kay, the Director of CREH, directed all research relating to the historic CREH catchment studies and, with Professor Ian Bateman, supervised the construction of the generic models reported in Chapter 1. Datasets for

[1] See, in particular, "Condition 2. At no point will any data, model parameters or any information be published or otherwise disseminated which could allow a third party to operationalise the FIO model," and, "Condition 3. Instead it is foreseen that publications will focus upon summary statistics and mapped outputs from the model and relationships described in terms of directional responses rather than quantified parameters." Although the quantified parameters are not published within this thesis (or within peer reviewed journal articles), they were scrutinized by the examiners of this PhD at viva.

independent variables (land use, human and livestock populations, rainfall Standard Percentage Runoff (SPR), and soil temperature) (Table 2) were generated by Danyel Hampson and Dr Paulette Posen and supplied to Dr John Crowther of CREH, who then developed the generic step-wise multiple regression models used to predict FIO concentrations.

Professor Ian Bateman directed the ChREAM economic analysis and, with Dr Carlo Fezzi, developed the econometric models underpinning the fertilizer tax, milk quota and Environmentally Sensitive Areas (ESA) scenarios used within ChREAM land use and hydrological modelling. The datasets for those scenarios were reapplied within this research. Dr Philip Jones developed the Land Use Allocation Model (LUAM) and, with Dr Joseph Tzanopoulos, generated the dataset for the Healthy Diet scenario using the LUAM. That research was originally undertaken to fulfil an entirely separate research project (Jones and Tranter, 2008). Dr Jones readily accepted an invitation by Danyel Hampson to have his dataset assessed using the generic FIO models, to estimate the impact on riverine FIO concentrations that adoption of the Healthy Diet regime would have. Dr Tzanopoulos increased the resolution of the Healthy Diet dataset to $10km^2$ and that grid square dataset was subsequently interpolated to HRUs by Danyel Hampson. Authorship of the datasets used in the scenario modelling (undertaken in chapter two) is summarised in Table 3.

Table 3: authorship of datasets used within the FIO scenario modelling

<table>
<tr><th>Author</th><th>Scenario type</th></tr>
<tr><td rowspan="5">Danyel Hampson</td><td>Livestock population reduction</td></tr>
<tr><td>Human population increase</td></tr>
<tr><td>Mixed scenario (20% decrease in dairy livestock population, 1.4% increase in human population and 5% improvement in wastewater treatment efficiency)</td></tr>
<tr><td>Fertilizer reduction</td></tr>
<tr><td>Streamside fencing</td></tr>
<tr><td rowspan="3">Ian Bateman and Carlo Fezzi</td><td>Fertilizer tax</td></tr>
<tr><td>Milk quota</td></tr>
<tr><td>Environmentally Sensitive Area</td></tr>
<tr><td>Philip Jones and Joseph Tzanopoulos</td><td>Nutritionally Driven Healthy Diet</td></tr>
</table>

Within the two chapters relating to FIO modelling, Danyel Hampson coordinated the research efforts between researchers based at the University of East Anglia, CREH and the University of Reading; developed the methodologies for interpolating human population data, LUAM Healthy Diet data, soil temperature data and SPR data to HRUs; produced catchment boundary, land use, livestock population, human population, soil temperature and SPR datasets of CREH catchments which were then used to develop the generic FIO models; produced catchment boundary, land use, livestock population and human population datasets of Humber catchments to facilitate the transfer of the FIO models to the Humber RBD; generated the Humber transfer function Predictor Variable Matrix (PVM); operationalised the FIO models' transfer to the Humber River Basin District and/or subcatchments using the PVM; interpolated Healthy Diet data to HRUs; generated datasets for fertilizer reduction, livestock reduction, human population increase, mixed effects and streamside fencing scenarios; modelled all Humber FIO input change scenarios; drafted, coordinated and acted as corresponding author for the two published papers arising from the FIO modelling research contained within the first two chapters of this thesis (see Appendix II). Danyel's contributions towards those two papers was c.90% and c.40%, respectively. Whilst having no direct input into the current research, Dr Carl Stapleton and Dr Mark Wyer were involved in the historic CREH catchment projects from which data were used to generate the generic FIO models. They were credited on the published papers accordingly.

Danyel was responsible for all of the research reported in the third chapter, with the following exception: Professor Dan Rigby, a ChREAM project partner, developed the choice experiment model using NGene (ChoiceMetrics PTY Ltd, 2012) so that it would be compatible with the choice experiments used within the ChREAM project.

Acknowledgements

Funding for this research was provided by SEER (the Social and Environmental Economic Research project, funded by the ESRC; Ref: RES-060-25-0063) and ChREAM (the Catchment hydrology, Resources, Economics and Management project, funded by the joint ESRC, BBSRC and NERC Rural Economy and Land programme; Ref: RES-227-25-0024, website: http://www.uea.ac.uk/env/cserge/research/relu). ESRC support for Danyel Hampson [grant number ES/F023693/1] is gratefully acknowledged. The FIO and discharge data which underpin this paper were generated during projects funded by Research Councils (EPSRC Grant Reference M99774), the European Union (Interreg IIa, IIIa and IIIb Programmes), the Environment Agency for England and Wales, the Scottish Environment Protection Agency, the Scottish Government, United Kingdom Water Industry Research and individual water companies, including West of Scotland Water (now Scottish Water), Anglian Water, United Utilities, Welsh Water Dŵr Cymru and Yorkshire Water. The grid square agricultural census data, as converted by EDINA at Edinburgh University Data Library, are derived from data obtained for recognised geographies from the Scottish Executive Environment and Rural Affairs Department (SEERAD), the Department of Environment, Food and Rural Affairs (Defra), and the Welsh Assembly, and are covered by © Crown Copyright. Census output is Crown copyright and is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland (Source: 2001 Census; Table KS001). This work is based on data provided through EDINA UKBORDERS with the support of the ESRC and JISC and uses boundary material which is copyright of the Crown. The Ordnance Survey OS Meridian 2 data set is © Crown Copyright 2008–an Ordnance Survey/EDINA supplied service. Copyright of the Land Cover Map 2000 is owned by the Natural Environment Research Council (NERC). All rights reserved. CEH LCM2000 data is only used within the context of this project. HOST soil data are based on digital spatial data licensed from the Centre for Ecology and Hydrology, © NERC, © Cranfield University and © Macaulay Land Use Research Institute. The LUAM modelling was funded by the joint ESRC, BBSRC and NERC Rural Economy and Land programme (reference RES-224-25-0073; website: http://www.relu.rdg.ac.uk/). This research is fully compliant with the ESRC *Research Ethics Framework* (2007).

My sincerest thanks go to the following people, without whom this work would not have been possible: Professors Ian Bateman, Dave Kay and Andrew Lovett, my supervisors, for providing expert guidance throughout, Dr Mike Hutchins, for invaluable hydrological advice and especially Dr Paulette Posen, Dr Silvia Ferrini and Professor Dan Rigby for their advice and assistance with methodological issues. Thanks also go to Dr John Crowther for producing simple, yet effective, FIO models and assisting on all aspects of CREH research.

Keywords

agro-environmental policy

Catchment hydrology, Resources, Economics and Management (ChREAM)

conditional logit model

ecological water quality

faecal indicator organisms

Land Use Allocation Model

latent class model

microbial source apportionment

microbial water quality

nutrition driven food policy

Quantitative Microbial Risk Assessment

recreational water quality

Water Framework Directive

water quality modelling

willingness to pay

Abbreviations and acronyms

AIC	Akaike Information Criteria
base-flow	river flow under low rainfall conditions
BIC	Bayesian Information Criteria
BMP	best management practice
CAIC	Consistent Akaike Information Criteria
CBA	cost benefit analysis
CE	choice experiment
CEH	Centre for Ecology and Hydrology
CFU	colony-forming unit
ChREAM	Catchment hydrology, Resources, Economics and Management
CL	conditional logit
CNL	cross-nested logit
Coef.	coefficient
CREH	Centre for Research into Environment and Health
CSO	combined sewerage overflows
CVM	contingent valuation method
Defra	Department for Environment, Food and Rural Affairs
DLUA	Developed Land Use Area
EA	Environment Agency
EC	*Escherichia coli*
ECL	error component logit
EN	intestinal enterococci
ESA	Environmentally Sensitive Area
EU	European Union
FC	faecal coliform
FIO	faecal indicator organism
GAMS	General Algebraic Modelling System
GIS	geographical information system
GM	geometric mean
high-flow	river flow under high rainfall conditions
HOST	Hydrology of Soil Type
HRU	Hydrological Response Unit (subdivisions within subcatchments)
HSPF	Hydrological Simulation Program Fortran

IIA	independence of irrelevant alternatives
LC	latent class analysis
LCM	land cover map
LM	Lagrange multiplier
LUAM	Land Use Allocation Model
MAFF	Ministry of Agriculture, Fisheries and Food
MXL	mixed logit
NERC	Natural Environment Research Council
non-visitor	A respondent who did not visit the survey river in the year prior to the choice experiment survey
OA	Output Area
ONS	Office for National Statistics
OS	Ordnance Survey
Prob.	probability
PVM	Predictor Variable Matrix
QMRA	Quantitative Microbial Risk Assessment
RELU	Rural Economy and Land Use
RBD	River Basin District
rBWD	revised Bathing Water Directive
RPL	random parameter logit
RUT	random utility theory
s.e.	standard error
SIMCAT	Simulated Catchments
SNIFFER	Scotland and Northern Ireland Forum for Environmental Research
SPR	Standard Percentage Runoff
SSSI	Site of Special Scientific Interest
SWAT	Soil and Water Assessment Tool
SWD	Shellfish Water Directive
SOA	Super Output Area
TC	travel cost
UKRCT	United Kingdom Randomized Controlled Trials
WFD	Water Framework Directive
WHO	World Health Organisation
WTA	willingness to accept

WTP	willingness to pay
WWT	waste water treatment
WwTW	wastewater treatment works

Introduction

Research imperative

The European Union (EU) Water Framework Directive (WFD) (EU, 2000) and its daughter directives, the revised Bathing Waters Directive (rBWD) (EU, 2006a) and the Shellfish Waters Directive (EU, 2006b), set out rules to halt deterioration in the status of EU water bodies. The WFD was adopted by the EU on 22 October 2000 with the requirement that the 'good ecological status' of groundwater and surface waters (rivers, lakes, transitional waters, and coastal waters) be achieved by December 2015[2]. 'Ecological status' is an expression of the quality of the structure and functioning of aquatic ecosystems associated with surface waters (classified in accordance with guideline parameters contained within Annex V of the WFD).

The policy framework replaced piecemeal EU and national legislation and introduced novel features to protect and enhance aquatic ecosystems, including: the establishment of planning and management of waterbodies at river basin district level; shifting emphasis from the assessment of water quality solely in terms of traditional limit-value approaches to chemical composition, to a more holistic assessment of the quality of the biological community, the hydrological characteristics and chemical characteristics of waterbodies; encouraging sustainable use of a renewable resource and widening public participation in water resource planning (Environment Directorate General, 2005).

The WFD was (and still is) extremely ambitious in its scope. For example, integrated river basin planning was adopted to facilitate characterisation and assessment of impacts on river basin districts, environmental monitoring, the setting of environmental objectives and the design and implementation of the programme of measures needed to achieve them (JNCC, 2010). The challenges of implementing the WFD were (and still are) immense, and, unsurprisingly, the December 2015 target for achieving 'good ecological status' was not met. In 2012, it was estimated that 53% of EU waterbodies were either 'good' or would

[2]Transposition into UK law occurred through the following regulations: The Water Environment (Water Framework Directive) (England and Wales) Regulations 2003 (Statutory Instrument 2003 No. 3242) for England and Wales; the Water Environment and Water Services (Scotland) Act 2003 (WEWS Act) and The Water Environment (Water Framework Directive) Regulations (Northern Ireland) 2003 (Statutory Rule 2003 No. 544) for Northern Ireland.

potentially be good (European Commission, 2012). In the UK the situation was worse: in 2011, 37% of UK surface waterbodies were assessed under the WFD as being at least of 'good' status. In 2016, the proportion had fallen to 35% (JNCC, 2016). Catchments are complex systems, which vary widely across the UK depending on geology, weather patterns and land use. This complexity makes it difficult to identify and monitor pollution sources, particularly in large catchments containing urban, industrial and agricultural activities. Initial classifications of UK water quality data were largely based on 'best available knowledge': it is likely that the decline to 35% can, in part, be attributed to the collection of additional data, rather than an actual decline in water quality (Houses of Parliament Parliamentary Office of Science and Technology, 2014).

We see that the policy shifts from assessing water in terms of its chemical composition in favour of its ecological quality has exacerbated the research problems surrounding the physical, social and economic consequences of modified relationships between land use and water quality (Kay et al., 2007a). Consequently, there are gaps in the research literature at all stages.

Policy now requires that the microbial quality of river water, measured by faecal indicator organism (FIO) concentrations, is to be assessed at catchment scale, rather than the previous method of point-source effluent quality regulation. This change encompasses the quantification and management of diffuse sources of microbial pollution derived from the farming community in addition to urban point sources (Kay et al., 2006a). Riverine microbiological water quality has not been systematically measured as part of previous regulatory monitoring programmes. Consequently, modelling of microbial water quality has been sporadic and is underdeveloped, particularly at catchment scale (Kay et al., 2007a).

One important motivation for the implementation of the WFD appears to be the creation of non-market social benefits such as improved provision of, and opportunities for, open-access recreation (Articles 4, 9 and 11 of the WFD) (EU, 2006). It is well reported that improvements in the quality of river water raises non-market benefit values, particularly in terms of increased opportunities for recreational use, but valuation experiments frequently fail to distinguish between respondents' preferences for the ecological or recreational aspects of water quality improvements, so that their trade-off is poorly understood (Bateman et al., 2011; Ferrini et al., 2014).

This thesis addresses these related themes within three substantive chapters. The first two chapters describe the construction and application of transferable models capable of accurately predicting microbial pollution concentrations in UK rivers, thus enabling insights into the optimal locations and cost-effective mix of implementation strategies for the delivery of WFD induced environmental improvements. The third chapter seeks to quantify the major socio-economic and environmental benefits for people which may arise if ecological and/or microbial river pollutants are reduced and considers the distributional and equity implications of alternate strategies for WFD implementation, including the spatial differentiation of policies.

Aims and objectives of the thesis

The following is a summary of the aims and objectives of this research. Detailed descriptions of the iterative steps required to fulfil these aims are provided in each chapter. While undertaking exploratory research it is natural to encounter methodological difficulties and identify potential refinements or entirely new research questions. Where this has been the case, the issues are discussed fully within the relevant chapter.

Some of the most successful catchment-scale FIO modelling has been undertaken using linear regression techniques to model relationships between GM FIO concentrations recorded at monitored sites and land use within their catchments, using variables such as the proportions of grassland and built-up land as proxies for key sources of faecal pollution. Such work has been primarily based on individual catchment studies (Crowther et al., 2003; Kay et al., 2005a). The aim of the first chapter is to extend this latter approach by (i) investigating whether improved models, that can predict base- and high-flow FIO concentrations across the UK, might be achieved by augmenting the predictor variables to include both direct measures of the key FIO sources (i.e., human population and livestock density data) and factors that may affect source strength and the mobilisation, transport, die-off and sedimentation of FIOs within catchments (e.g., volume of runoff, soil hydrology and catchment size); and (ii) assess the extent to which models developed by combining data from discrete UK catchment studies, sampled at different times and under different antecedent weather conditions, are truly generic and transferable across the UK.

The second chapter utilises the models produced in Chapter 1 to develop the regression modelling approach. The strengths and weaknesses of the generic models are assessed, using a range of land use and population change scenarios. The aims are to assess the effectiveness of the FIO models as (i) cost effective diagnostic tools capable of aiding source apportionment, (ii) assess water quality in terms of EU rBWD and WHO microbial water quality assessment categories, (iii) assess the relative effectiveness of a selection of microbial pollution remediation strategies, by quantifying the likely impact on riverine FIO concentrations following the implementation of land use policy measures designed to reduce livestock stocking densities (iv) at a range of spatial scales.

UK valuation studies typically assess WFD benefits in ways that conflate the value of ecological improvements with the value of microbial pollution reduction, therefore assessing water quality as a single attribute of preference (Bateman et al, 2011; Glenk et al., 2014). The main aim of the research in Chapter 3 is to further the knowledge on non-market valuation of river water by disaggregating the values people derive from ecological and microbial aspects of river water quality. Given the link between microbial quality and recreational river use, an investigation is conducted into how the values for these distinct attributes of river water quality differ over people who (i) engage with the river in different ways (rowers, swimmers, anglers) and (ii) who live at different distances from the river. The investigation is undertaken using a stated preference, forced choice discrete choice experiment (DCE), with discrete attributes for ecological and recreational water characteristics.

To achieve the above aims this research uses innovative integrated primary data analyses combining socio-economic survey techniques with geographical information system based analyses of secondary data concerning the spatial distribution of water resources, pollution emission sources and population characteristics. As such, it is a contemporary and highly interdisciplinary approach combining natural and social sciences.

Thesis overview

The substantive chapters refer to diverse research literatures and utilise disparate research methodologies. For this reason, the chapters are deliberately self-contained: each contain detailed reviews of the relevant research literature and each provides the details of the research methodologies used within that chapter. This introduction continues with a summary of the three chapters, followed by a statement of the novelty and contribution to knowledge of the thesis.

The research field of catchment microbial dynamics has been rapidly expanding due to the adoption of the WFD and its paradigm shift towards the integrated management of recreational water quality through the development of drainage basin-wide programmes of measures (Kay et al., 2006a). However, this has led to significant data gaps, due in part to a lack of funding. To meet WFD requirements, data are needed on FIO concentrations in rivers to enable the more heavily polluted to be targeted for remedial action. But due to the paucity of FIO data for UK rivers, especially under high-flow hydrograph event conditions, there is an urgent need by the policy community for generic models that can accurately predict riverine FIO concentrations, and thus inform integrated catchment management programmes.

Chapter 1 reports the development of regression models to predict base- and high-flow faecal coliform (FC) and enterococci (EN) concentrations for 153 monitoring points across 14 UK catchments, using land cover, population (human and livestock density) and other variables that may affect FIO source strength, transport and die-off. This research brings innovation to riverine faecal indicator modelling in several ways. Firstly, it remodelled existing catchment studies using consistent data, enabling a meta-analysis which successfully developed transferable predictive models, based on land use type. These models offer levels of explanation consistent with comparable research (e.g. Kay et al., 2005a). Secondly, although previous modelling has traditionally relied on land use variables as proxies for sources of faecal pollution, this research successfully pioneers the application of human population density and livestock population density as explanatory variables. These models offer superior levels of explanation of the sources of riverine FIO pollution and they represent the first transferable generic FIO models to be developed for the UK which incorporate direct measures of key FIO sources (namely human and livestock population data) as predictor variables. Statistically significant models are developed for both FC and EN, with greater explained variance achieved in the high-flow models. Both land cover and, in particular, population variables are significant predictors of FIO concentrations; with r^2 maxima for EN of 0.571 within the land cover model and 0.624 within the population model. The third innovation is the development of a transfer methodology, enabling the models to predict for actual, or simulated, levels of human or livestock population in hydrological catchments. This allows the generic models to be applied, with confidence, to other UK catchments, both to predict FIO concentrations in unmonitored watercourses and also to evaluate the likely impact of different land use/stocking level and human population change scenarios.

Chapter 2 begins by outlining the application of the FIO models to quantify geometric mean (GM) FC and EN concentrations for base- and high-flow during the summer bathing season in the (largely unmonitored) Humber River Basin District, and provides Quantitative Microbial Risk Assessment estimates for the same, using World Health Organisation (WHO) and EU water quality guidelines.

Because the FIO models incorporate explanatory variables which allow the effects of policy measures which influence livestock stocking rates to be assessed, empirical analyses are then made of the differential effects of seven land use management and policy instruments (fiscal constraint, production constraint, cost intervention, area intervention, demand-side constraint, input constraint, and micro-level land use management), all of which can be used to reduce riverine FIO concentrations. These assessments are conducted at a range of spatial scales, e.g. River Basin District, sub-catchment and at the level of individual hydrological response units (HRU). These analyses provide insights into FIO source apportionment and the spatial differentiation of land use policies which could be implemented to deliver river quality improvements. All of the policy tools are estimated to reduce FIO concentrations in rivers but this research suggests that the installation of streamside fencing in intensive milk producing areas may be the single most effective land management strategy to reduce riverine microbial pollution.

The benefits transfer method is advantageous in that it can be used to estimate economic values for environmental services by transferring available information (e.g. welfare estimates) from studies that have already been completed into other locations or contexts. Cost-benefit analysis in decision-making frequently makes use of benefit transfer for several reasons; it may be too expensive to undertake a full survey to collect primary data at the new site and there may be too little time to conduct an original valuation study, yet some measure of benefits is required. The research within chapters 1 and 2 is not an economic transfer (although it shares equivalent motivations), instead those chapters report the construction and application of transferable models that can be used to predict FIO concentration in unmonitored watercourses. The methodology employed in their construction shares many of the underlying principles of benefit transfer. The FIO models are underpinned by empirical primary data which is combined with readily available secondary data (i.e. land cover and population (human and livestock) variables). The resulting models are then applied to secondary data to predict FIO concentrations in unmonitored watercourses and evaluate the likely impacts of different land use scenarios.

The transferable FIO models reported within this thesis do share some of the disadvantages of economic benefit transfer models. There are limitations in scientific soundness: the transfer estimates are only as good as the methodology and assumptions employed in the original studies. It is advantageous that the procedures used by CREH for collecting primary data (i.e. water samples and enumerating FIO concentrations) from catchment studies were standardised. This primary data, coupled with standardised land use data surfaces and population profiles for catchments, has enabled a consistent meta-analysis of FIO data. However, the CREH datasets underpinning the models may be limited in terms of their temporal availability, catchment characteristics and geographic range. These factors (and other idiosyncrasies of the transfer methodology) may cause transfer errors and reduce the models' relevance to new contexts. Such limitations are discussed in greater detail within chapters 1 and 2.

Due to the commercially sensitive nature of the datasets used to construct the FIO model, the research covered in the first two chapters is subject to the data restrictions detailed in Appendix I. These restrictions prevent the FIO models' parameters from being published, as this could enable the unauthorised operationalisation of the FIO models. Consequently, this thesis focuses upon the mapped outputs from the models and the relationships between the sources of FIOs and riverine FIO concentrations are described in terms of directional responses rather than quantified parameters.

Assessments of potential WFD investments (e.g. for disproportionate cost extensions) typically require that non-market environmental benefits, such as improved ecological quality and greater opportunities for open-access river recreation are assessed within economic cost-benefit frameworks (Görlach and Pielen, 2007; Eftec, 2011). Recent UK valuation studies tend to assess WFD benefits in ways which conflate ecological improvements with the value of recreational gains (see for example Bateman et al. (2011), Doherty et al. (2014), Hanley et al. (2005, 2006) and Metcalfe et al. (2012)). The research within the third chapter seeks to disentangle these sources of value, thereby allowing decision makers to understand the consequences of adopting alternative investment strategies which favour either ecological, recreational or a mix of benefits. This is both feasible and necessary since, contrary to common conception, these facets of water quality can be uncorrelated (Haygarth et al., 2005).

The analysis used stated preference discrete choice experiment (DCE) methods to disaggregate the values respondents hold for recreational and ecological characteristics of river water quality. Face-to-face surveys presented respondents with choices across a range of future water quality scenarios for the River Yare in Norfolk, differentiated in terms of the river's ecological and recreational quality attributes and hypothetical remediation costs. The CE featured a D-efficient experimental design[3] (Ferrini and Scarpa, 2007) that differentiated between respondents' socio-economic and trip characteristics, and, of those respondents who identified as recreational users, captured a wide variety

[3] The combination of attributes and levels used on choice cards was derived following the D-efficient design strategy. D error is the determinant of the variance covariance matrix of the conditional model and is directly linked to parameters precision. D error was 0.306965. An alternative efficiency measure is the A error, which considers the trace of the variance-covariance method. A error in this design was 1.631721. The parameters precision is higher when these efficiency measures are closer to 0.

of recreational activities (e.g. rowing, swimming and angling). Both conditional logit (CL) and latent class (LC) analyses identified a number of significant preference predictors, including respondents' spatial relationships to the river and their socio-economic characteristics.

The willingness to pay estimates derived from the CL modelling revealed clear differences in preferences between respondent types. The non-specialised respondents (i.e. bankside visitors and non-visitors from the general public) hold higher values for improved ecological quality, rather than recreational enhancements. Similar preference orderings, but at higher levels of willingness to pay, were revealed by anglers. However, other specialised users, such as swimmers and rowers, prioritise recreational over ecological improvements. Other preference predictors were identified, including a clear distance decay in values away from the sites of any proposed investment.

LC modelling confirmed three significant classes of respondents and two significant socio-economic variables (number of environmental memberships held by the respondent and distance from the river) helping to define class membership. The results reveal heterogeneous preferences across the respondents; a majority preferring ecological over recreational improvements, while a substantial minority hold opposing preference orderings. The analysis also revealed a third group who have relatively low values for either form of river quality enhancement. Post-estimation results predict the likelihood of respondents' class membership and help identify class members' socio-economic characteristics.

Results were found to be stable over the alternative choice models estimated, confirming the significant heterogeneity in water quality preferences identified across the different groups. As such the research demonstrates that the non-market benefits which may accrue from different types of water quality improvements are nuanced in terms of their environmental impacts, their potential beneficiaries and, by inference, their overall value and policy implications.

The topic of this thesis was conceived before the UK government's decision to hold a referendum on EU membership. Consequently, the thesis was designed primarily to research the physical, social and economic consequences surrounding the relationship between land use and water quality from the perspective of the WFD. It is pertinent to state that the research presented in this thesis (e.g. modelling faecal pollution in UK rivers in response to land use change and quantifying the socio-economic and environmental benefits for people if riverine pollution levels are altered) can easily be decoupled from the legislative imperatives of the WFD.

The decision to leave the EU, providing the UK does indeed leave the EU, may potentially result in legislative changes in the mid- to long-term. Depending on the shape of future legislation, this research may, conversely, have greater relevance as UK policymakers become tasked with modifying legislation to suit an altered political and economic climate. The potential shape of the UK's post-Brexit water policy, and the implications for the value of this research, are discussed in a postscript.

Research outcomes and contributions made to the fields of study

Previous studies of individual catchments have used regression models based on relationships between GM FIO concentrations recorded at monitored sites and land use within their subcatchments. The work in Chapter 1 extends this approach by augmenting the predictor variables to, for the first time, include direct measures of key FIO sources (i.e., human population and livestock density data) and various factors (catchment size, runoff and soil hydrology) that may affect source strength and the mobilisation, transport and die-off of FIOs. Furthermore, the work pioneers the development of transferable generic models by combining data from 14 different catchment studies across the UK. By combining data from all 14 studies, which have a wide geographical distribution across UK and encompass a wide range of weather conditions, the effects of the temporal factors are minimised and the inter-catchment errors reduced. The resulting land cover- and population-based models can be employed, with some confidence, in UK catchments both to predict base- and high-flow FC and EN concentrations in unmonitored watercourses and to evaluate the likely impact of different land use/stocking level and human population change scenarios.

Chapter 2 develops a transfer methodology which allows the models to predict FIO concentrations in unmonitored UK watercourses. This enables the models to provide a cost-effective diagnostic tool capable of identifying and predicting the sources and spatial distributions of microbial pollution. By developing and incorporating human and livestock FIO sources as explanatory variables these models can be used to help apportion the responsibility for microbial pollution between the water industry and the agricultural sector. This research highlights issues of spatial scale surrounding the delivery of land use policy measures: the models can be used at a range of spatial scales, capable of expansion up to the scale of the UK, and are capable of identifying non-compliant HRUs which may benefit from micro-scale BMPs. The regression modelling approach developed here, by enabling spatially sensitive FIO function transfers, can inform integrated catchment management programmes, as required by the WFD.

The work in chapters 1 and 2 makes a valuable contribution to the field of catchment microbial dynamics. It is of great benefit to the policy and land management communities as it enables insights into the optimal locations and mix of implementation strategies for the delivery of WFD induced environmental improvements. The research also contributes to the emerging international debate on the use of farm best management practices and policy instruments to reduce FIOs and agricultural diffuse pollution (e.g., Bateman et al. (2006a), Chadwick et al. (2008), Monaghan et al. (2008), Helming and Reinhard (2009), Hutchins et al. (2009), Maringanti et al. (2009) and Oliver et al. (2009)).

The underlying rationale of Chapter 3 is that, given resource constraints, a focus on identifying and improving those river sites which yield the largest net benefits is entirely justified. This in turn requires estimates of the benefits of improvements to set against costs, and this analysis reveals the importance of key parameters (such as type of water quality attribute, respondent type and their distance from the proposed improvement) in determining those benefits.

As previous UK research has tended to conflate the value of ecological improvements with the value of recreational improvements, it is evident that decision-makers might target the ecological quality of water with little consideration of the impact on the recreational quality and vice versa. This is problematic if we don't know fully understand the contribution to value of each attribute. This research provides answers to this policy question by presenting two complementary models (CL and LC), that examine different aspects of the same data to disentangle and identify respondents' preferences for ecological and microbiological river water quality.

The research also ascertains how preferences for water quality characteristics differ across different types of recreational users'. The hypotheses of heterogeneity in preferences among the general public's WTP values is also confirmed. Respondents' perceptions of water quality are incorporated into the analyses of the data to provide welfare measures. The WTP measures derived from this research reveal clear differences in preferences between respondent groups, and so, from a policy perspective, enhances the ability of the policy-maker to more fully understand potential non-market benefits, in particular arising from improvements to the microbial quality of water, and thus produce more accurate cost-benefit analyses.

This research demonstrates that the non-market benefits that may accrue from different types of water quality improvements are nuanced in terms of their environmental impacts, their potential beneficiaries and, by inference, their overall value and policy implications. This information allows decision makers to better understand the consequences of adopting alternative investment strategies that favour either ecological or recreational improvements, or a mix of benefits, as these trade-offs were previously poorly understood.

While each of the empirical chapters provide novel contributions in their own right, it is their cumulative contribution against which should be judged. Taken together, these analyses link natural sciences, geographical analyses and economic valuation concerning related aspects of water quality. It is this interdisciplinary approach which is the hallmark of a contemporary natural capital approach to integrated analysis and decision making, and, as such, it is hoped that this thesis makes a significant contribution.

The challenges facing the field of catchment microbial dynamics have expanded rapidly in recent years. This thesis cannot fill all of the knowledge gaps, but we see a logical progression through the thesis. Chapters 1 and 2 examine the relationships between land use and microbial pollution. The methodology developed here is a significant advance and shows that FIO function transfers are theoretically possible. Chapter 3 develops a methodology for valuing change in microbial pollution for a range of respondent types. WFD compliance seems likely to yield spatial variation not just in the distribution of the implementation costs but also in the benefits of FIO reduction and in the willingness to pay (WTP) for river improvements.

The natural extension to the research presented here is to produce a spatially explicit valuation of the non-market benefits and WTP arising from FIO reduction and, within a cost benefit framework, to assess the relative cost effectiveness of different remediation strategies, such as those introduced in this paper. Analyses of this nature will be essential when assessing the optimum spatial differentiation and implementation of land use policies.

From a methodological perspective, the research within this thesis enhances the ongoing synthesis of geographic and economic social sciences. This work also addresses important policy questions which are of interest to a variety of stakeholders including government departments and agencies, the water industry, consumer groups and, importantly, the general public.

To avoid the contents of this thesis simply becoming a dusty tome on the author's supervisor's shelf, concerted effort has been made to disseminate this research as widely as possible via publications, conference presentations and press releases. Research output is detailed in Appendix II.

1 Chapter 1: Generic Modelling of Faecal Indicator Organism Concentrations in the UK

1.1 Introduction

The EU Water Framework Directive (WFD) (2000) established a framework for the protection of inland, transitional and coastal waters (Environment Directorate General, 2005). It placed a legal requirement on the UK's water regulators to manage pollution to achieve 'good ecological status'.

Good ecological status is defined in Annex V of the Water Framework Directive in terms of the quality of the biological community (i.e. the composition and abundance of aquatic flora, invertebrate fauna and fish fauna), the hydromorphological elements supporting the biological elements (i.e. the quantity and dynamics of water flow or the structure and substrate of the river bed), and the chemical characteristics (i.e. oxygenation and nutrient conditions) impacting upon the biological elements (EU, 2000). Nutrient pollution (e.g. phosphates and nitrates from farm fertilisers and phosphates from washing detergents) can cause algae to grow in rivers which, in turn, reduces the oxygen available for aquatic flora and fauna.

One important motivation for the implementation of the WFD appears to be the improved provision of, and opportunities for, open-access recreation (Articles 4, 9 and 11 of the WFD). Microbiological water quality, relevant for human health (and the quality of open-access recreation such as paddling, swimming or boating), is largely determined by faecal pollution (i.e. harmful bacteria, viruses and other infectious microorganisms), typically from livestock and/or human waste via wastewater treatment works. It is this aspect of river water quality that the first two chapters of this thesis addresses.

Bacterial/microbiological water quality is not specifically addressed in the WFD but the revised Bathing Water Directive (rBWD) (EU, 2006a) and Shellfish Water Directive (EU, 2006b), which complement the WFD, contain strict microbiological standards using faecal coliform (FC) and intestinal enterococci (EN) faecal indicator organisms (FIOs) as surrogates for infection risk in designated protected areas, such as bathing waters and shellfish harvesting areas (Edberg et al., 2000; IGES, 2008).

The use of FIOs as a measure of the safety of drinking water has a long history. In the 1890s many of the primary human pathogens were identified and categorized (Edberg, 1998). Concurrently, it was realised that public health protection required a cost effective indicator of faecal pollution, to avoid the expense of testing drinking water for all known pathogens. *E. coli* was chosen as the preferred biological indicator of water treatment safety for several reasons: it is universally present in the faeces of humans and mammals, is present in large numbers, is readily detectable by simple and inexpensive methods and would not multiply appreciably once voided into the environment (Prescott and Winslow, 1915). However, due to methodological deficiencies (e.g. it required several days and a number of subcultures in order to identify the bacterium), *E. coli* surrogates such as the FC test were developed (Eijkman, 1904). FC are considered to be present specifically in the gut and faeces of mammals and *E. coli* is a major species within the FC group.

Transmission of pathogens that can cause ill-health in recreational water is analogous to waterborne disease transmission in drinking water (WHO, 2003). As with drinking water quality, in both marine and freshwater studies of the impact of faecal pollution on the health of recreational water users, several faecal index bacteria have been used for describing water quality because they behave similarly to other harmful faecally derived pathogens (Prüss, 1998).

Whilst FC correlates well with health outcomes in freshwater, there are difficulties using it as an FIO in marine water because it is thought that some of its constituent index bacteria, in particular *E.coli,* may die off more rapidly in sea water than in freshwater, resulting in higher concentrations of harmful pathogens in seawater when index organism densities are identical (WHO, 2003). Such difficulties are discussed further in Chapter 2. One FIO that correlates well with health outcomes for both marine and freshwater is EN (Prüss, 1998).

The International Organization for Standardization (ISO, 1998) has defined EN as an appropriate FIO for use in both fresh and salt water environments. Alongside the adoption of both FC and EN as suitable FIOs by the EU, the United States Environmental Protection Agency (USEPA, 2012) has adopted EN and *E. coli* for fresh water, EN for marine water, and EN has been adopted by the WHO as the most suitable health criterion for both marine and freshwater environments (WHO, 2003).

Microbial pollution remediation is central to the WFD strategy for water quality improvements (Kay et al., 2007a). Under the WFD, EU member states are also legally required to design and implement catchment scale ’programmes of measures’ to manage non-compliant sources of microbiological pollution that could cause non-compliance of bathing and shellfish-harvesting waters with microbial standards (EU, 2000; Kay et al., 2006a).

The policy shift from assessing water in terms of its chemical composition in favour of its ecological quality (Bateman et al., 2006a) has caused unease among water regulators as, in the past, there has been little effort to measure the microbiological quality of our water (Kay et al., 2007a). Indeed, there are many within the research community who feel that policy is running ahead of the capabilities of water quality science (Chadwick et al., 2008). Despite this, it is acknowledged by the policy and management communities that significant reductions in diffuse agricultural pollution and substantial improvements to waste water treatment (WWT) infrastructure are required to achieve WFD compliance targets (Wither et al., 2005). There has been success in reducing pollution. Ofwat, the UK water industry regulator, is highly effective in compelling water companies to make WWT improvements, with over £5.5 billion invested on environmental schemes between 2005-10 (Ofwat, 2008). Record levels of rBWD compliance were achieved in 2006 (Defra, 2008a), but despite these improvements to WWT infrastructure and changes to farming methods leading to reductions in WFD non-compliance, many Environment Agency (EA) pollution monitoring sites continue to record high levels of FIO pollution, particularly from agricultural sources (Crowther et al., 2001; Aitken, 2003).

To drive further improvements to water quality, accurate data is needed to more accurately define faecal indicator organism concentrations and fluxes in individual rivers and streams. This will allow the magnitude of the problems in non-compliant rivers to be assessed, enabling heavily polluted waters to be identified and marked for priority remedial action. Accurate data is also required to assess the effectiveness of measures which have previously been implemented to reduce riverine pollution.

Given the legal requirements of the WFD, and the clearly defined knowledge gaps in current FIO modelling, there is an urgent imperative from the research and policy communities for generic transferable models that can accurately predict base- and high-flow FIO concentrations across the UK to better inform integrated catchment management programmes. One such programme, the Catchment hydrology, Resources, Economics and Management (ChREAM) project (Bateman et al., 2006a) required a transferable model capable of predicting riverine FIO concentrations. ChREAM also specified that such a model must be achieved within a standardised data framework, to enable full integration with other aspects of ChREAM land use and hydrological modelling.

This is quantitative theory-building exploratory research approached from the positivist standard view of science, as defined by Robson (2002). This chapter describes the design and construction of transferable models capable of accurately predicting microbial pollution concentrations in UK rivers, using nationally available data. The models developed here were subsequently reported in Crowther et al. (2011).

Following a review of contemporary FIO modelling and a statement of the research aims and objectives, the development of the datasets underpinning the transferable generic models is reported. Land use profiles have been synthesized from the Centre for Ecology and Hydrology (CEH) Land Cover Map 2000 (LCM2000) and the Ordnance Survey (OS) Meridian2 map (NERC, 2008a; Ordnance Survey, 2008) in order to combine their best features and minimise their inaccuracies. Previous FIO models have typically described ordinal, rather than cardinal, change within land use categories, i.e. they have recognised that an area is inhabited by humans, but revealed nothing about the concentration of humans within that area. For this reason, this research introduces innovative population density variables, derived from readily available national land use data (i.e. the ONS decennial census for England, Scotland and Wales (Office for National Statistics, 2001a and 2001b) and the June Agricultural Survey (EDINA, 2008a), to characterise the distributions of humans and a range of livestock types.

It was hypothesised that the creation of quantitative variables, that describe the distribution and population density patterns of humans and livestock within a catchment, may yield accuracy improvements over the simple binary designation of land use employed in previous research, e.g. improved characterisation of the populations of potential FIO sources may lead to improved estimates of FIO concentrations in rivers. The accuracy and suitability of these enhanced datasets is assessed. The chapter then describes the meta-analyses which remodelled FIO data, obtained during fifteen Centre for Research into Environment and Health (CREH) catchment scale studies, into generic land use and population based models capable of predicting FIO concentrations. Results from these models are reported and the models' suitability as transferable generic models is assessed. The chapter closes with a discussion which identifies the limitations, and provides potential improvements, to the research design.

1.2 A review of contemporary faecal indicator organism modelling

The research field of catchment microbial dynamics has been rapidly expanding due to the adoption of the WFD (IGES, 2008). Such is the importance of the field to the successful implementation of the WFD, Haygarth et al. (2005) have gone so far as to describe the policy imperative to understand catchment microbial pollution concentrations and fluxes as “the new challenge of the 21st century”. Despite the research field being of national and international importance, significant data gaps exist that hinder efforts to characterise riverine microbial pollution. These data gaps are now identified and the efforts to address them discussed.

The sources of river pollution are varied spatially. Much of our riverine microbial pollution comes from diffuse agricultural sources (Bateman et al., 2006a; Haygarth et al., 2005; Horsey, 2006) but urban point sources of pollution, such as wastewater treatment works (WwTW), account for substantial pollution discharges into rivers; particularly during periods of high rainfall when aging (and often inadequate) wastewater infrastructures overflow due to their inability to process high volumes of wastewater. It has been estimated that point source discharges from WwTWs can contribute significant proportions of the total phosphorus load in UK rivers, with significant increases in concentrations downstream from WwTWs under high-flow conditions (Young et al., 1999).

The ecological and microbial aspects of water quality can be distinct and typically have unique pollution sources requiring different remediation strategies (Haygarth et al., 2005). Remediation, particularly of microbial pollution, requires the correct identification of pollution vectors (typically agricultural livestock waste or overflows from human WwTWs) to ascribe liability and enforce accountability. Remedial action is hampered by a lack of accurate FIO modelling (Stapleton et al., 2008), a shortage of empirical measurement of sewage overflows (Wither et al., 2005) and a lack of research into the effectiveness of different sewage treatment types (Kay et al., 2008a). In addition, many routine water quality monitoring programmes tend to be systematically flawed as they habitually sample during low flow (base-flow) conditions, rather than capture the full range of river discharge rates (Crowther et al., 2011). This shortage of accurate empirical data leads to flawed assessments of the magnitude of high-flow FIO concentrations from both diffuse and point sources (Mattikalli and Richards, 1996; Kay et al., 2005a; Kay et al., 2008a). The lack of basic data on hydrological fluxes is disturbing. The mass movement of FIOs from urban sources is typically associated with relatively short duration high rainfall events. For example, Stapleton et al. (2008) found that urban point source discharges were directly responsible for 90% of the total organism load to the Ribble estuary during high rainfall. This poses significant risks to human health as elevated microbial pollution causes unacceptably poor recreational water quality (Wither et al., 2005; Defra, 2008a). Microbiologically polluted water has been shown to have a dose-response relationship with the risk of ill-health (i.e. the rate of infection among recreational users increases steadily with increasing concentrations of harmful microorganisms and, for a constant concentration of microorganisms, the rate of infection is higher for those recreational users who have higher exposure) (WHO, 2003). The evidence used to calculate the dose-response relationship and the detrimental effect elevated microbial pollution concentrations can have on human health is discussed in Chapter 2.

Microbial pollution from diffuse sources is also elevated under high-flow conditions. Kay et al. (2008a) found that FIO concentrations and discharge volumes typically increase by an order of magnitude in rural catchments under high-flow conditions. This c.100-fold increase in export coefficients is due to a range of factors, not limited to the increased run-off of faecal material from agricultural land or the increased mobilisation and transport of FIOs due to increased turbidity within watercourses (Wilkinson et al., 2006).

The investigative monitoring of thousands of discharge sites for the presence of FIOs presents an expensive logistical challenge for the regulator as it is simply infeasible to measure pollution concentrations at every location (Environment Agency, 2008a). Because of this, there is a real and necessary requirement for cost-effective diagnostic tools capable of predicting microbial pollution sources and distributions. Although a range of statistical methods have been developed and used to model riverine pollution (Fraser et al., 1998; Tian et al., 2002; Vinten et al., 2004; Lawler et al., 2006) they are not without inaccuracies or disadvantages when applied to modelling FIOs. Watershed modelling tools such as Hydrological Simulation Program Fortran (HSPF) (Bicknell et al., 1997;

Donigian et al. 1995), Simulated Catchments (SIMCAT) (Warn, 1987) or Soil and Water Assessment Tool (SWAT) (Gassman et al., 2007). are frequently used to assess nutrient or sediment loadings in watercourses. The use of these systems to reliably model FIOs in watercourses is restricted by the poor availability of empirical data with which to parameterise or assess the accuracy of modelled results (Crowther, 2011). The Scotland and Northern Ireland Forum for Environmental Research (SNIFFER) screening tool has been useful for identifying and characterising diffuse pollution, providing insights into the sources of FIO pollution and enabling FIO export coefficients for catchments to be determined (SNIFFER, 2006a and 2006b). However, SNIFFER does not characterise both base- and high-flow FIO concentrations and the accuracy of SNIFFER's predicted export coefficients, in common with the previous tools, has not been evaluated against data from monitored catchments.

Another approach to catchment-scale FIO modelling has been to use linear regression techniques to model relationships between the geometric mean (GM) FIO concentrations recorded at monitored sites and the dominant land use characteristics within the catchments draining into those monitored sites; e.g. the proportions of grassland or built-up land act as proxies for the key sources of faecal pollution.

This approach allows the correlations between FIO concentrations and land use types to be examined and water quality maps to be generated. By examining the locations of anomalous standardised residuals revealed by the spatially referenced regression models it is possible to identify pollution sources in need of remediation (Crowther et al., 2001; Kay et al., 2007b).

This methodological approach to FIO modelling has proved to be a cost effective exploratory tool to predict microbial concentrations both within and at subcatchment outlets (Crowther et al., 2003; Kay et al., 2005a). Another advantage of the regression modelling approach is that the costs of obtaining empirical data are minimised. Research by Crowther et al. (2001) proved it is possible to investigate and predict FIO concentrations in coastal water by combining only secondary data sources. Similar FIO studies have obtained good results with a minimum of primary data (Wither et al., 2005). However, desktop studies that rely on secondary data are not without complications. These complications are now discussed.

FIO studies have shown that land use is a statistically significant determinant of microbial concentrations in rivers. Spatial variations in water quality closely reflect the distributions of developed land, meaning that urban land, with its associated sewage outflows, is one of the most critical sources of microbial pollution (Crowther et al., 2003). Unfortunately, inaccuracies in remotely sensed data can cause land-use misclassification. A comparison of the CEH LCM1990 (NERC, 2008b) with field survey data has revealed marked discrepancies in that map, particularly in the urban land use category (Kay et al., 2005a). A similar comparison of the CEH LCM1990 with OS 1:50,000 maps revealed substantial misclassification of urban and woodland areas in the Ribble catchment (Kay et al., 2005a). Misclassification also occurs within agricultural land use categories. The problem is in part caused by the light reflectance values of similar surfaces. For example, mapping software may misclassify bare rock as urban areas (CEH, 2008a). The methods used to rectify these errors are not without their own inaccuracies. During one correction exercise, mapped areas of built-up land extracted from OS 1:50,000 maps systematically underestimated urban land (Stapleton et al., 2006).

If uncorrected, misclassification can cause significant systematic errors, particularly in heavily urbanized catchments. Some misclassification can be corrected manually by reclassifying the seventeen pre-defined CEH land use classes into a reduced number of principal land use categories. The accuracy of land use classification can also be further improved by using the most current data sets, such as the LCM2000 (Kay et al., 2005a) which offers accuracy improvements over its predecessor, the Land Cover Map of Great Britain (LCMGB). For example, procedures were developed and incorporated within the LCM2000 for segmenting satellite images to produce vector outlines (Fuller et al., 2005) and the LCM2000 also incorporates upgrade improvements in structure, thematic detail and associated metadata (Smith and Fuller, 2002).

Despite the restricted availability of primary empirical data, and the difficulties associated with the use of secondary data, successive generations of desktop studies have tended to predict FIO concentrations with increasing accuracy (Kay et al., 2005a; Stapleton et al., 2008). Since 1995 CREH have been assembling an empirical database of enumerated FIO concentrations and accurate runoff and discharge data for catchments. This database is continuing to expand and develop. As of 2010, FIO concentration and export coefficient data have been collected from 205 sampling points across 15 process based catchment studies. One of the most recent modelling exercises undertaken by CREH combined and reanalysed those datasets within a meta-analysis to improve characterization of FIO fluxes within those catchments and assess the effectiveness of different sewage treatment types (Kay et al., 2008a).

The ChREAM project examined the agricultural costs and key non-market benefits associated with the introduction of the WFD and considered the impacts of alternative implementations of that policy in terms of its impact upon rural land use and the farming sector (Bateman et al., 2006a). These impacts will involve geographically varied changes in land use patterns and water quality. This large-scale study enabled collaboration with experts in the field of catchment microbial dynamics based at CREH and enabled limited access to CREH's commercially sensitive FIO concentration and export coefficient database. As previously mentioned, ChREAM required a generic transferable model capable of predicting riverine FIO concentrations using standardised data surfaces, enabling integration with other aspects of ChREAM land use and hydrological modelling. The meta-analysis reported in Kay et al. (2008a) was calculated using the disparate data sources detailed in Table 6 and had not been assessed for its ability to predict FIO concentrations in other catchments. Consequently, that model could not meet ChREAM requirements. To achieve full integration with all aspects of ChREAM land use modelling it was necessary to reanalyse the CREH datasets using standardised data surfaces and standardised predictor variables.

In addition to the creation of standardised data surfaces from which CREHs primary FIO concentration and export coefficient database could be remodelled, this research also aimed to extend the regression modelling approach by investigating whether improved models might be achieved by including human population and livestock density data as direct measures of the key FIO sources as previous CREH research has relied solely upon ordinal change within land use categories.

Other independent variables that are known to affect FIO source strength, mobilisation, transport, die-off and sedimentation within catchments (e.g., volume of runoff, soil hydrology and catchment size) have also been incorporated within the meta-analyses reported here. This has resulted in comprehensive generic and transferable models which can be used to predict FIO concentrations in UK rivers.

1.3 Aims and objectives

The overall aim of this chapter is to produce a transferable generic model capable of accurately predicting EC and EN FIO concentrations in unmonitored UK rivers at both base- and high-flow during the summer bathing season.

In doing so, it will also attempt to meet the following objectives:

- Produce standardised land use variables, synthesised from the CEH LCM2000 and Ordnance Survey (OS) Meridian 2 maps, and apply the Centre for Ecology and Hydrology (CEH) land use classification system (grouping land into seven principal land use categories, so that the resultant datasets integrate with other aspects of ChREAM land use modelling). This synthesis of the OS Meridian2 and CEH LCM2000 maps may better represent the extent of built-up areas, providing accuracy improvements over existing land cover maps.
- Standardise the Centre for Research into Environment and Health (CREH) land use data to achieve consistent land use classification across previous CREH FIO studies. This standardisation may yield welcome accuracy improvements when producing the generic FIO models, but is also crucial in facilitating the full integration of the FIO modelling with other facets of the Catchment hydrology, Resources, Economics and Management project (ChREAM) land use and water quality modelling.
- Assess the accuracy of the standardised land use variables against CREH field survey land use data.
- Create a method by which decennial national census data may be interpolated to catchment areas. Incorporate that human population density variable into the FIO models. This innovation may explain the significance of human FIO discharges better than existing urban land use variables.
- Incorporate livestock population density variables into the FIO models for the first time. This may lead to a richer understanding of the distribution and importance of diffuse and point sources of pollution from agriculture.
- Identify, interpolate to catchment areas and incorporate additional datasets, with nationally available coverage (e.g. soil temperature or

Standard Percentage Runoff (SPR)), which may help to explain riverine FIO concentrations.

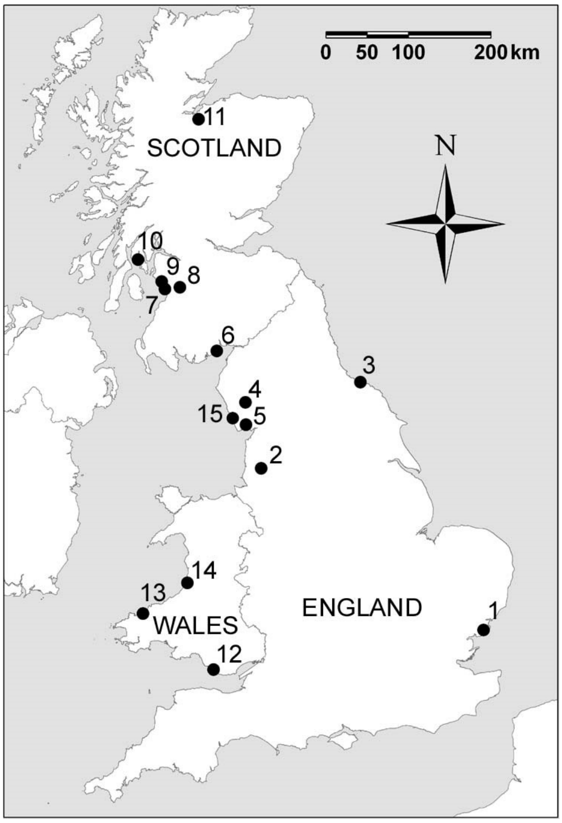
- To produce accurate digital catchment boundary files for CREH catchments, where necessary.
- CREH catchment studies will be calibrated and integrated to produce transferable models (using stepwise multiple regression techniques), capable of accurately predicting riverine FIO concentrations.
- Assess the relative performance (with reference to R^2 values) of land-use-based, population-density-based and all-variable FIO models.
- The transferability of the FIO models will be assessed via out-of-sample testing.

1.4 Methods

1.4.1 Study catchments

The primary empirical datasets of enumerated FIO concentrations and run off export coefficients used within this research are derived from monitoring undertaken between 1995-2005 in the 14 CREH catchment studies detailed in Table 4. Their locations are shown in Figure 1. The 15th catchment, Haverigg, was used to evaluate the transferability of the generic models.

Figure 1: locations of CREH catchment studies



Locations of the 14 study catchments (1–14, as detailed in Table 4) used within the modelling data set. The Haverigg catchment (15) was used for model evaluation.

Each of the studies detailed in Table 4 were conducted during the summer bathing season as they were aimed at improving understanding of bathing water compliance. To increase the robustness of the present modelling, only sites that met the following criteria were included: (1) Due to the relatively low resolution of livestock census and soil hydrology data only subcatchments with area ≥ 5 km^2 were included; (2) < 50% of land within the subcatchment is located upstream of lake and/or reservoir outlets. The reason for this is discussed below; (3) To produce accurate characterisation of FIO concentrations across different discharge conditions, ≥ 5 samples of FIO data must be available for both base- and high-flow at each sampling location; (4) river discharge records must be available; and (5) land within the subcatchments had not been subject to programmes of measures (e.g. riparian fencing and buffer strips) aimed at reducing FIO loadings. The data from the 153 base-flow sites and 134 high-flow sites which meet these criteria and were subsequently used to model base- and high-flow faecal coliform (FC) and intestinal enterococci (EN) concentrations during the summer bathing season, are shown on Table 4.

Table 4: water sampling points from previous CREH studies used in the present study

Study Catchment		Year Sampled	Number of water sampling locations	
			Base-flow	**High-Flow**
1	Holland Brook	1998	10	10
2	River Ribble	2002	37	37
3	Staithes Beck	1995	6	2
4	Lake Windermere	1999	3	3
5	River Leven/Crake	2005	16	16
6	Sandyhills	2004	4	4
7	Troon coastal inputs	2000	1	1
8	Killoch Burn	2004	1	14
9	River Irvine/Garnock	1998	23	23
10	Ettrick Bay	2004	1	1
11	River Nairn	2004	8	8
12	Afon Ogwr	1997	14	14
13	Afon Nyfer	1996	17	2
14	Afon Rheidol/Ystwyth	1999	12	12
	Total number of sampling sites		**153**	**134**

There are combined sewage overflows (CSOs) and wastewater treatment works (WwTWs) within many of the subcatchments. Consequently, the FIO concentrations recorded within those subcatchments reflect inputs from a range of both point and diffuse sources.

1.4.2 Water sampling and FIO enumeration methods

The procedures used by CREH for collecting water samples and enumerating FIO concentrations were largely standardised. Aseptic water sampling programmes, devised to collect samples at both base-flow and high-flow, were carried out at the EA sampling point network, where possible (Kay et al., 2008a).

Temporary staff gauges were used at any location where EA telemetry data was unavailable. High-flow periods were defined by standard base-flow separation of the hydrograph record for the nearest hydrometric station (Kay et al., 2005a). FC and EN concentrations were enumerated following industry standard membrane filtration methods (Environment Agency, 2000; HMSO, 1994; Kay et al., 1994). The discharge monitoring, base-/high-flow separation, water sampling and microbial analysis are discussed in detail in Kay et al. (2008a). The variables relating to base-flow runoff (m^3 km^{-2} h^{-1}), high-flow runoff (m^3 km^{-2} h^{-1}) and total runoff (m^3 km^{-2} h^{-1}) are derived from actual runoff data collected during the individual studies.

1.4.3 Outputs from lakes/reservoirs

Because of sedimentation and die-off of microbial organisms within reservoirs and lakes, previous research has shown that watercourses issuing from such waterbodies typically have very low FIO concentrations which may poorly reflect land use, livestock stocking levels or wastewater discharges within the contributing catchment (Kay and McDonald, 1980). For this reason, the geometric mean (GM) FC and EN colony-forming unit (CFU) concentrations in watercourses issuing from reservoirs/lakes have been set to the values reported in Table 5.

Table 5: GM FC and EN concentrations in waters issuing from lakes and reservoirs

Organism Type	GM concentrations (CFU 100 ml^{-1})	
	Base-flow	High-Flow
FC	26	83
EN	5	16

These values are based on research conducted at the Nant-y-Moch, Cwm Rheidol, Fewston and Thruscross reservoirs and Lake Windermere (Kay, 1979). Subcatchments in which over 50% of land is located upstream of reservoir/lake outlets were excluded from the modelling data set. Where reservoir/lake outlets are fed by less than 50% of the total area of the subcatchment, it has been assumed that the volume of flow recorded at the subcatchment monitoring point consists of two components proportional to the area derived from the waterbody and the area derived from the rest of the subcatchment.

1.4.4 Catchment boundary data

Twelve of the fifteen CREH catchment Hydrological Response Unit (HRU) boundary datasets had previously been georeferenced by CREH. Spatial boundaries corresponding with FIO relevant topographic features, such as reservoir outlets, were given precedence during that process specifically for the purpose of modelling FIOs. However, reservoir catchment boundaries were missing from the original CREH digital boundary files for the Windermere and Leven/Crake catchments and catchment boundaries for the Troon study (Wyer et al., 2001) were only available as 1:25000 paper maps. For this present study, catchment boundaries for the Windermere and Leven/Crake catchments were augmented to include the reservoir catchments and HRU boundaries for the Troon catchment were georeferenced using ArcMap 9.1 (ESRI Inc., 2005).

The georeferencing process followed the usual procedure for georeferencing a raster dataset (ESRI Inc., 2006). In short, this involved identifying a number of well distributed locations in the unreferenced image and adding these as coordinate control points, then linking the known raster dataset positions to known positions in map coordinates. Once this had been achieved it was straightforward to digitise the HRU and reservoir boundaries using the editor tools in ArcMap.

1.4.5 Construction of consistent data surfaces

As the lack of a consistent mapping standard within previous CREH studies may create data inconsistencies within a meta-analysis of those studies, standardised data surfaces were required for this research. Table 6 describes the disparate land use data surfaces used within the individual CREH studies and the standardised data surfaces created to replace them. The data sources for the land use and population profiles were chosen specifically because they are readily available and have national coverage. Therefore, by using the methodologies described below, consistent profiles can be generated for the entire United Kingdom.

Table 6: standardised variables used in the meta-analysis

	Study Catchment	Year Sampled	Original land use data sources	Standardised land use data source	Agricultural census years used for livestock enumeration	Census year used for population enumeration
1	Holland Brook	1998	Field mapping/OS	Synthesised Meridian2 and LCM2000 used throughout	England/Wales 1997	2001 census used throughout
2	River Ribble	2002	ITE1990/OS		England 2003	
3	Staithes Beck	1995	Field mapping/OS		England/Wales 1995	
4	Lake Windermere	1999	LCM2000/OS		England/Wales 1997	
5	River Leven/Crake	2005	LCM2000/OS		England 2004	
6	Sandyhills	2004	Scottish Executive		Scotland 2004	
7	Troon coastal inputs	2000	Estimated		Scotland 2000	
8	Killoch Burn	2004	Scottish Executive		Scotland 2004	
9	River Irvine/Garnock	1998	Field mapping/MLCM		Scotland 1997	
10	Ettrick Bay	2004	Scottish Executive		Scotland 2004	
11	River Nairn	2004	Scottish Executive		Scotland 2004	
12	Afon Ogwr	1997	Field mapping/OS		England/Wales 1997	
13	Afon Nyfer	1996	Field mapping/OS		England/Wales 1996	
14	Afon Rheidol/Ystwyth	1999	Field mapping/OS		England/Wales 1997	
15	Haverigg	2008	LCM2000/OS		England/Wales 2004	

Notes on data sources: OS = 1:50,000 mapping for built up land and woodland; MLCM = Macaulay LCM1988; ITE1990 = Institute of Terrestrial Ecology LCM1990; LCM2000 = CEH Land Cover Map 2000; Estimated = estimates of built up land (OS) and improved pasture.

The next section of this chapter describes how the standardised land use variables were synthesised from the CEH LCM2000 and OS Meridian 2 maps (NERC, 2008a; Ordnance Survey, 2008). This is followed by a description of the construction of the human population density, livestock population density and Standard Percentage Runoff (SPR) for soil types variables, using the Office for National Statistics (ONS) decennial census for England, Scotland and Wales (Office for National Statistics, 2001a and 2001b), the June Agricultural Survey (EDINA, 2008a) and the Institute of Hydrology's Hydrology of Soil Type (HOST) database (Boorman et al., 1995) respectively, within a spatially explicit geographical information systems (GIS) framework using ArcGIS 9.1 (ESRI Inc., 2005).

1.4.6 Land use data reclassification

The purpose of this land use reclassification of CREH catchment areas was to produce accurate land use profiles of the number of hectares of each type of land cover in each HRU from which a meta-analysis of the CREH FIO data could be made. In addition to providing accuracy improvements over existing land cover maps, the reclassification scheme described here applies the CEH classification system which groups land into seven principal land use categories. These are urban/suburban, rough grazing, temporary/permanent grassland, woodland, arable/set aside, water and all other land. The 'arable' category can be subdivided and reclassified for other aspects of ChREAM land use modelling as necessary. Supplying CREH with standardised land use datasets, generated by the same method as for other aspects of the ChREAM project, enabled CREH FIO models to be fully integrated into all aspects of ChREAM land use modelling.

By verifying Developed Land Use Area (DLUA) polygons from the OS Meridian2 map (Ordnance Survey, 2008) and urban land use data from the LCM2000 (NERC, 2008a) against OS 1:25,000 paper maps, the method described in Posen et al. (2011) was further developed so that the synthesis of the OS Meridian2 and CEH LCM2000 maps may better represent the extent of built-up areas. For example, although DLUA polygons tend to capture the outlines of large urban areas very well they occasionally misrepresent some aspects of urban land use, e.g. not capturing the full extent of urban sprawl in some newly-developed areas, or incorrectly identifying some very small settlements as DLUAs. In contrast, the

LCM2000 is good at capturing the extent of urban sprawl but also captures features that are not necessarily urban, such as quarries or motorway interchanges. The synthesis and proportional interpolation methods used to allocate land use data to subcatchments described here combines the best features from both maps while minimising their inaccuracies.

1.4.7 Method of reclassifying the 'urban' total in the Land Cover Map 2000

Preparation of LCM 2000 data:

1a, Using ArcGIS 9.1 (ESRI Inc., 2005) extract classes for 'urban' and 'suburban' from LCM 2000. Reclassify 'urban' and 'suburban' areas as 1, all other classes as 0.

2a, Isolate densely urban areas from very small rural settlements using a focal statistics operation to clump groups of cells based on values ranging from 0 and 1, where 1 = urban and 0 = rural, using a grid of 7*7 cells. The focal statistics operation assigns a value to each cell based on the values of surrounding cells, therefore cells surrounded by other cells of the same value will have a high value and vice versa. Assigning the 'urban' classification to Focal Statistics values > 0.9 removes of all the LCM 'urban speckles' while preserving the true urban areas. Cells with values between 0.9 and 1 are reclassified as 'urban'. Cells with values lower than 0.9 are reclassified as 'all other land'.

Preparation of OS Meridian2 data:

1b, Use the 'feature to polygon' function to convert the DLUA layer to polygons.

2b, Convert DLUA polygons to grid resolution (25m) and reclassify DLUA grid data. DLUA areas = 1, all other classes = 0.

Merging and reclassifying the prepared LCM 2000 and OS Meridian urban data:

3, Add together the grid layer LCM2000 urban areas (2a) and DLUA grid data (2b). The resulting 'urban' layer is assigned values of 1 for urban and 0 for all other land uses.

4, Very small settlements, i.e. the remaining 'speckles' from LCM 'urban' categories, are now classified as 'all other land'. Raster clumps are grouped into regions. Regional groups > 80 cells are retained (value = 1), all those areas with a cell count < 80, corresponding to an area < 5ha., are classed as ‘all other land’ (value = 0).

5, Occasionally small clumps of peripheral cells, that are clearly part of a larger urban area, separate from the urban area. To correct this problem, the original DLUA grid of Meridian defined urban areas, obtained at step 2b, is added to the output layer.

6, The non-urban categories from LCM2000 are now reclassified into the 7 principal categories used by CEH and consistent with ChREAM land use modelling. This reclassified layer is then mosaiced with the new ‘urban’ and ‘all other land’ categories. The reclassified ‘urban’ and ‘all other land’ categories take precedence over the previous ‘urban’ and ‘all other land’ categories and any resulting small gaps are classed as ‘all other land’.

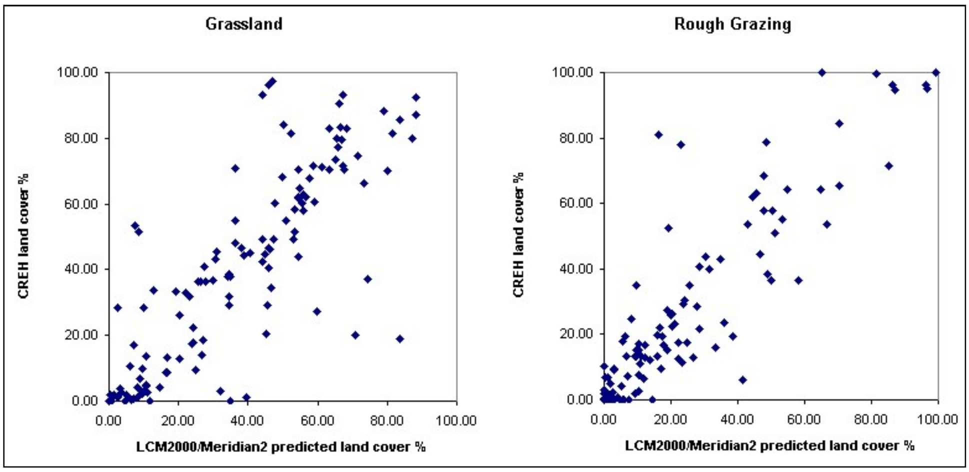
7, By cross-tabulating the HRU polygons with the CEH categories it is then possible to extract the number of hectares and the proportion of each land use type within each HRU.

1.4.8 Assessing the suitability of the synthesised land use data

This section of the chapter assesses the accuracy of the synthesised LCM2000/Meridian2 land use data surfaces produced during this research by comparing them against field survey data, collected by CREH, for the HRUs in six catchments.

At the scale of individual sub-catchments, analysis revealed some discrepancies between the LCM 2000/Meridian2 and CREH datasets. Although the scatter plots in Figure 2 reveal errors in the correlation between the two data sets for the grassland and rough grazing categories, there is no clear evidence of systematic bias.

Figure 2: comparing the proportions of grassland and rough grazing obtained by CREH field survey data against those predicted by the synthesised LCM2000/Meridian2 maps

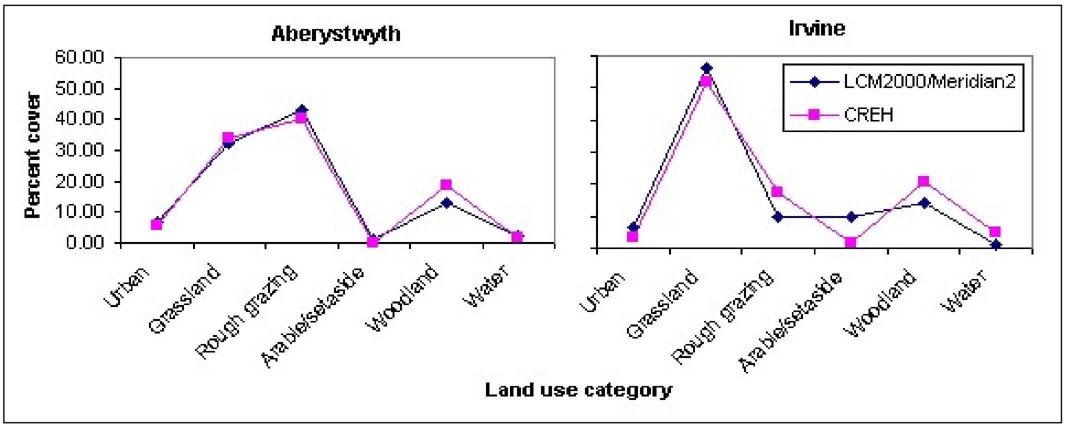


CREH data based on field survey and supplemented by OS mapping (Crowther, 2008a)

At sub-catchment scale the woodland category showed no significant differences. Paired sample t-tests suggested that the total area of urban land calculated by CREH in each sub-catchment is lower than that predicted by the synthesised land use data. These types of fine-scale discrepancies between the two datasets were not unexpected as the CREH field survey data was gathered at a much higher resolution than the LCM2000/Meridian2 data.

Further analysis found that as spatial scale increases (and resolution becomes coarser) differences in land use patterns between the two datasets tended to converge. At catchment scale the LCM2000/Meridian2 dataset closely corresponds with the CREH field survey data, particularly for the Aberystwyth catchment, shown in Figure 3.

Figure 3: comparing the LCM 2000/Meridian2 data with CREH field survey data at catchment scale



Despite the fine-scale differences at sub-catchment level, there are strongly significant ($p<0.01$) positive correlations between the land use categories across the two data sets at catchment scale, particularly for those land use categories (urban, grassland, rough grazing) that have been demonstrated to be significant sources of FIOs (see for example Kay et al, 2005a). Correlations between the datasets range from 0.814 for grassland to 0.938 for woodland. The correlation between the synthesised urban and the CREH urban is 0.894.

In conclusion, despite some reservations about the accuracy of the synthesised urban total for individual sub-catchments when compared with the fine resolution data collected by CREH (which may cause slightly lower r^2 values within the FIO models), the reclassified and synthesised land use data generated in this research is almost certainly fit for producing accurate land use profiles for use in larger scale modelling exercises.

1.4.9 Human population data

Urban land use is one of the single most important variables used in previous CREH FIO modelling (Kay et al., 2005a). However, areas such as retail parks and industrial complexes have relatively low population density but are classed as urban land. It was hypothesised that the creation of quantitative variables describing the distribution and density patterns of the human population within urban areas, at the resolution of individual HRUs, may yield accuracy improvements over the simple binary designation of urban land use employed in previous research.

Population profiles for each HRU were derived from the ONS decennial census for England, Scotland and Wales (Office for National Statistics, 2001a and 2001b). 2001 was the most appropriate census year, being at the approximate mid-point of the period that the 15 catchment studies were undertaken. ONS Census data is available in a range of resolutions. Output Area (OA) data was chosen because it is the most detailed geographic level for which 2001 census data are available. Each OA has approximately 300 residents and, importantly, OA boundaries enclose as compact an area as possible (Office for National Statistics, 2005). Because population is not distributed evenly in space it was important to use the most spatially compact and detailed data available, in order to reduce errors when interpolating population. The rationale is that points close together in space are more likely to have more similar values than points further apart (Waters, 1997). It was anticipated that the larger the spatial area, the larger the interpolation errors would be. These interpolation errors were confirmed by preliminary investigations using coarse resolution Super Output Area (SOA) census data, which yielded less accurate results.

1.4.10 Human population profile calculations

1. Intersect OA shapefile (1) with the subcatchment HRU shapefile (2). Output = shapefile 3.

2. Edit shapefile 3 attribute table. Create fields 'oa_m^2' and 'oa_ha' for OA area values. Calculate 'm^2', then 'oa_ha' as m^2/10000. This calculates the area, in hectares, of each OA.

3. Clip, then union shapefile 3 to the HRU shapefile (2). Output = shapefile 4.

4. Edit shapefile 4 attribute table. Create and calculate fields 'union_m^2' and 'union_ha' for OA unioned polygon area values expressed in hectares. Create a field 'proportion' where 'proportion'='union_ha'/'oa_ha'.

5. Add 2001 OA population data as a new layer. Join the population data to the shapefile 4 attribute table based on a common field i.e. 'label'. Output = shapefile 5.

6. Edit shapefile 5 attribute table. Create a field ‘polygon_population’ where ‘polygon_population’=‘population’*‘proportion’. This calculates the proportion of OA population estimated to reside in the unioned polygon area.

7. Dissolve shapefile 5. Sum ‘polygon_population’ based on the HRU shapefile (2) HRU ID field. This final step dissolves the unioned OA polygons into the HRU polygons and sums the population estimated to reside within each HRU.

8. Population density for each HRU is obtained by dividing HRU population by HRU area.

1.4.11 Livestock population data

The proportion of grassland in agricultural catchments is a dominant predictor variable, particularly at high-flow (Crowther et al., 2002; Crowther et al., 2003; Kay et al., 2005a). However, the algorithms used by CEH to categorise grassland in the Land Cover Map also capture other areas of short grass (CEH, 2008a). Consequently, the land use category ‘grassland’ will include playing fields, golf courses and urban green spaces, all of which rarely contain dairy herds, the primary source of agricultural FIOs. The purpose of this classification exercise is to capture livestock types, populations and densities within HRUs. It is hypothesised that by quantifying livestock populations in this way, significant variables can be produced which will improve the accuracy of predictive FIO models.

To integrate consistently with the land use and human population density variables described previously, a derivation of the previously described method has been use to generate livestock profiles for CREH catchment HRUs. Agricultural census data (Agcensus) from the June Agricultural Survey (EDINA, 2008a) was used to generate these profiles. The livestock categories are dairy herd, beef herd, bulls, other cattle over one-year-old, other cattle under one-year-old, sheep (which also includes goats, deer and horses), total pigs, indoor pigs, outdoor pigs, total fowl, indoor fowl and free-range fowl. It should be noted, as detailed in Table 7, that some of these categories were aggregated for modelling purposes.

Just as the smallest census enumeration units were used to minimize interpolation errors in the human population datasets, the smallest available Agcensus grid square resolution, 2km*2km, was used in these calculations for the same reason.

Livestock populations fluctuate to a far greater extent than human populations, as animals are reared and slaughtered. Fortunately, unlike the decennial human census, the Agcensus is produced more regularly, often annually or biannually. Agcensus data corresponding to the year of the CREH catchment studies was used, where available, to minimize enumeration errors. There are several instances where the corresponding Agcensus year has not been used. There are two main reasons for this. Either there was no Agcensus that year, or the data within the census had been coarsely aggregated to preserve the confidentiality of individual farmers (EDINA, 2008b), as was the case for the England 2000 and Wales 1999 censuses. In these cases, the closest census year was used. Table 6 shows the Agcensus year used to generate each of the livestock profiles.

There were several other unavoidable sources of enumeration error caused by differences between individual Agcensus datasets. Data for bulls was unavailable on the census for Eng./Wales 1996. To rectify this the number of bulls was calculated as 4.4% of the 'other cattle over one-year-old', as this was the average proportion of bulls in other Agcensus years. Data for fowl was unavailable on the census for Eng./Wales 1997, so this variable was derived from the Eng./Wales 1996 census data. Due to differences in agricultural policy the questionnaires used in England, Wales and Scotland occasionally differed slightly, which also led to minor inconsistencies in the data (EDINA, 2008b).

1.4.12 Livestock population profile calculations

1. Download relevant Agcensus data at 2km*2km resolution.

2. Reclassify livestock data into the 8 main ChREAM livestock categories i.e. 'dairy herd', etc.

3. Centre the coordinates. The grid coordinates supplied with Agcensus data relate to the south west corner of each grid square. By adding 1000 to both eastings and northings the grid coordinates then apply to the centre of the 2km*2km squares.

4. Divide the totals of the livestock categories by 4.

5. Add Agcensus data to the project as a .dbf file (1). Display as x,y data. Add a 1km * 1km mesh (2) to the project. Add the catchment shapefile (3) to the project.

6. Intersect the mesh (2) with the catchment shapefile (3). Output = 'mesh intersect' (4). Edit 'mesh intersect' (4) attribute table. Create a field 'mesh_ha' and set field value equal to 100. (Each mesh square = 100 hectares).

7. Spatially join the Agcensus .dbf file (1) to 'mesh intersect' (4). Give the 1km mesh the attributes of the point closest to its boundary. (Therefore, 4 * 1km squares at one quarter of the original livestock value equal the original livestock value.) Export as 'livestock_mesh' (5).

8. Clip 'livestock_mesh' (5) by CREH catchment shapefile (3). Export as shapefile (6). Union shapefile (6) with CREH catchment shapefile (3). Export as shapefile (7).

9. Edit shapefile (7) attribute table. Create fields 'area_m^2' and 'area_ha'. Calculate 'area_m^2', then 'area_ha' as m^2/10000. This calculates the area, in hectares, of each unioned area.

10. Create a field 'proportion', where 'proportion' = 'area_ha'/'mesh_ha'.

11. Create 8 new fields (i.e. 'dairy population', etc.) to calculate the proportion of each livestock category contained in each unioned area. Calculate 'dairy population' as 'proportion'*'dairy herd'.

12. Dissolve shapefile (7) using the catchment shapefile (3) HRU identifier as the dissolve field. Set the statistics field to each of the eight livestock population fields i.e. 'dairy population', and the statistics type of each to 'SUM'. The output will be the total populations of each of the eight livestock types contained in the areas demarcated by the shapefile (3) catchment HRU boundaries. Export the attribute table.

13. Edit the attribute table within a spreadsheet. Add four columns for the remaining ChREAM categories; 'indoor pigs', 'outdoor pigs', 'indoor fowl' and 'free-range fowl'. Set 'indoor pigs' and 'outdoor pigs' as 70% and 30% of 'total pigs' respectively. Set 'indoor fowl' and 'free-range fowl' as 90% and 10% of 'total fowl'. These proportions are based on national averages calculated by the National Pig Association and Defra (Posen, 2008).

1.4.13 Other variables that influence riverine FIO concentrations

Following the calculation of human and livestock population profiles for each subcatchment, further independent variables on *E. coli* inputs were created using the data on *E. coli* production for different animal types in Jones and White (1984), where *E. coli* input = number of each livestock type (km^{-2}) × mean *E. coli* output for each livestock type. Composite variables combining different livestock types and combining human population with livestock, were also created.

Soil type, moisture content and moisture retention are known to affect FIO survival and, more importantly for the present study, their transport through the soil into receiving watercourses via hydrological processes (Deeks et al., 2005; Hagedorn et al., 1978; Jenkins et al., 1984). The use of one of the hydrological properties of soil, namely soil runoff, as an explanatory variable with national coverage, was explored. Data from the Institute of Hydrology's HOST database (Boorman et al., 1995) at a grid-square resolution of 1*1 km was prepared, using the proportional procedures described previously, to calculate the mean SPR for each HRU. The SPR categories corresponded to those described in Table 4.17 of Boorman et al. SPR is the percentage runoff derived from rainfall event data, adjusted to standard rainfall and catchment conditions, and averaged for a subcatchment. It was hypothesised that differences in soil runoff may help to explain riverine FIO concentrations. However, soil runoff was insufficiently significant to warrant inclusion in the models. Reasons for this are explored within the discussion of this chapter.

There are many factors that affect FIO concentrations, mobilisation and die-off in soils and receiving watercourses. The aim of this research is to generate transferable predictive models using readily available data which has national coverage, and this requirement considerably narrows the types of data that can be used. For example, micro- and small-scale processes that affect FIO survival, such as the localised effects of soil pH (Nichols et al., 1983) and antagonism by soil microflora (Gerba et al., 1975) must be excluded from this analysis as nationally available data is unavailable.

There was one other macro-variable that was explored, soil temperature, but this was excluded from the meta-analysis following a preliminary investigation. This variable, and the reasons for its exclusion, is now discussed.

It is widely reported that soil temperature affects FIO dieback rates (Filip et al., 1988; Gerba et al., 1975). Soil Temperature data were sourced from the MIDAS database (Met Office, 2010) and prepared using the proportional procedures described above. The majority of the catchments used within this research are relatively small. Because of this there were only very small variations in mean soil temperatures across those catchments. Given that all of CREH's FIO samples were collected during the narrow temperature range of the summer bathing season and that FIOs are known to be able to survive in a wide range of temperatures outside of the gut (Jones, 1999), it was apparent that coarse resolution soil temperature data would be an insignificant determinant of riverine FIO concentrations. Soil temperature, as an independent variable within this analysis, was rejected.

1.4.14 Statistical analysis

This section of the chapter describes the statistical methods used to produce the predictive models, discusses the rationale governing the selection of the predictor variables used to develop those models and describes the procedures used to assess the transferability of the models.

The generic approach to FIO modelling uses stepwise selection multiple regression techniques to model the relationships between GM FIO concentrations at base- and high-flow (the dependent variables, y) and the various independent variables (x) listed in Table 7. These variables were entered into a Predictor Variable Matrix (PVM), the construction of which is described fully in the following chapter. Statistical analysis was undertaken using SPSS v15.0 for Windows (SPSS Inc., 2006).

FIO concentrations in contaminated water have a normal distribution when $\log_{10}$ transformed (Kay et al., 2005a). Therefore, $\log_{10}$ transformations were applied to those independent variables for which skewness exceeded 1.00. Because the data is transformed, the GM has a greater validity as a measure of central tendency than the more commonly used arithmetic mean (Bishop, 1966; Kay et al., 2008a). Both $\log_{10}$ transformations and the GM are typically used within the field of catchment microbial dynamics. For examples, see the analyses within the review of epidemiological studies on health effects from exposure to recreational water by Prüss (1998). FIO enumerations are expressed as CFU 100ml^{-1}. The geometric mean (GM, calculated as: GM = 10^x, where x = the mean of $\log_{10}$ transformed values) concentration is used to characterise microbial water quality under base- and high-flow conditions for each sampling point.

As with previous FIO studies conducted by CREH (Crowther et al., 2001; Stapleton et al., 2008), relationships of the following form were generated: $y = a + b_1x_1 + b_2x_2 + ... + b_ix_i + e$ where a is the intercept (y at $x = 0$), b is the slope (change in y per unit change in x) and e is the random error term. The regression analysis in this study was parameterised as follows: independent variables with a variance inflation factor > 5 (i.e. tolerance, 0.200) were excluded to minimise multicollinearity (Rogerson, 2001); the level of significance for a predictor variable to enter a model was set at 0.05; the level of explained variance was assessed using the coefficient of determination (r^2 adjusted for degrees of freedom, expressed as a percentage); and the normal probability plot of standardised residuals was examined to confirm the validity of each model. All statistical tests were assessed at the 95% confidence level (Crowther et al., 2011).

Previous CREH studies have found 'urban' and 'grassland' land use variables to be significant FIO predictor variables. Logically, our *a priori* belief was that we could expect close correlation between land use and population variables, namely human with urban and dairy with grassland as land use variables are essentially surrogates explaining patterns of population due to their spatial relationships with those populations.

Three sets of regression models were developed using the independent predictor variables detailed in Table 7, on p.70. These were (1) models using all variables and, to avoid any occurrence of multicollinearity, models that used either (2) land cover or (3) population variables. In the last two cases only those variables that consistently offered significant explanation of the sources of FIOs were included. Specifying the models in these ways allowed the most parsimonious models to be identified and enabled the models to be integrated within the ChREAM and Land Use Allocation Model (LUAM) land-use change models, as will be reported in the next chapter (Fezzi et al., 2008; Jones and Trantor, 2008).

1.4.15 Out of sample testing

A programme of out-of-sample testing was undertaken to evaluate the extent to which the models are truly generic and transferable to other UK catchments (Bateman et al., 2009; Tashman, 2000). In order to minimise the effects of unexplained variance in the models, attention focused on the model that provided the highest level of explained variance. This model was re-run seven times with data for one of the seven catchment studies with ≥ 5 subcatchments omitted in turn, so that the hold-out sample was large enough to be representative (Nau, 2005).

The resulting models were then used to predict the GM concentration for subcatchments in the omitted catchments (termed “test catchments”), and the mean error (predicted–actual concentration) and mean absolute error (absolute difference between predicted and actual concentration) were calculated for each study catchment. The mean error provides a measure of whether a model is under- (+ve values) or over-estimating (−ve values) GM FIO concentrations within the test catchments. In cases where the mean error and mean absolute error have the same value, then the GM concentrations for all of the subcatchments in the test catchment are either over- or under-estimated. As a further independent check, this model was applied to three sampling points in a further catchment (the Haverigg catchment, Cumbria), which was monitored in 2008.

Table 7: independent variables used to predict variations in $\log_{10}$ geometric mean FC and EN concentrations under base- and high-flow conditions, and selected summary statistics for the 153 subcatchments

Variable (unit of measurement)	Code	Mean	Minimum	Maximum	Std. Dev.
Population and population-related variables					
Human population (no. km^{-2})	HUMAN*	206.0	2.1	2,178.3	398.4
Dairy cattle (no. km^{-2})	DAIRY*	22.9	0	84.1	21.0
Other, *i.e.*, non-dairy, cattle (no. km^{-2})	OTHCATTLE*	44.7	0.2	109.3	29.0
All cattle (no. km^{-2})	ALLCATTLE*	67.6	0.2	189.5	47.5
Sheep (no. km^{-2})	SHEEP	292.8	6.7	953.5	206.1
E. coli input: all livestock (cfu km^{-2} h^{-1})	LSTOCKEC	2.4×10^{11}	5.1×10^{9}	7.4×10^{11}	1.5×10^{11}
E. coli input: humans + livestock (cfu km^{-2} h^{-1})	TOTALEC	2.6×10^{11}	1.0×10^{10}	7.4×10^{11}	1.5×10^{11}
Land use variables					
Temporary/permanent grassland (%)	GRASSLAND	45.66	0.45	88.33	23.92
Rough grazing (%)	RGRAZING*	18.69	0	85.33	19.50
Arable, including set-aside (%)	ARABLE*	12.39	0	81.14	19.72
Woodland (%)	WOODLAND*	11.84	0.38	65.26	12.85
Urban/suburban (%)	URBAN*	8.74	0	56.88	10.67
Other variables					
Subcatchment area (km^2)	AREA*	71.19	5.01	1013.18	137.49
Base-flow runoff (m^3 km^{-2} h^{-1})	LFRUNOFF	49.47	2.43	196.33	38.99
High-flow runoff (m^3 km^{-2} h^{-1})	HFRUNOFF	216.56	7.90	845.51	170.91
Total runoff (m^3 km^{-2} h^{-1})	TOTRUNOFF	83.61	3.94	211.88	60.85
Standard percentage runoff (SPR)[c]: all land (%)	TOTSPR	40.28	22.47	59.41	40.28
Standard percentage runoff (SPR): grassland (%)	GRSPR	38.39	18.38	58.44	6.64

* indicates $\log_{10}$ transformation applied in cases where this reduces skewness of data set

1.5 Results

Results of differently specified models are now reported. First, the results of models that use a mix of all variables. These are followed by models that use only land-use variables and then models that use only population-based variables. Finally, results of transfer testing are presented. Descriptions of variables used in the modelling are presented below.

Table 8: variables used within FIO modelling

ALLCATTLE	All cattle population density (no. km^{-2})
ARABLE	Arable land including set-aside (% of total subcatchment area)
AREA	Total subcatchment area (km^2)
DAIRY	Dairy cattle population density (no. km^{-2})
GRASSLAND	Temporary/permanent grassland (% of total subcatchment area)
GRSPR	Standard Percentage Runoff, grassland only (%)
HFRUNOFF	High rainfall river runoff (m^3 km^{-2} h^{-1})
HUMAN	Human population density (no. km^{-2})
LFRUNOFF	Low rainfall river runoff (m^3 km^{-2} h^{-1})
LSTOCKEC	*E. coli* input: all livestock sources (CFU km^{-2} h^{-1})
OTHCATTLE	Non-dairy cattle population density (no. km^{-2})
RGRAZING	Rough grazing (% of total subcatchment area)
SHEEP	Sheep population density (no. km^{-2})
TOTALEC	*E. coli* input: humans and livestock sources (CFU km^{-2} h^{-1})
TOTRUNOFF	Total river runoff (m3 km^{-2} h^{-1})
TOTSPR	Standard Percentage Runoff, all land (%)
URBAN	Urban and suburban land (% of total subcatchment area)
WOODLAND	Woodland (% of total subcatchment area)

1.5.1 All variable models

The results, summarised in Table 9, show statistically significant ($p < 0.05$) base- and high-flow regression models for both FC and EN.

Table 9: summary of results of stepwise multiple regression models of relationship between log_{10} geometric mean FC and EN concentrations at base- and high-flow and all of the independent variables

Step	Variable	Sign of b^a	Adjusted r^2	Significance level (p)
Base-flow models (n = 153)				
Faecal coliforms (FC)				
1	HUMAN	+	0.363	
2	DAIRY	+	0.418	
3	ALLCATTLE	−?	0.481	
4	ARABLE	+?	0.505	
5	AREA	+?	0.518	<0.001
Enterococci (EN)				
1	URBAN	+	0.294	
2	DAIRY	+	0.325	
3	ALLCATTLE	−?	0.369	
4	ARABLE	+?	0.394	<0.001
High-flow models (n =134)				
Faecal coliforms				
1	DAIRY	+	0.439	
2	HUMAN	+	0.595	
3	TOTALEC	+	0.631	
4	OTHCATTLE	−?	0.653	<0.001
Enterococci				
1	DAIRY	+	0.388	
2	HUMAN	+	0.598	
3	TOTALEC	+	0.632	<0.001

? indicates that the sign does not conform with prior expectation.

The levels of explained variance are higher in the two high-flow models than the base-flow models. In each case at least three independent variables were entered. With the exception of AREA, which is entered at Step 5 in the base-flow FC model, all the variables entered are either population- or land cover-related variables. Runoff during the study period and soil hydrology (SPR) were insufficiently significant to warrant inclusion in the models.

Overall, the models are dominated by the population variables, particularly HUMAN and DAIRY, though some land cover variables are also significant. DAIRY is entered first in the high-flow models, whereas HUMAN or URBAN are entered at Step 1 in the base-flow models. For all of the more significant predictors the sign of the slope (b) value is consistent with prior expectations. However, some of the less significant variables (labelled "?" in Table 9) have unexpected effects.

1.5.2 Land cover based models

Table 10 outlines four statistically significant regression models with URBAN and GRASSLAND being the only two variables entered.

Table 10: results of stepwise multiple regression models of relationship between $\log_{10}$ geometric mean FC and EN concentrations at base- and high-flow and the land use variables

	Step	Variable	Sign of b	Adjusted r^2	Significance level (p)
Base-flow models (n = 153)					
Faecal coliforms (FC)					
	1	URBAN	+	0.339	
	2	GRASSLAND	+	0.388	<0.001
Enterococci (EN)					
	1	URBAN	+	0.294	
	2	GRASSLAND#	+	0.301	<0.001
High-flow models (n =134)					
Faecal coliforms					
	1	GRASSLAND	+	0.316	
	2	URBAN	+	0.540	<0.001
Enterococci					
	1	URBAN	+	0.331	
	2	GRASSLAND	+	0.571	<0.001

Only URBAN is entered with PIN = 0.05. In order to include GRASSLAND (the key agricultural FIO-source variable), PIN was relaxed to 0.12. (PIN represents the criteria for variable selection, i.e. the probability of the variable to enter the stepwise regression within SPSS (Probability IN). The default value is 0.05.)

Rough grazing, the other potentially significant FIO source, proved insufficiently significant to be included. While these models inevitably have lower levels of explained variance than those including all potential predictors, it is notable that all the models generated using land cover variables conform to prior expectations, with both URBAN and GRASSLAND land cover types being significant.

1.5.3 Population based models

Within these models the composite variables LSTOCKEC and TOTEC were excluded in order to remove the inevitable overlap with the individual population variables. The results, shown in Table 11, highlight the importance of HUMAN and DAIRY, which are entered at Steps 1 and 2 in all four models.

Table 11: results of stepwise multiple regression models of relationship between log_{10} geometric mean FC and EN concentrations at base- and high-flow and the population variables

Step	Variable	Sign of *b*	Adjusted r^2	Significance. level (p)
Base-flow models (n = 153)				
Faecal coliforms (FC)				
1	HUMAN	+	0.363	
2	DAIRY	+	0.418	<0.001
Enterococci (EN)				
1	HUMAN	+	0.290	
2	DAIRY	+	0.311	
3	SHEEP	−?	0.326	<0.001
High-flow models (n =134)				
Faecal coliforms				
1	DAIRY	+	0.439	
2	HUMAN	+	0.595	
3	SHEEP	+	0.622	<0.001
Enterococci				
1	DAIRY	+	0.388	
2	HUMAN	+	0.598	
3	SHEEP	+	0.624	<0.001

? indicates that the sign does not conform with prior expectation.

HUMAN is entered first in the base-flow models, whereas DAIRY is the key variable at high-flow. SHEEP is also entered at Step 3 in three of the models: with a positive b value in the two high-flow models, though with a counter-intuitive negative b value (for EN) at base-flow. The levels of explained variance are notably higher in the high-flow models.

SPR values for individual soil types in the UK typically range between 2 – 60%. There was much less variation in CREH subcatchments at the resolution of 1*1km. As we can see in Table 12, the mean SPR for the individual catchments ranges between 24.1 – 57.7%, with a mean for all catchments of 40.2%.

Table 12: mean SPR of CREH subcatchments

Catchment	Mean SPR (%) within subcatchments
Holland Brook	41.3
Ribble	41.4
Staithes Beck	38.7
Lake Windermere	45.2
River Leven/Crake	37.4
Sandyhills	39.6
Troon coastal inputs	24.1
Killoch Burn	37.3
River Irvine/Garnock	44.0
Ettrick Bay	40.9
River Nairn	38.0
Afon Ogwr	40.2
Afon Nyfer	36.2
Afon Rheidol/Ystwyth	41.7
Mean SPR	**40.2**

1.5.4 Inter-study transfer errors

Transfer errors were investigated for the high-flow, population-based EN model, since this had the highest level of explained variance of the more parsimonious land cover- and population-based models[4].

Table 13: inter-study transfer errors[a] in the high-flow population-based EN model

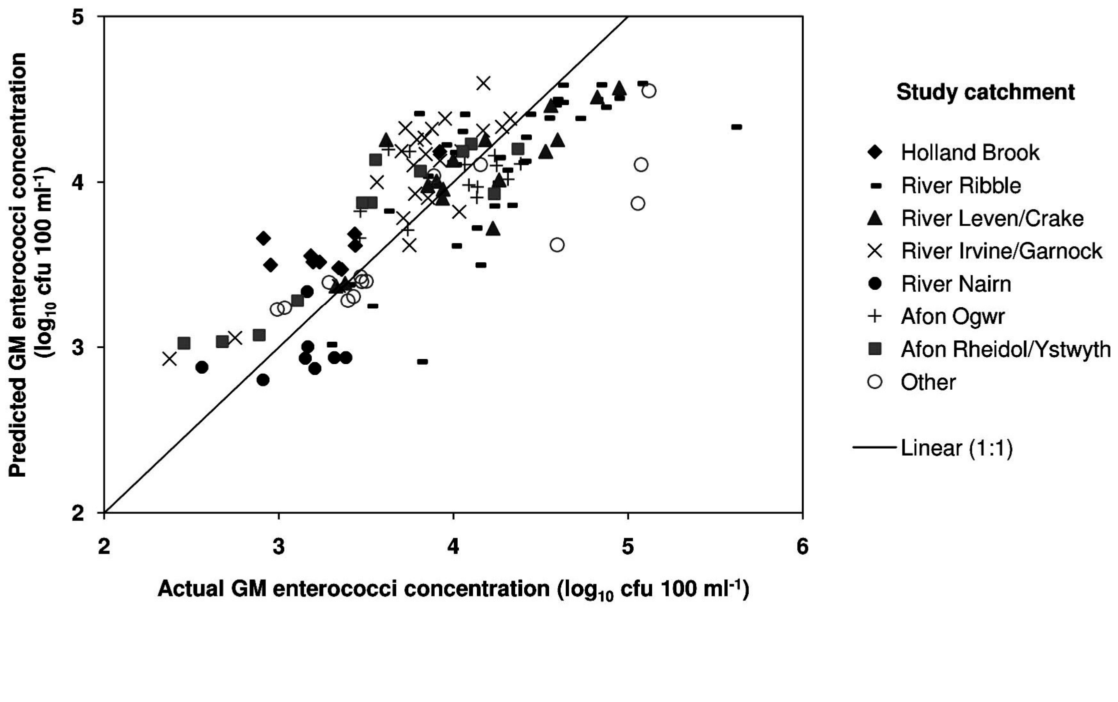
	Study catchment tested[b]	Mean error[c] ($\log_{10}$ CFU 100ml^{-1})	Mean absolute error[d] ($\log_{10}$ CFU 100ml^{-1})
1	Holland Brook	0.4975	0.4975
2	River Ribble	−0.1883	0.2985
5	River Leven/Crake	−0.0513	0.2215
9	River Irvine/Garnock	0.6227	0.6227
11	River Nairn	−0.2126	0.3059
12	Afon Ogwr	0.0417	0.2116
14	Afon Rheidol/Ystwyth	0.4609	0.4912
	Mean	**0.1672**	**0.3784**

[a] Determined by deriving a model with data for the tested study catchment omitted and using the resulting model to predict the geometric mean concentration for subcatchments in the omitted study; [b]Only study catchments with ≥ 5 subcatchments with valid high-flow data were included, see Table 4; [c]Mean of predicted–actual $\log_{10}$ EN concentrations for each of the subcatchments in the study catchment being tested; [d] Mean of absolute difference between predicted and actual $\log_{10}$ EN concentrations for each of the subcatchments in the study catchment being tested.

[4] Other models were not assessed. The high-flow population-based FC model has a very similar explained variance (r^2=0.622) so it is likely that transfer errors for that model will be broadly similar. Transfer testing was not undertaken on the base-flow population-based models. It is acknowledged that transfer errors may differ in those models. The remaining models (i.e. those using land use variables) were not assessed, as those models were inferior to the population-based models and were not used within the transfer analyses reported in Chapter 2.

The results in Table 13 reveal inter-study variability that is not accounted for by the model. Only the Leven/Crake and Ogwr studies have mean errors close to zero. For the Holland Brook, Irvine/Garnock and Rheidol/Ystwyth studies the models based on the other study catchments tend to overestimate the actual EN concentrations that were recorded (mean errors: 0.4975, 0.6227 and 0.4609 $\log_{10}$ CFU $100\mathrm{ml}^{-1}$, respectively); whereas for the Ribble and Nairn studies the models tend to underestimate actual EN concentrations (mean errors: −0.1883 and −0.2126 $\log_{10}$ CFU $100\mathrm{ml}^{-1}$, respectively).

Figure 4: plot of actual high-flow $\log_{10}$ GM EN concentration against predicted values using the population-based model reported in Table 11, with values from those studies showing clear +ve or –ve anomalies from transferability testing (Table 13) identified.



The mean absolute error recorded is 0.3784 $\log_{10}$ CFU $100\mathrm{ml}^{-1}$, with values ranging from 0.2116 (Ogwr) to 0.6227 (Irvine/Garnock) $\log_{10}$ CFU $100\mathrm{ml}^{-1}$. The pattern in these results is closely reflected in the plot of predicted against actual high-flow EN concentrations based on the overall model, shown in Figure 4. Application of the model to the three sites in the Haverigg catchment produced a mean error of $-0.1810 \log_{10}$ CFU $100\mathrm{ml}^{-1}$ (ranging from -0.0513 to $-0.2467 \log_{10}$ CFU $100\mathrm{ml}^{-1}$). It should be noted that inter-study transfer errors will tend to be greater where levels of explained variance in the models are lower, notably in the base-flow models.

A linear assessment of transfer errors fitted very badly (i.e. was an inappropriate functional form). Figure 4 shows a $\log_{10}$ transformed plot of the values predicted by the high-flow GM enterococci population based model (reported in Table 11) versus actual GM enterococci concentrations. The models underperform when predicting outliers: the model is tending to over-predict very low concentrations and under-predict very high concentrations. The primary reason for this is because the models are compiled using data from several river sites (each of which have differing ambient characteristics), which impacts on the fit of predicted concentrations. This transfer error may result in less accurate predictions of extreme values, but does not prevent the models from predicting generalised patterns of pollution or identifying potential 'pollution hotspots' that may require further investigation. A more sophisticated non-parametric approach to model construction may have resulted in more accurate predictions but this option was not explored as the parsimonious nature of the parametric approach provides tractable models which aids model transfer.

The models reported here represent the first generic transferable models which can be used to predict FIO pollution across unmonitored UK watercourses. They are the first models for which any transferability testing has been undertaken (so transfer errors cannot be directly compared to previous research): e.g. the Scotland and Northern Ireland Forum for Environmental Research (SNIFFER, 2006) screening tool provides insights into FIO export coefficients for catchments (in only Scotland and Northern Ireland, not elsewhere within the UK) but does not provide a basis for characterising base- and high-flow FIO concentrations separately, and the SNIFFER export coefficient calculations have yet to be fully evaluated against out of sample tests or data from monitored catchments. The immediate predecessors to this research – the meta-analyses conducted by CREH and reported in Kay et al. 2008a and 2008b – did not provide any assessment of transferability (e.g. out of sample analyses). Both of the Kay et al. studies are qualitatively different from this research. Kay et al. 2008a is not directly comparable as it focused exclusively on FIOs from sewage and treated effluents (i.e. no FIOs from agricultural sources). Comparison with Kay et al. 2008b is not possible as it focused on the significance of differences in FIO export coefficients (cfu km^{-2} h^{-1}) between base-flow and high-flow river flow conditions. Furthermore the results of Kay et al. 2008b are somewhat obfuscated by examining the relationship between overall (i.e. base and high flow) FC export coefficients whereas within this research base- and high-flow are modelled separately and FIOs are expressed as concentrations (CFU $100ml^{-1}$), not export coefficients.

1.6 Discussion

All of the regression models clearly identify both humans and livestock as key FIO sources within catchments. It should be noted that some FIOs from both sources, especially particle-attached FIOs, may be deposited on the stream bed under base-flow conditions and re-suspended at times of high-flow. The FIO concentrations reported are therefore derived from both newly entrained and newly added organisms into the water column. Indeed, a significant proportion of the elevated concentration at high-flow may well be from the stream bed as increased water velocities increase water turbidity, entraining deposited sediments and the stream-bed store of FIOs (Wilkinson et al., 2006).

Under base-flow conditions human sources (as reflected in the HUMAN and URBAN variables) are more important than livestock sources in accounting for the observed variance in FC and EN concentrations. Indeed, the DAIRY or GRASSLAND variables that are entered at Step 2 in the base-flow regression models provide only very limited additional explanation. This suggests that sewage-related sources are dominant at base-flow, with relatively little FIO input from agricultural sources. The former will be largely treated effluents from WwTWs, which generally have much lower FIO concentrations under base-flow conditions than high-flow (Kay et al., 2008a). The relatively low levels of explained variance in the base-flow models probably reflects the fact that in this 'black box' modelling, no account is taken of the nature of the effluent quality discharged by individual WwTWs, which varies with the type of treatment (Kay et al., 2008a); and also that the URBAN and HUMAN data for individual subcatchments will poorly reflect the magnitude of sewage effluent inputs to the subcatchment watercourses in cases where WwTWs serving a significant proportion of the built-up area are located downstream of the monitoring point (i.e., sewage is exported out of the subcatchment for treatment). It is also interesting to note that the HUMAN and URBAN variables provide very similar levels of explained variance, which suggests that, for the purpose of catchment-scale modelling, built-up land is a relatively good proxy for human population.

At high-flow both human and livestock sources assume importance, with the latter generally being the more dominant. Under such conditions some untreated sewage from combined sewerage overflows or overflows from WwTW storage tanks is likely to be discharged to watercourses, and the quality of treated effluents from many WwTWs will be reduced due to more rapid transmission through the plant (Kay et al., 2008a). The importance of human sources is evidenced by the inclusion of URBAN and HUMAN as key variables in the various high-flow models.

The general importance of livestock sources at high-flow is reflected in the land cover-based models by the prominence of GRASSLAND, which is entered first for FC and makes a major contribution to the explained variance achieved for EN (Table 10). This is in keeping with previous studies which have shown that the dominant sources of FIOs at high-flow tend to be of agricultural origin (Stapleton et al., 2008). It should be noted that the GRASSLAND land use category comprises all temporary/permanent grassland, other than that which is mapped as rough grazing. As such it encompasses quite a wide range in terms of quality and productivity, extending from very fertile lowland pastures, which tend to be dominated by dairy farming, up to quite high altitudes in some subcatchments, where beef and sheep production systems tend to dominate. Because of this, GRASSLAND is simply a proxy variable for the more intensive areas of livestock production. Consequently, land cover data, as are traditionally used in FIO modelling, inevitably have limited explanatory power and potential for scenario modelling. By incorporating livestock density data, the present study provides insight into the relative significance of different production systems. Of the various livestock variables used in the modelling (Table 7), DAIRY emerges consistently as the key variable, with levels of explained variance that are consistently higher than GRASSLAND. In the case of the high-flow FC models, for example, the DAIRY has an r^2 value of 0.439, compared with 0.316 for GRASSLAND, which clearly highlights the importance dairy farming systems (*cf*. beef cattle and sheep) as a FIO source. This presumably reflects the high intensity of most dairy farming operations, which tend to be largely confined to the better land in the lowlands; the concentration of animals close to farm buildings for milking; and the storage and disposal to land of large quantities of waste (mostly in form of slurry) from yard areas and indoor winter housing–all of which pose potential pollution risks in terms of both diffuse sources (e.g., faeces voided directly in fields and slurry/manure applications to land) and point-source pollution (e.g., runoff from farmyards and milking parlours, slurry stores and manure heaps). By contrast, beef and sheep systems are not so confined to the better land, are often less intensive, and generate smaller amounts of waste for disposal. Sheep may, however, be present in quite large numbers in some catchments, both in areas of temporary/permanent grassland and rough grazing. They therefore represent a potentially significant FIO source, and this is reflected in SHEEP being entered at Step 3 with a +ve b value in both high-flow population-based models (Table 11). On the basis of these results, the design and implementation of measures to address FIO pollution from agricultural sources should be targeted initially on areas of dairy production.

Several non-source variables (Table 7) were included in the all-variable modelling. These relate to three catchment characteristics that may affect source strength and the mobilisation, transport, die-off and sedimentation of FIOs within catchments, namely: runoff volume, soil hydrology and catchment size.

Volume of runoff during the study period may be an important factor since, during prolonged periods of wet weather, certain FIO sources (especially those associated with diffuse sources, such as animal faeces in fields and stream source contributory areas) will tend to become depleted. It might be anticipated, therefore, that a period of high-flow will tend to be associated with higher FIO concentrations if preceded by a long spell of dry weather than if it followed a relatively wet period. Due to differences in weather conditions between the 6–8 weeks of each of the 14 catchment studies, there is very marked inter-subcatchment variability in runoff volumes (e.g., TOTRUNOFF: range, 3.94–211.88 m^3 km^{-2} h^{-1}) (Table 7). In the case of soil hydrology, in subcatchments with more poorly drained soils (*i.e.*, with a higher mean SPR) there will likely be more surface runoff per unit rainfall and hence increased mobilisation and transport of FIOs from land to adjacent watercourses, which may well lead to increases in FIO concentrations. In the present study the SPR for both the subcatchments as a whole (TOTSPR: range, 22.47–59.41%) and for the areas of permanent/temporary grassland (GRSPR: range, 18.38–58.44%) were used as predictor variables (Table 7). Catchment size may also be an important factor, since the opportunity for die-off of FIOs along watercourses as a result of exposure to UV light is increased within larger catchments as a result of the greater length of channel flow. This is particularly likely under base-flow conditions when flow velocities, water depth and turbidity are all at a minimum, thereby maximising UV exposure. The 153 subcatchments used in the modelling range in size (AREA) from 5.01–1,013.18 km^2 (Table 7).

Despite the marked inter-subcatchment variability of runoff volume, soil hydrology and catchment area, only AREA was entered in any of the models, and that with a (counter-intuitive) +ve b value in the base-flow FC model (Table 9). Clearly, controlled experimental studies are needed to assess more fully the effects of these factors and any interaction effects between them. Given that these variables were either insignificant or produced a counter-intuitive b value, interaction effects were not explored further during this research. On present evidence, it would seem that their role in affecting FIO concentrations in watercourses at the regional and national scales is minor compared with differences in human population density, stocking levels and associated land use types (URBAN and GRASSLAND), –*i.e.* those factors that relate directly or indirectly to the key FIO sources.

The out-of-sample testing reveals some degree of inter-study variability in the model evaluated, and this will inevitably tend to be greater in models with lower levels of explained variance, notably the base-flow models. This is not unexpected and is likely to be attributable to a combination of both inter-catchment and temporal factors. The former reflect systematic differences between the catchments affecting the sources, survival and mobility of FIOs that are not accounted for by the variables in the final regression models (*i.e.*, the unexplained variance). For example, there may be inter-catchment variations in livestock farming facilities and management practices that limit the extent to which key predictor variables such as GRASSLAND and DAIRY provide a measure of FIO sources. Also, soil hydrology (as outlined above) seems likely to account for some degree of inter-catchment variability, but its influence is not sufficiently strong to be included in the all-variable models; and other factors that were not included as potential predictor variables (e.g., temperature and topography) are likely to have a similar effect. The temporal factors, on the other hand, reflect the fact that the individual studies were undertaken over 6–8 week monitoring periods with markedly contrasting weather conditions, both before and during the studies; and at different times during the bathing season, which could, for example, affect FIO source strength in grazed fields as a result of the progressive accumulation over the summer months of faeces from dairy cattle (which are housed over winter). Volume of runoff during the individual study periods, which was considered most likely to be the key temporal factor, was included in the predictor variable set, but, as with soil hydrology, was not sufficiently significant to be entered in the all-variable models.

The strength of the present models lies in the fact that they are based on a FIO database that has extensive geographical coverage (land use, climate, topography, soils, etc.) and encompasses a wide range of weather conditions during the individual monitoring periods. Some of the inter-study transfer errors are inevitably quite high, and these are partly attributable to temporal factors. Clearly, by combining the data from all 14 catchment studies the effects of the temporal factors are minimised and the inter-catchment errors reduced. The resulting land cover- and population-based models developed in the present study can therefore be applied with some confidence for predicting base- and high-flow GM FC and EN concentrations during the summer bathing season in UK watercourses with catchments areas between 5 and approximately 1,000 km^2. While the lower size threshold is determined by the level of resolution of the available agricultural census data, the upper limit simply reflects the size of the larger catchments used in the present modelling.

By combining the predicted GM FIO concentrations with discharge data, the contribution that an individual rivers/streams makes to overall FIO loadings to coastal waters can be estimated.

The models can also be used to evaluate the likely impact of different land use/stocking level and human population change scenarios, as might result from the implementation of measures designed to reduce FIO loadings, or reforms in agricultural policy/funding, as reported in Hampson et al. (2010).

1.7 Conclusions

In order to meet European WFD requirements there is an urgent need for transferable models that can accurately predict base- and high-flow GM FIO concentrations in UK watercourses. Previous studies of individual catchments have successfully developed regression models based on relationships between GM FIO concentrations recorded at monitored sites and the land use type within their subcatchments. The present study has extended this approach by combining data from 14 different catchment studies within a meta-analysis to develop generic models and augmenting the predictor variables to include direct measures of key FIO sources (i.e., human population and livestock density data) and various other factors (catchment size, runoff and soil hydrology) that may affect FIO mobilisation, transport and die-off.

Statistically significant base- and high-flow regression models have been developed for both FC and EN, with levels of explained variance consistently higher in the latter models. Population variables (notably HUMAN and DAIRY) generally provide higher levels of explained variance than the land cover variables. Under base-flow conditions human, sewage-related, sources are dominant, whereas livestock sources tend to assume greater significance at high flow, with dairy farming systems (*cf.* beef cattle and sheep) being particularly important sources. Neither runoff, soil hydrology or catchment size were significant predictor variables. In the more parsimonious land cover or population-based models, developed for ease of transferability to other UK catchments, relatively high levels of explained variance were achieved for all of the high-flow models, with adjusted r^2 values ranging from 0.540 (land use model for FC, Table 10) to 0.624 (population model for EN, Table 11).

A programme of out-of-sampling testing on the high-flow EN model indicated some degree of inter-study variability, which is likely attributable to a combination of: (i) inter-catchment factors, which reflect systematic differences between the catchments that affect the sources, survival and mobility of FIOs that are not accounted for by the variables in the models; and (ii) temporal factors, which reflect the fact that the FIO monitoring was undertaken under different weather conditions and at different times during the summer bathing season. However, it is argued that by combining data from all 14 studies, which have a wide geographical distribution across UK and encompass a wide range of weather conditions, the effects of the temporal factors are minimised and the inter-catchment errors reduced.

The resulting land cover- and population-based models can be used, with some confidence, in UK catchments to predict base- and high-flow FC and EN concentrations in unmonitored watercourses and to evaluate the likely impacts of different land uses, livestocking levels and human population change scenarios. In so doing, these models help provide valuable insights into the key sources of FIOs at catchment scale and can therefore help inform the development of policies and the prioritisation of investments to reduce microbial pollution, given that a mix of cost-effective regional and site specific policy remediation strategies will be required to achieve the highest reductions. This theme is further explored in the next chapter.

1.7.1 Limitations and potential improvements to the research within Chapter 1

The data in several of the CREH catchments are nested and consequently violate the Gauss-Markov assumption of no autocorrelation within the residuals (Gujarati, 2003). As they stand the models are misspecified. Although not traditionally used in FIO studies, multi-level modelling techniques have been used within the field of epidemiology to develop models which explicitly incorporate hierarchies or levels within which data is clustered (Goldstein, 1995; Duncan et al., 1993). These methods may be appropriate given the nested properties of the catchment data and may enable models to be developed which control for contextual effects within catchments.

The majority of CREH catchment studies have been undertaken in rural catchments and tend to underrepresent highly populated urban areas. Therefore the data underpinning the models may cause systematic inaccuracies when the models are used to extrapolate to areas of very high population density.

The assessment of transfer errors reported in this chapter was performed in two ways: firstly, the high-flow population-based EN model was re-run seven times with data for one of the seven catchment studies with ≥ 5 subcatchments omitted in turn, so that the hold-out sample was large enough to be representative.

Secondly, the model was applied to three sampling points from the Haverigg catchment (a catchment from which no empirical data was used to construct the models). Although transfer errors were assessed for the high-flow population-based EN model, transfer errors in the other models were not assessed. Future improvements to this research could include analyses of transfer errors in the other models, to assess the ways in which transfer errors may differ across models (i.e. population-based vs. land use-based) and across flow rates (i.e. high-flow vs. base-flow). Further improvements to the testing of transfer errors could include using the models to assess FIO concentrations at alternative monitored watercourses and comparing predicted vs. actual values. Ideally, transfer errors for watercourses within the Humber RBD should be assessed and alternative watercourses should be subject to a wide range of site-specific geographic and climatic conditions (ideally including watercourses in the south and east of the UK and in highly urbanised areas, as these locations are underrepresented in the models) to test how well the models perform at different locations. Transfer testing should also be undertaken at different times during the year to assess temporal variability as the majority of the data underpinning the models was collected during the summer bathing season.

The models may not show the best association between water quality and the impact of humans. There are two main sources of error in the models that are associated with the vagaries of the sewerage network. Firstly, errors are introduced because the models cannot account for sewage being piped across HRU boundaries and this has the potential to cause large errors in FIO enumeration. Within the models FIO discharges are always attributed to the originating HRU, but, in reality, this does not always happen. For example, the Minworth sewerage treatment works processes the sewage from 2.5 million people, the majority of which is piped from Birmingham via the Black Country Trunk Sewer system (Severn Trent Water, 2005). These transboundary flows cause the models to overestimate concentrations emitted from source HRUs (i.e. Birmingham) and underestimate concentrations in receiving HRUs (Minworth). Secondly, the models fail to recognise the relative efficiency of different sewage treatment types. Tertiary wastewater treatment methods are more effective than secondary or primary methods (WHO, 2003; Kay et al., 2008a) and this impacts on riverine FIO concentrations. These errors may be reduced by identifying the sewage infrastructure catchment areas, discharge points and treatment types, and then incorporating this additional data into the models. The EA database of discharge sites may be a potential source of some of the required data (Harley, 2008).

There are a number of other variables that could be explored in an attempt to explain more of the variation in the models. For example, a composite variable, based on human population density and urban land use area, could be used. This approach is used by the Office for National Statistics when trying to minimize the misclassification of electoral districts (2002). A similar composite of dairy and grassland could be created, where the grassland variable is adjusted for the number of cattle present.

It has been calculated that *E. coli 0157* is endemic within 1-15% of UK dairy herds, with distinct regional variations in infection rates (Jones, 1999). By incorporating data on the locations where *E. coli* is endemic, the variables for *E. coli* inputs used within this research may be improved.

The ability of the riverine environment to assimilate microbial pollution is poorly understood. Many organisms have remarkable survival rates (Burton et al., 1987; Ogden et al., 2002) and precise characteristics of their entrainment and deposition are unknown (Wilkinson et al., 2006). Turbidity has been positively correlated with FIO concentrations (Wilkinson et al., 1995; Lawler et al., 2006), as have other factors such as gradient, slope shape, stream proximity, soil moisture (Fraser et al., 1998), FIO inactivation, transportation through soils (Vinten et al., 2004), temperature and in-stream mobilization (Tian et al., 2002). If these variables could be generated, on a consistent national scale, they may warrant further investigation and possible inclusion in the models described here.

2 Chapter 2: Predicting Microbial Pollution Concentrations in UK Rivers in Response to Land Use Change

2.1 Introduction

The previous chapter introduced the shift towards the integrated management of recreational water quality through the development of drainage basin wide programmes of measures, prompted by the Water Framework Directive (WFD) (EU, 2000). The WFD and its daughter directives, the revised Bathing Waters Directive (rBWD) (EU, 2006a) and the Shellfish Waters Directive (EU, 2006b), have increased the need for cost-effective diagnostic tools capable of accurately predicting riverine faecal indicator organism (FIO) concentrations and fluxes (Kay et al., 2007a and 2008b). The last chapter described the construction of models designed to fulfil that requirement. This chapter demonstrates several ways in which those generic regression models can be applied: to predict riverine FIO concentrations within unmonitored watercourses, to produce a Quantitative Microbial Risk Assessment (QMRA) of the water within those watercourses and to quantify the likely impact on FIO concentrations (and health risk) following the implementation of hypothetical land use policy measures designed to reduce faecal pollution. These assessments are made at both River Basin District (RBD) and catchment scale.

This chapter begins by outlining, in terms of health impacts and external economic costs, the need for models capable of apportioning the sources of excessively high FIO concentrations in watercourses. The literature review then examines the dose-response relationship between FIO concentrations and ill-health, before providing a critique of the capabilities and limitations of QMRA, a statistical tool used to numerically simulate and improve estimations of the risk of ill health from the use of recreational water contaminated with faecal pollution (Pond, 2005).

Following a statement of the aims and objectives, the methodology section describes in detail the methodology used to transfer the model to the Humber RBD and a sub-catchment scale example is used to illustrate how the models are applied. An overview of the mathematical and conceptual structures of the econometric and linear programming models used to generate land management scenario data is also provided.

The results are in two sections. Firstly, the FIO models are used to predict FIO concentrations in the Humber River Basin District during base- and high-flow conditions, in the summer bathing season, 2004. The resulting patterns of microbial pollution are presented and interpreted. The EU and the WHO provide slightly different guideline values for recreational water quality. Nevertheless, the models used within this research produce results compatible with both: QMRAs (guided by WHO compliance values) for different land use change scenarios in the Humber RBD are undertaken and the impact on water quality of those scenarios is also examined with reference to the EU guideline values.

Secondly, and because the FIO models incorporate explanatory variables which allow the effects of policy measures which influence livestock stocking rates to be assessed, the effects of seven land use management and policy instruments (fiscal constraint, production constraint, cost intervention, area intervention, demand-side constraint, input constraint, and micro-level land use management) are modelled. All of these scenarios are qualitatively very different from one another but all have the potential to reduce microbial pollution in rivers. An assessment is made of the relative effectiveness of these microbial pollution remediation strategies.

The discussion examines the results. This is followed by the conclusions which highlights some of the challenges faced by the policy and management communities in devising suitable strategies to reduce riverine microbial pollution. The conclusion also contains a discussion identifying some of the limitations of, and provides suggestions for potential improvements to, the research design.

As this chapter demonstrates, the diagnostic tool reported here can provide significant insights which aid microbial source apportionment and help to identify watercourses at elevated risk of pollution.

2.2 Literature review

Given the relevance of the literature review in the previous chapter, this review is specific in its focus and, in order to remain concise and relevant, it briefly discusses the issues relevant to each of the three main tasks (health risk assessment, FIO source apportionment and land use scenario modelling) covered within this chapter.

Unidirectional FIO discharges impose a variety of uncompensated external costs on a variety of downstream users (Pearce and Turner, 1990). These costs include the degradation of recreational water quality and drinking water supplies and unacceptably high levels of pollution received by shellfish harvesting waters. Microbiological pollution is potentially best prioritised in terms of its health impacts and external economic costs (Larsen and Ipsen, 1997). Microbiological pollution can cause a variety of illnesses ranging from nausea and diarrhoea to, very occasionally, more serious illnesses which can, very rarely, result in death.[5]. Illnesses arising from exposure to polluted bathing water have large associated healthcare costs, estimated at an annual $12 billion globally (Shuval, 2003). Dwight et al. (2005) estimated a public cost of $3.3million per year as a result of illnesses acquired from localized contaminated water at just two Californian beaches. In the UK, Mourato et al. (2003) calculated that c.1.3 million excess cases of gastroenteritis each year may be attributable to poor bathing water quality.

The positive correlation between microbial concentrations in recreational water and increased ill health has been established through epidemiological studies of recreational water and its adverse effects on recreational users (Kay et al., 1994; US EPA, 2003; WHO, 2003). Such studies support the idea that the rate of infection and disease among recreational users increases steadily with increasing concentrations of harmful microorganisms within a dose-response relationship (Ferley et al., 1989; Fleisher et al., 1996; Kay et al., 1994; Prüss,

[5] The Centers for Disease Control and Prevention (CDC) maintain the Waterborne Disease and Outbreak Surveillance System for collecting and reporting waterborne disease and outbreak related data. During the 2001-2002 reporting year 65 waterborne disease outbreaks associated with recreational water were reported by 23 US states. These 65 outbreaks caused illness among an estimated 2,536 persons; 61 persons were hospitalized, eight of whom died (CDC, 2004).

1998)[6]. The UK Randomised Controlled Trials (UKRCT) for gastroenteritis (Kay et al., 1994) and acute febrile respiratory illness (Fleisher et al., 1996) were the two key studies identified by the WHO as providing the most accurate and unbiased data from which to compile the *Guidelines for Safe Recreational Water Environments* (2003). Both studies found significant dose-response relationships between EN and ill health and the slopes of the dose-response curves in both studies were broadly consistent. Following those trials, the WHO adopted intestinal enterococci (EN) concentrations as the most suitable health criterion for both marine and freshwater environments (WHO, 2003) and the EU adopted FC and EN concentrations as the main parameters to measure compliance with the rBWD (Environment Agency, 2008b; EU, 2006a).

Table 14: EU Bathing Water Directive inland water quality compliance values (EU, 2006a)

	Water Quality Classification		
Organism	**Excellent***	**Good***	**Sufficient****
EN (CFU $100\ \mathrm{ml}^{-1}$)	200	400	330
FC (CFU $100\ \mathrm{ml}^{-1}$)	500	1000	900

Notes: * based on 95th percentile compliance, ** based on 90th percentile compliance. Guidelines based on geometric mean values of EN concentration are unavailable. It is acknowledged that the geometric mean values calculated in the FIO models do not correspond exactly with percentile values.

This research is guided by both the WHO and the EU rBWD compliance values. A range of scenarios are modelled in order to examine how land use changes have the potential to improve the quality of river water quality. These scenarios are assessed with reference to the EU compliance values shown in Table 14. But, as those values have no associated health risk, the QMRAs performed within this chapter refer to the WHO QMRA compliance parameters, described below.

QMRAs are used to predict infection or illness rates from given concentrations of particular pathogens, assumed rates of ingestion and the most appropriate dose-response models for the population exposed (Haas et al., 1999). However, due to the limited empirical data upon which the WHO guideline values are based, there are four key areas in which data is lacking, each of which are now briefly

[6] i.e. the rate of certain enteric and respiratory infections and disease among bathers, compared with unexposed non-bathers, increases steadily with increasing concentrations of indicator microorganisms of faecal pollution.

discussed. It is acknowledged that each of these issues have the potential to affect the validity of the QMRA results presented later within this chapter.

For ethical reasons the subjects of the UKRCT were healthy adult volunteers (Kay et al., 2004a). Diseases that are normally mild can have severe outcomes in susceptible sub-populations, e.g. those with weakened immune systems, the infirm, the elderly or young, pregnant women, etc. (Carr and Bartrum, 2004). Several studies suggest that illness rates are higher for children (Cabelli, 1983), the elderly and the infirm (Prüss, 1998). Consequently, the WHO guidelines may systematically underestimate risks to these groups.

To simulate bathing conditions the subjects in the UKRCT studies bathed for a minimum of ten minutes and during that period immersed their heads three times (Kay et al., 1994). Although broadly representative of the actions performed when bathing, this does underestimate the risks associated with increased time in the water or higher risk activities. There is a growing body of evidence that longer exposure to polluted water leads to increased probability of illness (Bradley and Hancock, 2003). Philipp et al. (1985) and Dwight et al. (2004) found significantly higher rates of ill health in snorkel swimmers and surfers. Similarly, Fewtrell et al. (1994) found that health risks decrease with lower exposure, and lower risk, recreational activities.

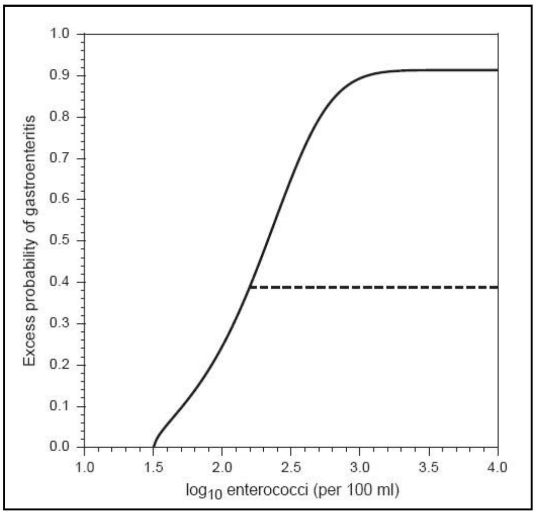
This next limitation has the potential to introduce a significant systematic error into the accuracy of freshwater QMRAs based on the WHO guidelines. Dufour (1984) suggested that the risk of ill health from sea water bathing may be twice that of freshwater bathing. A comparison of the data produced by Kay et al. (1994) and Ferley et al. (1989), although using different methodologies, suggests that illness rates are five times higher in sea water. One explanation is that some FIOs, particularly *E. coli*, may die off more rapidly in sea water than in freshwater, resulting in higher concentrations of harmful pathogens in seawater when FIO densities are identical (WHO, 2003)[7]. If this is the case then the application of guidelines derived for seawater would result in lower rates of illness in freshwater users. This phenomenon is acknowledged within the rBWD (EU, 2006a), which

[7] The use of FIOs as surrogates for the presence of other harmful microorganisms is discussed in Chapter 1.

provides inland water quality parameters that are twice as high as coastal water parameters.

The fourth limitation of the WHO QMRA guidelines concerns the mathematical form of the dose-response relationship used by the WHO, shown in Figure 5.

Figure 5: dose-response relationship in the 1994 sea-bathing trials (Kay et al., 2004a)



The solid line represents the mathematical form of the dose–response relationship. The dotted line represents the functional form used in the derivation of the WHO Guideline values.

The maximum concentration of EN detected in the UKRCT was 158 CFU 100 ml^{-1} (2.198 $\log_{10}$ CFU 100 ml^{-1}) (Kay et al., 1994). To avoid extrapolating the dose-response curve into areas of no data, the WHO guidelines assume that the excess probability of ill health (0.388) remains constant above the level of 158 CFU 100 ml^{-1} (i.e. above the level of the dotted line), rather than continue to increase as the mathematical relationship suggests (Kay et al., 2004a). This assumption may lead to gross underestimates of the risks posed by highly contaminated water. For a full explanation of the probability density function of the disease burden assessment method please see Kay et al., 2004a.

Because of the relative scarcity of evidence in the four areas described above, the WHO adopted a single series of microbial values, for both coastal and fresh water, irrespective of the type of recreational activity. These values are shown in Table 15.

Table 15: microbial water quality assessment categories and associated probability of gastrointestinal illness (WHO, 2003)

Water quality assessment category	95th percentile value of enterococci per 100 ml (rounded values)	Basis of derivation	Estimated risk of gastroenteritis per exposure
A very good	< 40	This value is below the no-observed-adverse-effect level in most epidemiological studies	<1%. An average probability of 1 case in 100 exposures
B good	41-200	The 200/100 ml value is above the threshold of illness transmission reported in most epidemiological studies that have attempted to define a lowest-observed-adverse-effect level	1-5%. An average probability of 1 case in 20 exposures
C fair	201-500	This level represents a substantial elevation in the probability of all adverse health outcomes for which dose–response data are available	5-10%. An average probability of 1 case in 10 exposures
D poor	> 500	Above this level, there may be a significant risk of high levels of minor illness transmission	>10%. There is a greater than 1 in 10 chance of illness per single exposure

The WHO classification underpins the QMRA assessments undertaken in this research because, in contrast to the rBWD classification, the WHO guidelines contain clearly defined health risks. This alone lends greater policy relevance to the QMRA results as it provides a greater degree of quantification upon which policy makers can make more informed decisions.

Despite ongoing improvements to wastewater treatment facilities in the UK, noncompliance with microbial guidelines still occurs at many designated bathing and shellfish sites, particularly after high rainfall when there are increased emissions of untreated sewage from combined sewer overflows or wastewater storm tanks (Crowther et al., 2001). As water companies increasingly treat sewage to higher standards, it may be argued by the water industry that a greater proportion of non-compliance may, in the future, be attributable to the agricultural sector (Chadwick et al., 2008).

Poor agricultural practices have the potential to contaminate watercourses with enteric microorganisms, a proportion of which are pathogenic to humans, and this potential is exacerbated within intensively farmed catchments (Crowther et al., 2001; Oliver et al., 2007). Pollution of water sources by agricultural wastes such as cattle slurry has been a major problem in the UK and accounted for 28% of all agriculturally related water pollution incidents between 1987 and 1989 (MAFF, 1991). However, the results of the models developed in the previous chapter demonstrated that the water industry and the agricultural sector are both responsible for FIO emissions. Failure to comply with the water quality parameters of the WFD may result in infraction proceedings being instigated by the European Commission (Kay et al., 2005b) so there is a growing need for the UK regulator to correctly identify FIO sources and apportion liability to either human or agricultural sources. As was discussed in the previous chapter, sources and concentrations of FIOs are dependent on a complex interplay of factors, which creates uncertainty when attempting to apportion liability for faecal pollution sources.

Progress has been made in apportioning liability for FIOs, and other potentially harmful pathogens found in watercourses, to agricultural or urban sources using regression modelling (e.g. Kay et al., 2005a; Crowther et al., 2011) and microbial source tracking techniques (e.g. Stapleton et al., 2007a).

Microbial source tracking is a forensic technique used to identify the source (i.e. human, bovine, etc.) of microbial pollution in watercourses. Methods rely on the identification of signature molecules (markers) such as DNA sequences of host-associated microorganisms. Several techniques are available, e.g., the identification of F+RNA coliphage groups (types II and III are predominantly human and types I and IV are animal-associated) (Havelaar et al., 1990), or the genomic analysis of bacteroidetes to identify the presence/absence of specific genotype markers (i.e. bovine CF128 and human HF183) (Bernhard and Field, 2000). The field of microbial source tracking has advanced and expanded considerably over the last two decades (see, for example, the reviews by Harwood et al. (2014) and Shanks et al. (2016)). However, such analyses are time consuming and typically used for small-scale applications: they would be prohibitively expensive and impractical at catchment scale.

Regression modelling has been a useful tool for identifying potential pollution ‘hotspots’ and modelling spatial trends at the catchment or subcatchment scale. By incorporating livestock and human population density variables into regression models, the predictive models used in this research yield accuracy improvements which may help inform the identification of FIO sources in the environment. For example, by identifying trends in the distribution of FIOs, the models could be used as a starting point to inform more intensive empirical microbial source tracking investigations that determine whether the ‘sterol fingerprint’ of faecal pollution is of human or animal origin (Leeming et al., 1996).

Cuttle et al. (2007) have identified 44 different methods of controlling diffuse pollution from agriculture. The effects of many of these, i.e. maintain or enhance soil organic matter levels, cannot be estimated by this research. However, the dairy and human population variables do lend themselves to modelling the potential effects of different levels of livestocking, human population change or alterations to wastewater treatment infrastructure.

By altering the values of the explanatory variables used within the FIO models, hypothetical land use change scenarios can be simulated to assess the potential impacts on riverine FIO concentrations. For example, if the impact of a 20% reduction in dairy cattle were to be assessed, the dairy stocking density parameter of each target HRU would be reduced by 20% within the model: an HRU which previously had 100 dairy cows per km^2, for the purposes of the scenario modelling, would be assumed to have 80 cows per km^2 and the effect of this change on riverine FIO concentrations can be assessed. The land use management and policy instruments modelled in this research are discussed further within the methodology.

The models used here were developed to inform the Catchment, Hydrology, Resources, Economics and Management (ChREAM) project (Bateman et al., 2006a). The construction of the models was reported in Crowther et al. (2011) and the scenario analyses are detailed in Hampson et al. (2010). As these models are based on readily available data with national coverage, they are capable of accurately predicting base- and high-flow FIO concentrations nationally (Crowther et al., 2011). As such, the modelling approach presented here is of great benefit to the policy and land management communities when planning basin wide or national programmes of measures.

2.3 Aims and objectives

This chapter utilises the models produced in Chapter 1 with the aim of fully exploring the strengths and weaknesses of the generic models, using a range of land use and population change scenarios. The aims are to assess the effectiveness of the FIO models as (i) cost effective diagnostic tools capable of aiding source apportionment and (ii) assess the effectiveness of different pollution remediation strategies, (iii) at a range of spatial scales.

In doing so, the following objectives are necessary:

- Develop a transfer methodology, capable of expansion up to the scale of the UK, from which the FIO models can estimate FIO concentrations in watercourses.
- Generate human and livestock population profiles for Humber catchments.
- Create a theoretical sampling point network and a Predictor Variable Matrix (PVM) for the Humber RBD, from which predictions can be made.
- Predict baseline base- and high-flow FIO concentrations of EC and EN FIO indicators within the watercourses of the Humber RBD, for the summer bathing season 2004.
- By altering the values of the explanatory variables, generate datasets for hypothetical land-use and population change scenarios (20% decrease in dairy livestock; 1.4% increase in human population; Mixed effects: 20% decrease in dairy livestock with a 1.4% increase in human population and 5% improvement in wastewater treatment efficiency). Simulate and assess the potential impacts of those scenarios on base- and high-flow FIO concentrations of EC and EN FIO indicators within the watercourses of the Humber RBD, for the summer bathing season 2004.
- Map water quality in terms of EU rBWD inland water quality compliance values for baseline base- and high-flow FIO concentrations of EC and EN FIO indicators within the watercourses of the Humber RBD, for the summer bathing season 2004.
- Assess and compare the estimated baseline EC and EN concentrations in the Humber RBD with the rBWD compliance water quality categories and compare those baseline estimates against the three land use change scenario detailed above.

- Produce QMRAs (using WHO microbial water quality assessment categories) for the baseline base- and high-flow EN concentrations in the Humber RBD and for base- and high-flow EN concentrations for the three land use change scenario detailed above. Compare the effects of the risks to human health arising from the simulations.
- Assess the impact on riverine FIO concentrations across the Humber RBD if the UK government were to adopt a nutrition driven food policy.
- Assess the relative effectiveness of a range of remediation strategies (Taxing fertilizer by £50/tonne; ESA designation in Aire; Increase milk quota cost by £20; Reduce dairy stocking by 20%; Reduce fertilizer application by 20%; Installation of stream bank fencing) against the baseline high-flow FC estimate, at subcatchment scale (Aire subcatchment).
- Simulate and map the reductions in high-flow FC concentrations resulting from stream bank fencing erected in intensive milk producing HRUs in the Aire subcatchment.

2.4 Methods

2.4.1 Outline of the method for generating the generic FIO model

To summarise the key findings of the previous chapter, the statistical models used here predict geometric mean (GM) FC and EN concentrations under both base- and high-flow conditions using human and livestock population density data as explanatory variables. These population based models represent the first transferable generic FIO models to be developed for the UK to incorporate direct measures of key FIO sources (namely human and livestock population data) as predictor variables. Under high-flow conditions, which are the critical times in terms of FIO loadings in watercourses (due to a combination of increased FIO concentrations and volumes of flow), levels of explained variance of up to 62.2% have been achieved for FC and up to 62.4% for EN (see Table 16). Both discharge volumes and FIO concentrations are approximately an order of magnitude higher at high-flow and, in line with previous studies (Wither et al., 2005; Kay et al., 2008a), high-flow models show a greater level of explanatory power for both FC and EN.

Table 16: the models used to predict FC and EN at base- and high-flow within the transfer exercise, QMRA assessments and scenario modelling.

River discharge type	Intercept	Primary Coefficient	Secondary Coefficient	Tertiary Coefficient	r^2 (Adj.)
GM faecal coliforms (FC) (Log_{10} CFU 100 ml^{-1})					
High-flow	+**	+Log_{10}Dairy/km^2**	+Log_{10}Human/km^2**	+Sheep/km^2**	0.622
Base-Flow	+**	+Log_{10}Human/km^2**	+Log_{10}Dairy/km^2**	-	0.418
GM enterococci (EN) (Log_{10} CFU 100 ml^{-1})					
High-flow	+**	+Log_{10}Dairy/km^2**	+Log_{10}Human/km^2**	+Sheep/km^2**	0.624
Base-flow	+**	+Log_{10}Human/km^2**	+Log_{10}Dairy/km^2*	-	0.311

Notes: +/- indicates the sign of the coefficient. ** indicates strong significance at p<0.01, * = significance at p<0.05. All coefficients have low standard errors and high t-values. Values of beta coefficients are confidential, see Appendix I for details.

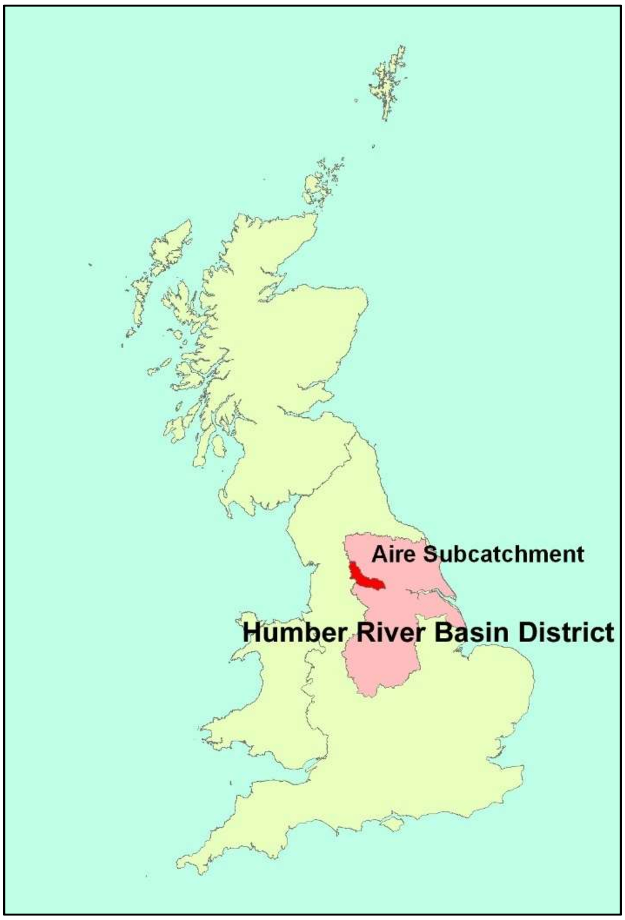
At high-flow the FC and EN models are dominated by dairy sources. Previous studies have shown that the dominant sources of FIOs at high-flow tend to be of agricultural origin when large quantities of manure or slurry are washed off fields and farmyard hardstandings into rivers (Stapleton et al., 2008). The quantity of human derived FIOs is also elevated at high-flow but discharges and spills are dependent on the capacity and efficiency of waste water treatment works (WwTWs). In contrast, humans are the most significant source of FIOs at base-flow (Table 16). This is probably due to constant background inputs from WwTWs. A full account of the development of the FIO models underpinning this analysis is described in the previous chapter and reported in Crowther et al. (2011).

2.4.2 Study areas

One change prompted by the WFD was the introduction of the River Basin District (RBD) planning system to enable integrated catchment management strategies. Given that the FIO models used here were expressly developed to predict FIO concentrations in watercourses for which detailed FIO sampling has not taken place, the transfer site for the present study is the Humber RBD which covers 26,000km^2 from Birmingham to the North York Moors (Appleton, 2008). This is the largest RBD in the UK, draining 28% of the land surface of England into the North Sea via the Humber estuary. Varied physical landscape characteristics coupled with diverse livestock farming operations and extremes of human population density make the Humber RBD an ideal case study area within which to predict riverine FIO concentrations at RBD scale.

The FIO models have also been used to predict FIO concentrations at the smaller, subcatchment scale. For these assessments the Aire subcatchment of the Humber RBD has been used. The Aire subcatchment covers 1100km^2 and is an ideal case study for smaller scale applications of the FIO models, as the River Aire passes through three distinct land use types before its confluence with the River Calder in the Aire Calder subcatchment (Environment Agency, 2010a). The River Aire rises from the relatively clean waters of the Malham Tarn glacial lake in the southern Yorkshire Dales before receiving livestock derived FIOs as it flows through areas of intensive dairy farming to the west and south of Skipton. Further downstream the river receives large quantities of FIOs from the WwTWs serving the urban conurbations of Bradford and Leeds (CaBA, 2016). The relative scale of the Aire subcatchment and the Humber RBD are shown in Figure 6.

Figure 6: locations of the Humber River Basin District and the Aire subcatchment



2.4.3 Humber catchment boundary data

The Humber catchment boundary polygons were supplied to the ChREAM project (Bateman et al., 2006a) by the Centre for Ecology and Hydrology (CEH). They were produced using a hydrological digital terrain model, based on a 50 m grid interval (NERC, 2009). To be consistent with other aspects of ChREAM modelling, the spatial information defining Hydrological Response Unit (HRU) boundary polygons used in this study are configured so that the major Humber subcatchments have an EA river monitoring point at their outlet (Hutchins, 2008a). As the generic FIO models cannot be applied to tidal areas the Humber Estuary subcatchment has not been modelled.

2.4.4 Theoretical sampling points

To conduct the transfer exercise a network of theoretical sampling points is devised. Within the context of this research the term 'sampling point' simply means a location on the river network at which predictions of FC or EN concentrations are made. Sampling sites are 'theoretical' in that they do not correspond spatially with the locations of the actual EA river monitoring network sites. There are several reasons why this is both necessary and desirable, which are now discussed.

The EA river monitoring network does not correspond with the boundaries of the smaller HRUs within the Humber RBD (Hutchins, 2008b). In this respect, the use of actual EA monitoring locations would have been limiting because the FIO models require that HRU sample points can only be located at their downstream exit as sample points must include all water draining from that HRU. Fortunately, the use of empirical EA sampling data (or the actual locations of EA monitoring stations) was unnecessary for the present modelling exercises because actual river discharge data is not required for the models to function: the FIO models themselves are underpinned by discharge data and are calibrated to predict GM FIO concentrations for both base- and high-flow antecedent runoff conditions (Crowther, 2008b). The fact that discharge data from EA monitoring stations was not needed provided greater flexibility when choosing the locations of sampling points. Subsequently, a network of 613 theoretical water sampling points, capturing data from 988 HRUs from the 18 subcatchments within the Humber RBD, was devised to achieve a balance between two main sampling objectives. The first requirement was for an evenly distributed network of sample points along the length of the main rivers to provide an overview of water quality as it becomes progressively aggregated downstream. The second sampling objective targeted tributaries upstream of the confluence with the main rivers to capture diverse variability in water quality relating specifically to those tributaries.

The sampling scheme can be adjusted to meet specific sampling objectives as required in the future. For example, if there was a need for intensive sampling in urban tributaries the sampling network can be amended to accommodate this.

2.4.5 Livestock and human population data

Human and livestock population data for each of the HRUs within the Humber RBD were constructed using the methods described in the previous chapter. 2004 was used as the base year for two reasons. Firstly, this was the most recent year for which agricultural survey data were available at the time that the modelling was undertaken and, secondly, 4 of the 15 CREH catchment datasets, underpinning the FIO models, use water quality data from 2004 (see Table 4 in Chapter 1) and, within the meta-analysis, 6 of the 15 CREH catchment sites use livestock populations derived from 2004 agricultural survey data (see Table 6 in Chapter 1). Table 17 shows the variables used to calculate human and livestock population profiles.

Table 17: data sources used in the transfer analysis and scenario modelling

Study area	Number of theoretical sampling points	Agricultural census data used for livestock enumeration	Census year used for population enumeration
Humber RBD	613	England 2004	2001

2.4.6 Transfer methodology

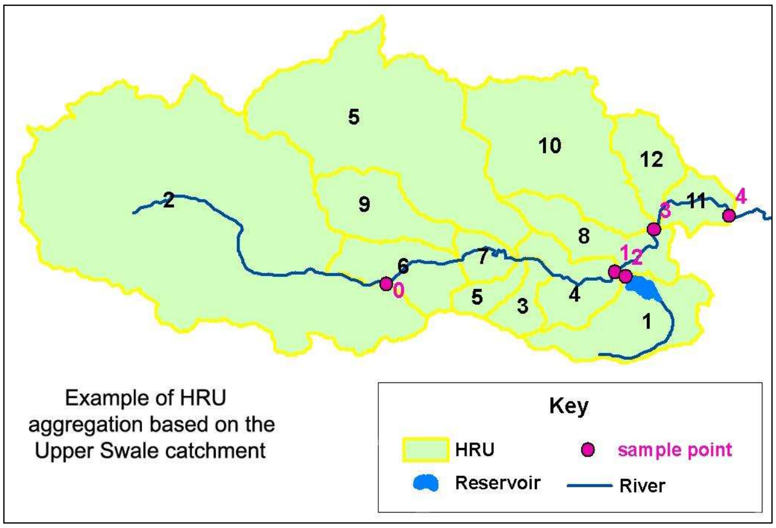
A transfer methodology enabling the models (Table 16) to quantify predicted riverine FIO concentrations in the UK employs the same algorithms used to generate the models. The only difference is that the predictor variable matrices (PVMs) of the models use empirical FIO concentrations to calculate model parameters, whereas the transfer PVMs use the models' parameters to predict FIO concentrations. The next section of this chapter provides a worked example to comprehensively describe how the transfer methodology is applied.

2.4.7 An overview of the subcatchment used within the worked example

Figure 7 shows Humber's Upper Swale catchment. The catchment contains 12 HRUs. Minor tributaries to the River Swale have been omitted for clarity. The headwater is in HRU 2 in the west and the river exits the catchment via HRU 11 in the east. The catchment contains 5 sample points, numbered 0-4. Sample point

0 is relatively close to the headwater and sample point 4 is at the catchment exit. Within Figure 7 we see that HRU 1 contains a major tributary which passes through a reservoir[8] immediately before its confluence with the Swale. Sample point 2 is located at the reservoir exit, above the confluence of the two rivers.

Figure 7: map of the Upper Swale catchment, used within the example. Relevant topographic features and the locations of the theoretical sample points are shown



The PVM is composed of three tables: (1) an HRU data table, (2) an aggregation table and (3) a sample point data table. Each of the three tables for the worked example are now described.

2.4.8 The example subcatchment's HRU data table

Table 18 shows the example's HRU data table. The table shows the predicted numbers of humans and dairy livestock within each HRU and the area of each HRU. The table also provides a column which shows the proportion of each HRU that drains into a reservoir. The issues surrounding FIO attenuation rates due to reservoir catchments are discussed within Section 1.4.3 in Chapter 1. If an HRU contains a reservoir the population values for that HRU and the area of the HRU are adjusted to account for the attenuation of FIOs by that reservoir (Stapleton

[8] In order to provide a simple illustration of the way in which reservoir catchments are treated within the PVM a hypothetical reservoir was added into the catchment. This exemplar attenuates FIOs from 100% of the water emitted at the HRU exit sampling point, enabling a straightforward calculation of predicted FIOs.

and Kay, 2007b).If 100% of the HRU drains into the reservoir (as is the case with HRU 1), the human and dairy populations of that HRU are reduced to zero to simulate FIOs from those sources being attenuated by the reservoir. If the reservoir is higher up in the HRU, so that 50% of the area of the HRU drains into the reservoir, the population values are reduced by 50% to simulate the removal of 50% of human and dairy FIOs. This method is subject to interpolation errors, but, as reservoirs are typically in upland areas of low population density, the errors that do arise tend to be negligible. The area of the HRU is also adjusted: if 100% of the watercourses within the HRU drain through the reservoir, the non-reservoir area is set to zero (e.g. HRU 1 in Table 18). If 50% of the HRU drains into the reservoir, the size of the HRU is reduced by 50%.

Table 18: the example PVM HRU data table

HRU	Populations		% of HRU draining into reservoir	Populations adjusted for reservoirs		non-reservoir area (km^2) exiting the HRU
	Human	Dairy		Human	Dairy	
1	322	76	100	0	0	0
2	540	60	0	540	60	151.47
3	30	8	0	30	8	5.83
4	89	135	0	89	135	13.38
5	452	29	0	452	29	73.37
6	111	45	0	111	45	15.22
7	320	57	0	320	57	5.61
8	89	108	0	89	108	15.57
9	192	18	0	192	18	16.30
10	243	32	0	243	32	42.88
11	162	102	0	162	102	7.43
12	38	62	0	38	62	11.28

2.4.9 The example subcatchment's aggregation table

The aggregation table is shown in Table 19. It simply tells us which HRUs drain into each sample point. So, by examining Table 19, it can be seen that HRU 2 drains into sample point 0; HRUs 2 - 7 and 9 drain into sample point 1, and so on. Sample point 4 at the subcatchment exit contains data from all 12 HRUs in the subcatchment.

Table 19: the example PVM aggregation table

Sample point	HRUs aggregated within each sample point
0	2
1	2, 3, 4, 5, 6, 7, 9
2	1
3	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
4	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

2.4.10 The example subcatchment's sample point data table

Table 20 shows the third, and final, PVM data table: the sample point data table. This table takes the data for each HRU (Table 18, HRU data table) and aggregates that data into each sample point according to the aggregation rules defined by the aggregation table (Table 19).

Table 19: the example PVM sample point data table

sample point	Total population		Proportion reservoir/ non-reservoir		Populations adjusted for reservoir catchments		Non-reservoir area of catchment draining into sample point (km^2)	Log_{10} population densities (km^2)		Model coefficients			Non-reservoir output: predicted coliform concentration		Reservoir output: predicted coliform concentration	Total output: (Reservoir plus non-reservoir) predicted coliform concentration
	human	dairy	Proportion reservoir	Proportion non-reservoir	human	dairy herd		log_{10} human	log_{10} dairy	intercept	beta human	beta dairy	log_{10} CFU 100 ml^{-1}	CFU 100 ml^{-1}	(CFU 100 ml^{-1})	(CFU 100 ml^{-1})
0	540	60	0	1	540	60	151.47	0.66	0.14	2	1	1	2.80	637	83	637
1	1733	351	0	1	1733	351	281.17	0.86	0.35	2	1	1	3.20	1612	83	1612
2	322	76	1	0	0	0	0	-	-	2	1	1	-	0	83	83
3	2387	567	0.060	0.940	2065	491	339.61	0.85	0.39	2	1	1	3.24	1732	83	1632
4	2588	730	0.057	0.943	2266	654	358.32	0.86	0.45	2	1	1	3.31	2069	83	1955

Note: arbitrary coefficients are for illustration purposes only (please see Appendix I).

The first column displays the sample point. The next two columns show the aggregated total populations within each sample point. The next two columns indicate the proportion of the aggregated land area that is either reservoir or non-reservoir. The next two columns adjust the populations according to the impact of reservoirs, based on the rules described above. Note that the populations of sample points 2, 3 and 4 have all been adjusted to compensate for the water emitted from the reservoir in HRU 1, which enters the sampling network at Sample Point 2. The adjusted human population in the entire catchment is 2266 and the adjusted dairy population is 654. The next column aggregates the non-reservoir area of the catchment draining into each sample point. The total non-reservoir area of the entire catchment is 358.32km^2. The next two columns convert the adjusted human and livestock populations of non-reservoir areas into population densities, expressed as Log_{10} densities per km^2 (e.g. $\mathrm{LOG}_{10}((540/151)+1)$ for the human population for Sample Point 0).

The next three columns provide the model's intercept term and coefficients (anonymised within this example), from which predictions of the coliform concentrations of non-reservoir areas can be made.

The regression equation used to predict FIO concentrations takes the following form[9]:

$\mathrm{Y}\ (\log_{10}\ \mathrm{CFU}\ 100\ \mathrm{ml}^{-1})$ = Intercept + (b human*$\log_{10}$human/km^2) + (b dairy*$\log_{10}$dairy/km^2).

Equation 1: the regression model used to predict FIO concentrations

Y for each sample point is displayed in the next column. The next column shows the linear value for non-reservoir watercourse coliform concentrations expressed as CFU $100\ \mathrm{ml}^{-1}$, transformed from the $\log_{10}$ value (e.g. 637 $\mathrm{CFU}\ 100\ \mathrm{ml}^{-1}$ at Sample Point 0).

[9] Although the quantified parameters are not published within this thesis (or within peer reviewed journal articles) due to their commercially sensitive nature, they have been seen by the examiners of this PhD. The reasons for this confidentiality are discussed in the Author's declaration.

Coliform concentrations emitted from reservoirs are now calculated. Reservoir FIO output concentrations are determined depending on FIO type and river flow conditions. Reservoir output values used in this research are reported in Table 5. For this example the value for high-flow FC is used (i.e. 83CFU $100\ \mathrm{mL}^{-1}$).

The final column provides the total FIO concentration, derived from a mix of reservoir and non-reservoir sources (i.e. (proportion of non-reservoir land*non-reservoir output)+(proportion of reservoir land*reservoir output)). For Sample Point 4, at the subcatchment exit, this is (0.943*2069)+(0.057*83), or 1955CFU $100\ \mathrm{mL}^{-1}$.

As mentioned previously, providing that each sampling point is at the exit of an HRU, the sampling scheme can be adjusted to meet specific sampling objectives. Within the example, sample point 0, in the headwater, and sample point 2, in a tributary, provide more diverse results than those obtained from the progressively aggregated sampling points 1, 3 and 4 along the main river. There is no reason why sampling points could not, for example, have been created at the exits of different HRUs, in order to capture the FIO concentrations within tributaries to the Swale, if that were the sampling objective.

2.4.11 Outline of methods for generating data for land use management strategies

The PVMs make predictions of FIO concentrations in the Humber RBD using human population and livestock population variables. These independent variables can be adjusted, as required, to estimate FIO concentrations for a variety of land use scenarios at both base- and high-flow.

Cuttle et al. (2007) propose a 1:1 reduction in FIO emissions resulting from policy measures to reduce dairy cattle stocking density rates. To assess the impact of a 20% reduction in dairy livestock, the dairy population of all HRUs was reduced by 20%. The riverine FIO concentrations which may arise are modelled.

According to the Ministry of Agriculture, Fisheries and Food (MAFF) (2000) many farms are over-fertilizing. In line with Defra (2004), a 20% cut in all fertilizer application is assumed across all farming activities, including all grassland fields. By impacting on grassland productivity, this measure is designed to encourage producers to switch away from dairy farming to more extensive activities. Assumptions within this scenario are based on the changes in livestock populations associated with a 20% cut in fertilizer application: the expected change in dairy livestock numbers is -8.98% and sheep numbers is -10.00% (Defra, 2004). The parameters in the FIO models are adjusted accordingly (i.e. for all HRUs, dairy livestock populations are reduced by 8.98% and sheep are reduced by 10%). As Cuttle et al. (2007) indicate, this measure might prompt increases in manure applications but, as in their analysis, this possibility is not considered here.

To model the change in riverine FIO concentrations due to changes in human population, the human populations within each HRU entered into the model are adjusted as required.

Farm profits are determined by a variety of fixed factors (e.g. physical environment); input costs (e.g. fertilizers); output prices (e.g. milk price); subsidies and taxes (e.g. single farm payment); and other factors (e.g. expectations). A highly flexible model, described in detail in Fezzi and Bateman (2009), was estimated and used to generate changes from baseline livestock populations in the Environmentally Sensitive Area (ESA), fertilizer tax and milk quota scenarios. Agricultural Census data was combined with data from the Farm Business Survey to provide agricultural land use and livestock numbers (EDINA, 2008a). Environmental and climatic variables, policy determinants and input and output prices were then added. The profit (π) function associated with the optimal land allocation can be expressed as:

$$\pi(\mathbf{p},\mathbf{w},\mathbf{z},L) = \max_{l_1,\ldots,l_h}\{\pi(\mathbf{p},\mathbf{w},\mathbf{z},l_1,\ldots,l_h) : \textstyle\sum_{i=1}^{h} l_h = L\}$$

Equation 2: the profit (π) function associated with the optimal land allocation

Where p is a vector of output prices, w is a vector of the input prices, z is a vector of other fixed factors, l is the vector of h land use allocations, with L the total land available.

ESAs were introduced in 1987 to safeguard and enhance areas of particularly high landscape, wildlife or historic value. ESA payments encourage switching to extensive grassland types such as permanent grassland and rough grazing (Natural England, 2009a). Within this scenario the effects of designating the entire Humber RBD as an ESA is modelled. It is acknowledged that the ESA scheme has been replaced by Higher Level options of the Environmental Stewardship Scheme (Lobley and Potter, 1998): the focus here is on the use of the ESA concept as an area intervention.

The econometric model (Equation 2) underpinning the fertilizer tax scenario predicts that an increase in fertilizer price by £50/tonne decreases the optimal shares of nutrient-intensive activities. This encourages producers to switch away from intensive activities (e.g. dairy farming) to more extensive activities, such as rough grazing.

The EU milk quota scheme was introduced in 1984 to reduce the imbalance between supply and demand for milk and milk products. It controls milk production and stabilizes milk prices for both consumers and producers. It is acknowledged that milk quota entitlements were withdrawn in 2015: the milk quota is used purely as an example of a quantity restriction policy. It is hypothesized that the additional cost to the farmer of raising the price of the EU milk quota will discourage milk production and the FIO model is used to predict the impact on riverine FIO concentrations arising from the adjusted dairy stocking levels generated in the econometric model.

Farm best management practices (BMPs) can significantly reduce the delivery of FIOs to watercourses and have the potential to be effective and cost efficient: Meals (1996) observed 70% reductions in FIOs from dairy sources after BMPs were adopted. By preventing livestock from voiding directly into watercourses it has been demonstrated that stream bank fencing, as a micro-level policy, is highly successful at reducing microbial pollution received by watercourses (Larsen et al., 1994; Oliver et al., 2007) and may be more beneficial than simply reducing the stocking density of livestock (Vinten et al., 2004).

To simulate the effect of stream bank fencing (erected a minimum distance of 2.13 m from watercourses) attenuating FC in runoff from fields and other farmyard sources, a 95% reduction in dairy and sheep derived FC concentrations entering watercourses is modelled, in line with Larsen et al. (1994). These reductions are applied to the parameters of the FIO model in those HRUs in the Aire subcatchment having dairy cow densities above 25 per km^2, as this threshold captures the region of intensive dairy farming (near Keighley) to the west of the Aire (outlined in red on Figure 20)[10].

The Nutritionally Driven Food Policy scenario investigates the effect on riverine FIO levels of the adoption by the UK population of a healthier diet, e.g. one consistent with Department of Health Guidelines on healthy eating (RELU, 2009). While not strictly a measure designed to reduce diffuse pollution, adoption of this policy would see large reductions in milk (42.2%) and mutton and lamb (28.2%) consumption (Jones et al., 2009) with the positive consequence of reducing microbial pollution discharged to watercourses. The dairy and sheep populations in this scenario are generated using the Land Use Allocation Model (LUAM), which models the decoupling of production from support payments under the reformed EU Common Agricultural Policy and the dietary change-inspired reduction in demand for milk and sheep products. The LUAM is constructed using the General Algebraic Modelling System (GAMS) software package (GAMS Software GmbH, 2008) and the methodology underpinning the LUAM is fully described in Jones and Tranter (2008).

The mathematical structure of the LUAM is that of an ordinary linear programming model, shown in Equation 3.

$$\text{Maximize: } \mathrm{Z} = \mathrm{cx}$$
$$\text{subject to: } \mathrm{Ax} \leq \mathrm{b}$$
$$\mathrm{x} \geq 0$$

Equation 3: the linear programming model underpinning the LUAM

Where Z is the objective function given as the scalar product of c and x vectors, b is the resource endowment and input availability vector, c is the vector whose

[10] The Aire subcatchment contains 47 HRUs. Mean dairy cattle density for the Aire is 12.85 per km^2 (range within the HRUs = min 0 per km^2 – max 42.18 per km^2). 9 of the 47 HRUs had mean values above 25 per km^2. The locations of those HRUs correspond closely with the region of intensive dairy farming near Keighley (Agbotui et al., 2014).

elements are returns and costs, and x is the output vector. A is the matrix of input/output coefficients (aij) representing the amount of input i required per unit of output (j).

In this scenario livestock populations predicted by a 'reference run', defined by the economics of production and the market and the policy environment observed in 2006, are compared against livestock populations predicted by the 'scenario run', which also assumes that the work of changing people's diets in line with the UK Food Standards Agency (2009) guidelines has been completed.

As the FIO models cannot be applied to tidal areas, the Humber estuary subcatchment is excluded from the analysis. ArcGIS v.9.3 (ESRI Inc., 2008) and Microsoft Office Excel 2007 (Microsoft Corporation, 2006) are used to generate and map predicted FIO concentrations for all scenarios.

The results of the transfer function, QMRAs and scenario modelling are now reported.

2.5 Results

The results are reported in three sections. These are summarised in Table 21.

Table 21: list of modelling results

Figure or Table number	Pathogen type	Output scale	Flow type	Notes/scenario type
Section 1: Transfer to the Humber River Basin District				
Figure 8	FC	Humber RBD	base-flow	Transfer exercise
Figure 9	FC	Humber RBD	high-flow	Transfer exercise
Figure 10	EN	Humber RBD	base-flow	Transfer exercise
Figure 11	EN	Humber RBD	high-flow	Transfer exercise
Section 2: Scenario models examining rBWD compliant sites and QMRA results				
Table 22	both	Humber RBD	both	Changes to FIO concentrations
Figure 12	FC	Humber RBD	high-flow	20% decrease in dairy livestock
Figure 13	EN	Humber RBD	base-flow	1.4% increase in human population
Figure 14	FC	Humber RBD	high-flow	Mixed: 20% decrease in dairy livestock, 1.4% increase in human population and 5% improvement in WwTW efficiency
Figure 15	both	Humber RBD	both	rBWD compliance following scenarios 1-3
Figure 16	EN	Humber RBD	base-flow	QMRA
Figure 17	EN	Humber RBD	base-flow	QMRA
Figure 18	EN	Humber RBD	high-flow	QMRA
Figure 19	FC	Humber RBD	high-flow	Adopting the nutrition driven food policy
Section 3: Assessments of water quality at subcatchment scale				
Figure 20	FC	Aire catchment	high-flow	Stream bank fencing in HRUs with high dairy cow populations
Table 23	FC	Aire catchment	high-flow	Relative effectiveness of different remediation strategies

The first section reports the results of the transfer exercise, providing predictions of FC and EN concentrations in the Humber RBD during summer 2004, for both base- and high-flow river conditions.

The results in section two primarily examine the impact on water quality at RBD scale in response to three simple scenarios. These are: (scenario 1) a 20% decrease in dairy livestocking density, (scenario 2) a 1.4% increase in human population, and to demonstrate the ability of the models to model simultaneous changes to more than one independent variable, (scenario 3) a mixed scenario which applies three changes to the independent population variables: a 20% decrease in dairy livestock, a 1.4% increase in human population and a 5% improvement in WwTW efficiency. The effect these scenarios have on the water quality of the Humber RBD in terms of rBWD compliance and the change in risk to human health are assessed. The last assessment within section two examines the impact on the microbial quality of the watercourses of the Humber RBD, if the UK government were to adopt a nutrition driven food policy.

To highlight the versatility of the FIO models, the third section of the results moves from macro scale assessments of the water quality across the Humber RBD to more focussed assessments of changes in water quality within only the Aire subcatchment in response to a range of pollution remediation strategies. Within this section the relative effectiveness of several different remediation, including stream bank fencing is assessed.

2.5.1 Section 1: results of the transfer to the Humber River Basin District

This section describes the results of the transfer exercise predicting FIO concentrations in the Humber RBD during summer 2004. The four maps, shown in Figures 8 - 11, represent estimated concentrations of FC and EN during summer 2004 at base- and high-flow. Green and orange labels represent the sample points predicted to comply with rBWD 'Excellent' and 'Good' water quality categories, shown on Table 14.

Figure 8: predicted FC concentrations $(\mathrm{CFU}\ 100\ \mathrm{ml}^{-1})$ **in the Humber RBD under base-flow conditions, summer 2004**

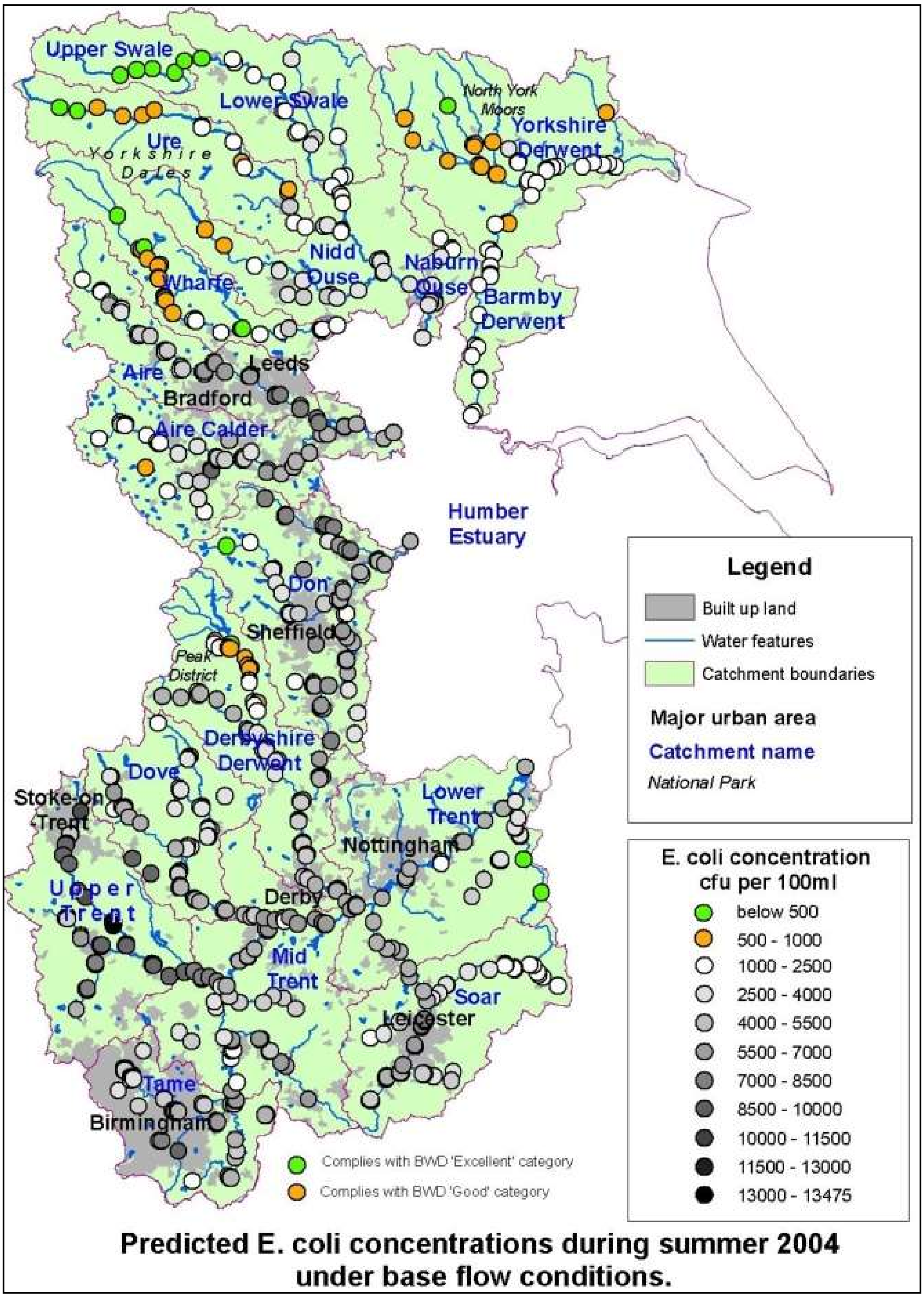


Figure 9: predicted FC concentrations (CFU 100 ml^{-1}) in the Humber RBD under high-flow conditions, summer 2004

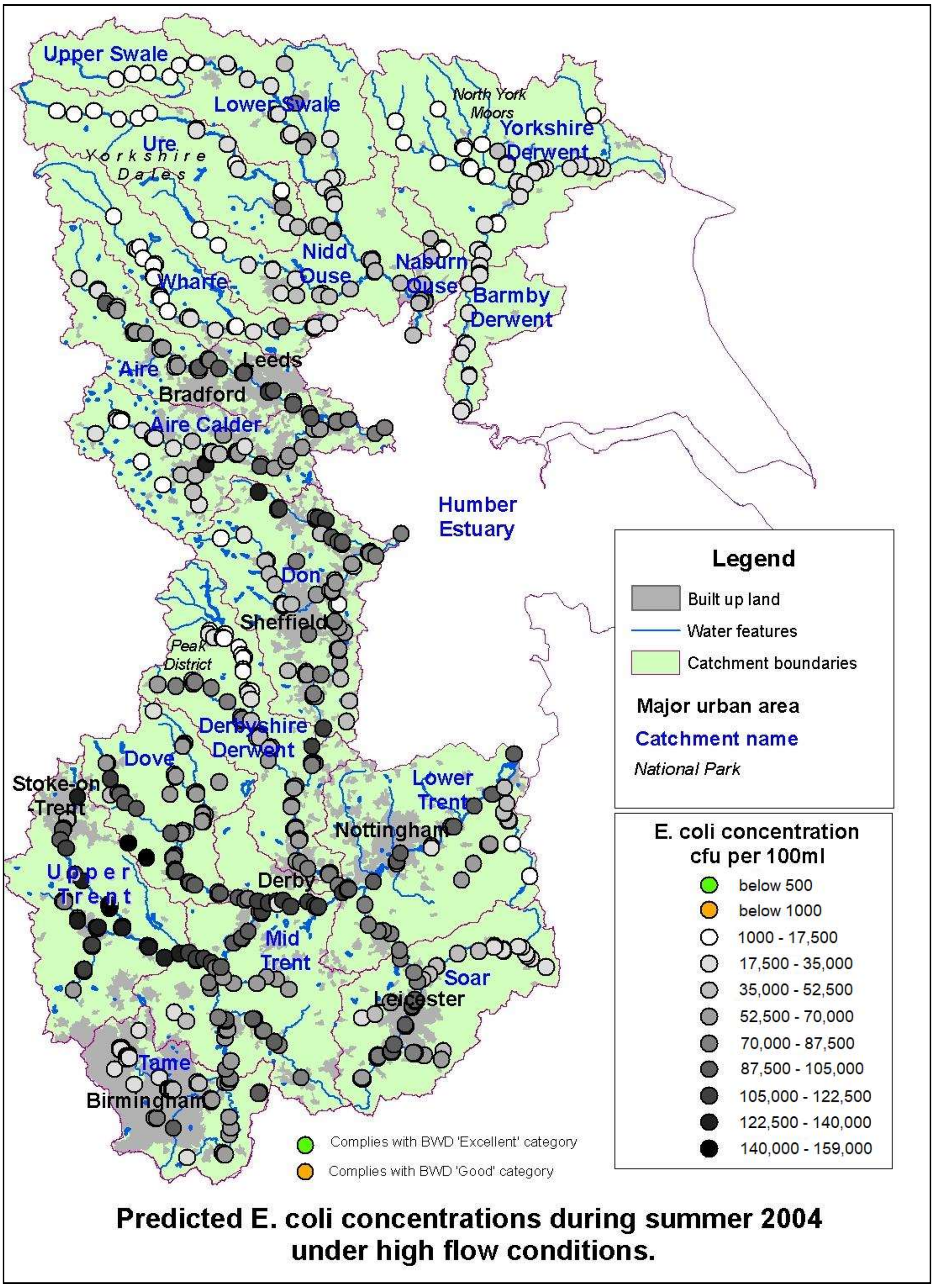
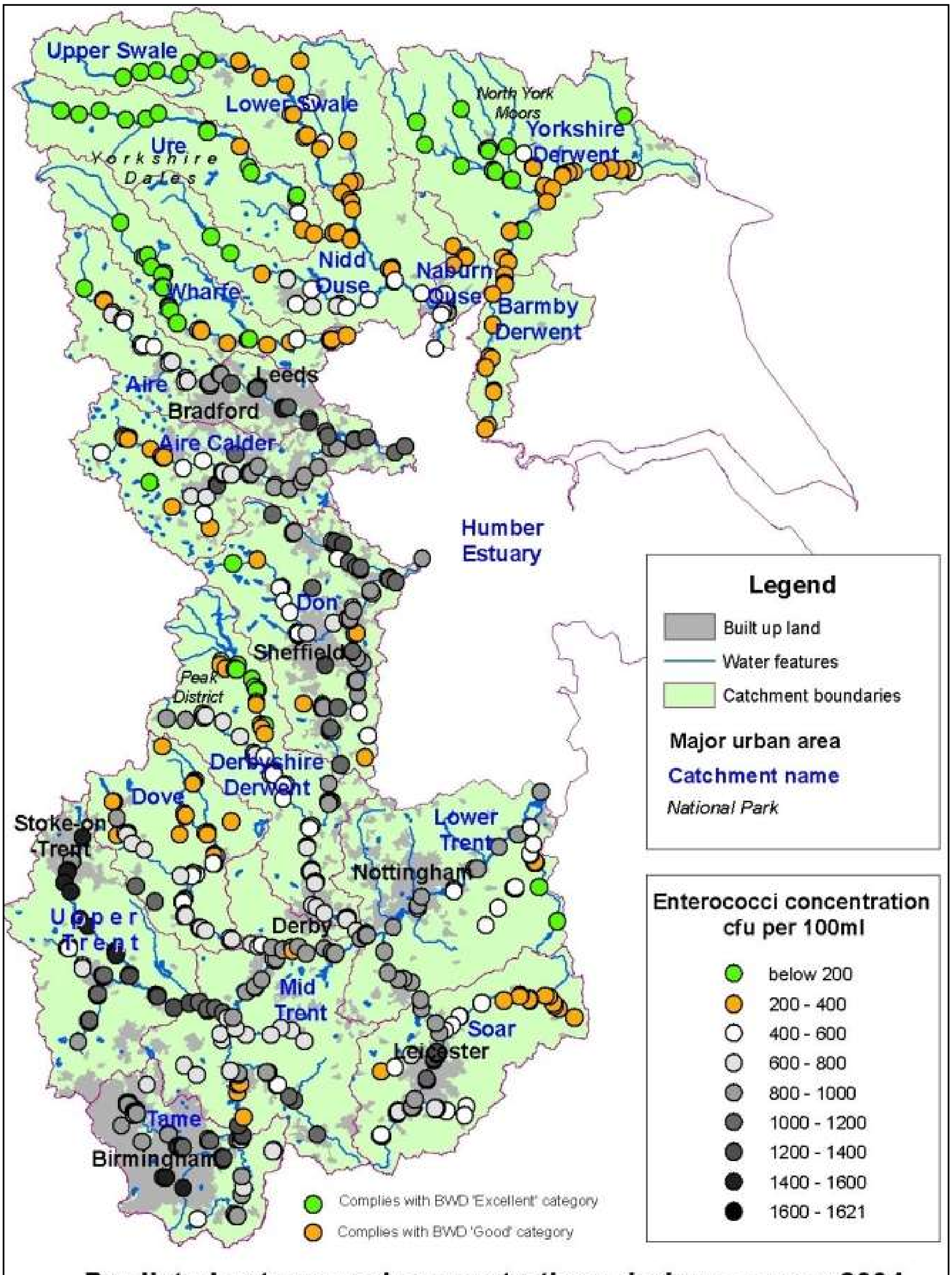
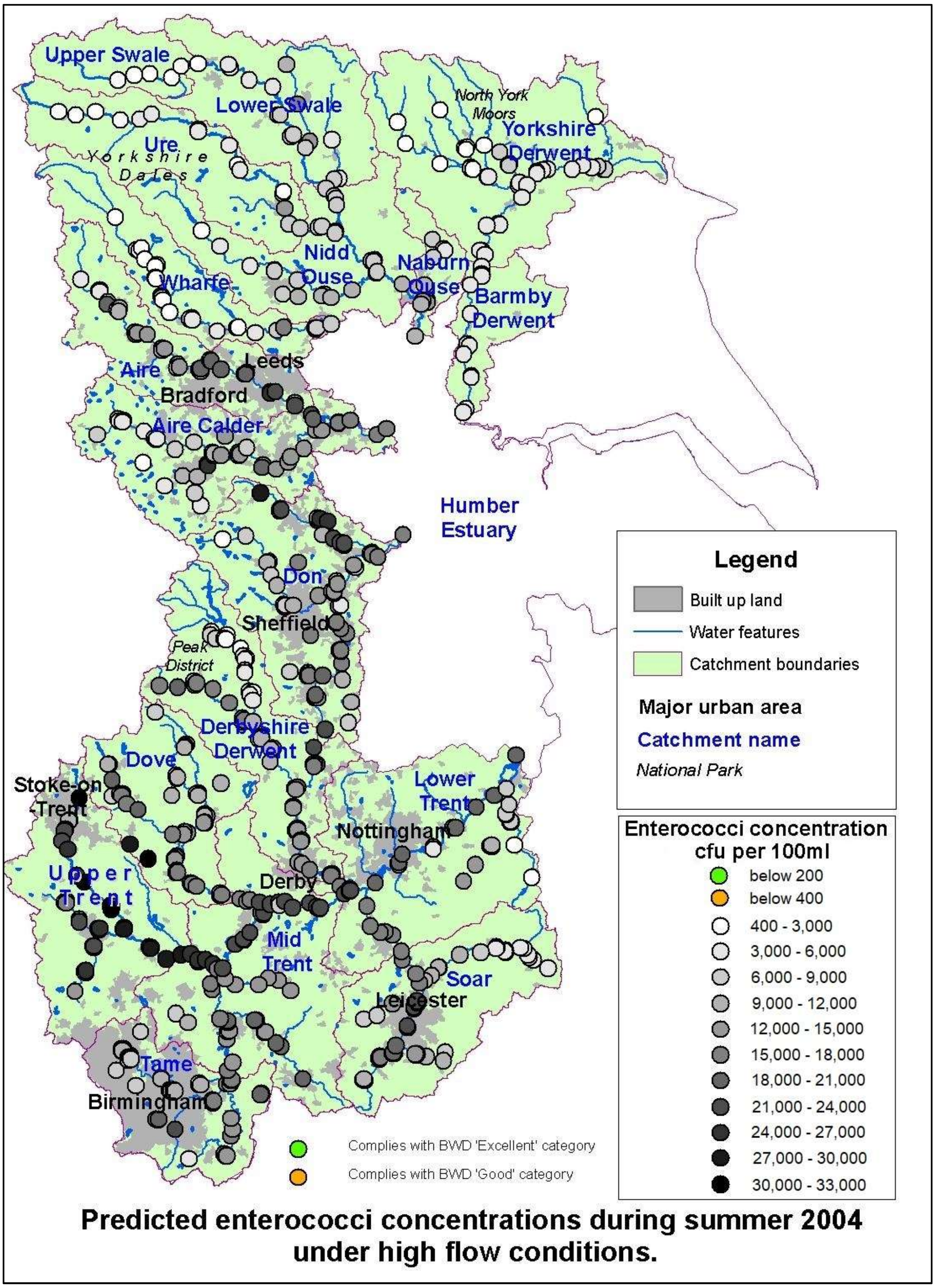


Figure 10: predicted EN concentrations $(\mathrm{CFU}\ 100\ \mathrm{ml}^{-1})$ **in the Humber RBD under base-flow conditions, summer 2004**



Predicted enterococci concentrations during summer 2004 under base flow conditions.

Figure 11: predicted EN concentrations $(\mathrm{CFU}\ 100\ \mathrm{ml}^{-1})$ in the Humber RBD under high-flow conditions, summer 2004



A cursory inspection of the maps (Figures 8 – 11) confirms that the transfer methodology appears to work well at the RBD scale of application. The two main trends shown on the maps are in line with previous research: FC concentrations are roughly an order of magnitude higher than EN concentrations and high-flow concentrations of both organisms are roughly an order of magnitude higher than base-flow concentrations. At high-flow no sites comply with the rBWD ‘Good’ water criteria.

There are large variations in the spatial distributions of FIO concentrations. As the patterns are broadly similar on all of the four transfer maps, regardless of FIO type or flow conditions, the following discussion applies to all four maps, unless indicated otherwise.

The two lowest concentrations of FC and EN, at both base- and high-flow, are found at the Knipton Reservoir tributary in the south-east of the Lower Trent subcatchment, and the Derwent Reservoir tributary in the north of the Derbyshire Derwent subcatchment. The rivers Ure, Wharfe and Upper Swale, rising from the Yorkshire Dales in the north-west of the Humber RBD and the tributaries to the Yorkshire Derwent (Seph, Dove, Severn, Hodge Beck), rising from the North York Moors have the lowest FIO concentrations. At base-flow, long stretches of these rivers comply with the rBWD ‘Excellent’ and ‘Good’ categories on Table 14. All of these rivers and tributaries have the same two characteristics in common: low human and low dairy population density.

In the absence of any major settlements, and with very low density dairy farming, the River Derwent has the lowest concentration of FIOs as it flows into the Humber Estuary via the Barmby Derwent subcatchment. It has ‘Good’ water status, in terms of base-flow EN contamination, along its entire length.

The highest FIO concentrations emitted into the Estuary subcatchment are from the rivers Don, Aire and Trent. Base-flow rBWD compliant sites are rare along these rivers, confined only to headwaters. Aggregated concentrations in the River Aire are raised by the presence of high dairy populations in two HRUs in the west of the subcatchment, before receiving large human FIO inputs as it passes through first Bradford then Leeds. Similarly, the River Don is loaded with human

FIO inputs from Sheffield and the neighbouring Rotherham before its confluence with the River Dearne.

The 13 sample points with the highest density dairy populations are found in the Dove and Upper Trent subcatchments. These correspond closely with the 10 highest base-flow FC concentrations and the 14 highest high-flow EN concentrations.

The most polluted river stretch in the Humber estuary is the River Trent between Stoke-on-Trent and its confluence with the River Sow. This stretch, in addition to receiving FIO inputs from some of the highest dairy concentrations, also receives very high human inputs from its headwaters near Stoke-on-Trent. Although the south and south-east of the Dove subcatchment has high density dairy farming, the FIO concentrations in the River Dove are lower than the River Trent because it does not have the high density human inputs found in the Stoke-on-Trent stretch.

Twenty-seven of the thirty most densely populated human areas are in the Tame subcatchment, more specifically within Birmingham. As these areas typically have near zero dairy populations they receive virtually no FC or EN from agricultural sources during high-flow conditions (Dairy being the most significant source of FC and EN at high-flow, according to the models on Table 16). Therefore the concentrations of agricultural EN and FC are noticeably lower in Birmingham than in the intensive dairy farming areas of the Upper Trent, particularly the tributary rivers Sow and Penk in the south-west of the subcatchment.

In contrast, the base-flow EN map, Figure 10, reveals that two of the tributaries to the Tame, the Rea and the Cole, in the south-east of Birmingham, have very high EN concentrations under base-flow. This is in line with predictions from the base-flow EN model, on Table 16, which has human density as its most significant source of FIOs. This phenomenon is obvious in Birmingham, but with closer inspection it can be seen that all major urban areas, particularly Stoke-on-Trent and Leicester, are responsible for elevated levels of EN at base-flow.

The Lea Marston purification lakes, on the border of the Tame and Mid-Trent subcatchments, are effective at reducing base-flow FIO concentrations (Martin and Brewin, 1994; Environment Agency, 2004). Base-flow EN concentrations, Figure 10, are reduced considerably, to the level of rBWD 'Good' status, as they pass through Lea Marston. Although the Lea Marston lakes are discussed in detail in the final chapter, suffice to say here that engineered purification schemes make tangible improvements to downstream water quality at base-flow. The FIO reductions achieved at Lea Marston are largely responsible for diluting the high levels of FIOs in the River Trent at its confluence with the Tame in the east of the Mid-Trent catchment. This is noticeable on all four maps.

2.5.2 Section 2: predicting changes in FIO concentrations in response to land use change scenarios

The results of three simple scenarios are now considered. These are: (scenario 1) a 20% decrease in dairy livestocking density; (scenario 2) a 1.4% increase in human population; and to demonstrate the ability of the models to model simultaneous changes, a mixed scenario which applies three changes to FIO sources: a 20% decrease in dairy livestock, a 1.4% increase in human population and a 5% improvement in WwTW efficiency (scenario 3).

For simplicity each of these three scenarios assume a blanket application of the amended FIO inputs across the Humber RBD. It is acknowledged that any changes to dairy herd size are unlikely to be proportional across space, that human population increases are unlikely to be uniformly distributed but will probably be concentrated into the main urban areas (RERC, 2007) and that any improvements to WwTW infrastructure are likely to be targeted into areas that would produce the most cost effective improvement.

Table 22: predicted changes to maximum and mean FIO concentrations within the Humber RBD, in response to land use change scenarios

Pathogen type	Flow conditions	Percent change in pathogen concentration	
		Maximum	Mean
Scenario 1: a 20% reduction in dairy livestock within the Humber RBD			
FC	High-flow	-13.72	-11.76
EN	High-flow	-11.96	-10.23
FC	Low-flow	-11.37	-9.69
EN	Low-flow	-6.79	-5.77
Scenario 2: a 1.4% increase in human population			
FC	High-flow	0.43	0.42
EN	High-flow	0.48	0.47
FC	Low-flow	0.60	0.59
EN	Low-flow	0.58	0.57
Scenario 3: a 1.4% increase in human population, a 20% decrease in dairy livestock and a 5% improvement in WwTW efficiency			
FC	High-flow	-14.67	-12.74
EN	High-flow	-13.06	-11.37
FC	Low-flow	-12.76	-11.11
EN	Low-flow	-8.2	-7.20

Although the spatial patterns of water quality in response to each scenario is very similar, regardless of flow rate or pathogen type, the magnitude of the changes in water quality is different for each flow rate or pathogen type within each scenario. Table 22 shows these differences.

The greatest FIO reductions in scenarios 1 and 3 occur under high-flow conditions. FC reductions are typically 1.5% higher than EN reductions at high-flow and 4% higher at base-flow. Note that for all scenarios and all models the distribution of values is highly skewed towards the maximum, particularly in scenario 2. This is in part due to the fact that there are very few zero values, or values close to the minimum. Those that do exist have been generated by sample points containing high proportions of reservoir catchment.

Rather than report four output maps for each scenario (corresponding to the results of each of the four models), the results of the FC high-flow model are reported for both the change to livestocking density and the mixed scenario (Figures 12 and 14), as it is the model which provides the highest level of explanation (Table 16). The results of the base-flow EN model, applied to model an increase in human population (Figure 13), are reported as humans are the most significant source of FIOs at base-flow and this model may better describe the relationship between increased human FIO outputs and riverine EN concentrations. The changes in water quality are displayed in quantiles.

Scenario 1 simulates a livestock destocking option, proposed by Cuttle et al. (2007). FIO concentrations are predicted to reduce significantly because of the measure. A 20% decrease in dairy cattle results in mean high-flow reductions across the Humber RBD of 11.76% and 10.23% for FC and EN respectively, shown on Table 22.

Figure 12 shows that areas of intensive dairy farming, such as those found along the rivers Penk and Sow in the Upper Trent subcatchment, or the rivers Hamps and Chumet in the Dove subcatchment, respond particularly well to compulsory destocking, with FC reductions along these rivers of c.13.5%.

Urban areas, particularly urban headwaters, i.e. Birmingham and Stoke, and areas with sparse dairy populations, such as the Upper Swale and Wharfe subcatchments, respond less well, with below average FIO reductions.

By applying the FIO models at RBD scale pollution hotspots can be identified: two HRUs in the headwaters of the Aire which are densely populated with dairy livestock, respond well to the destocking, with FIO reductions in the top quantile. These high reductions are progressively diminished, particularly as the Aire passes through Bradford and Leeds, until, at the confluence with the Calder, the reductions are close to the mean.

Figure 12: predicted reductions in FC concentrations in response to scenario 1, a 20% reduction in dairy livestock within the Humber RBD, summer 2004, high-flow conditions.

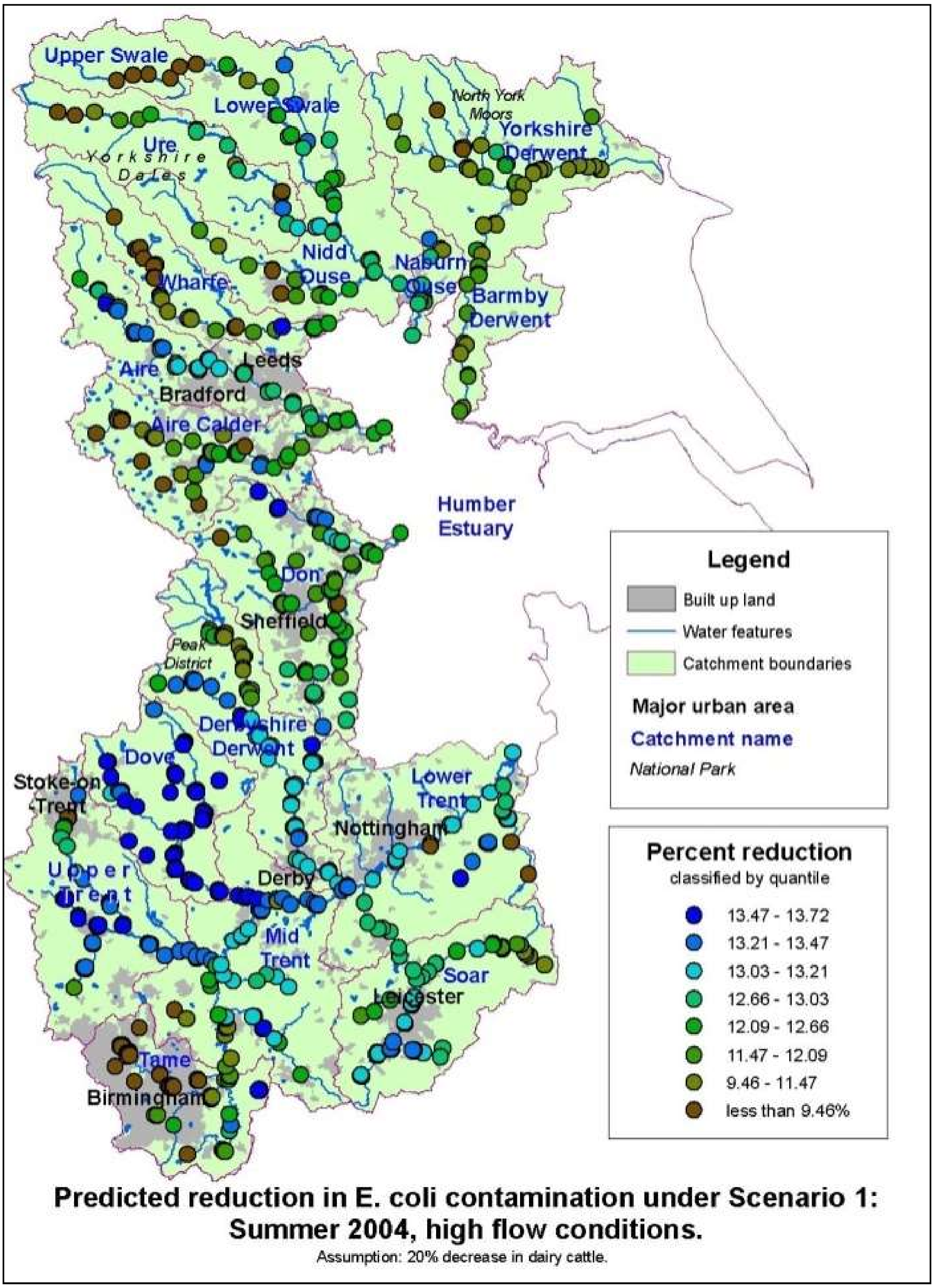


Figure 13: predicted increase in EN concentrations in response to scenario 2, a 1.4% increase in human population within the Humber RBD, summer 2004, base-flow.

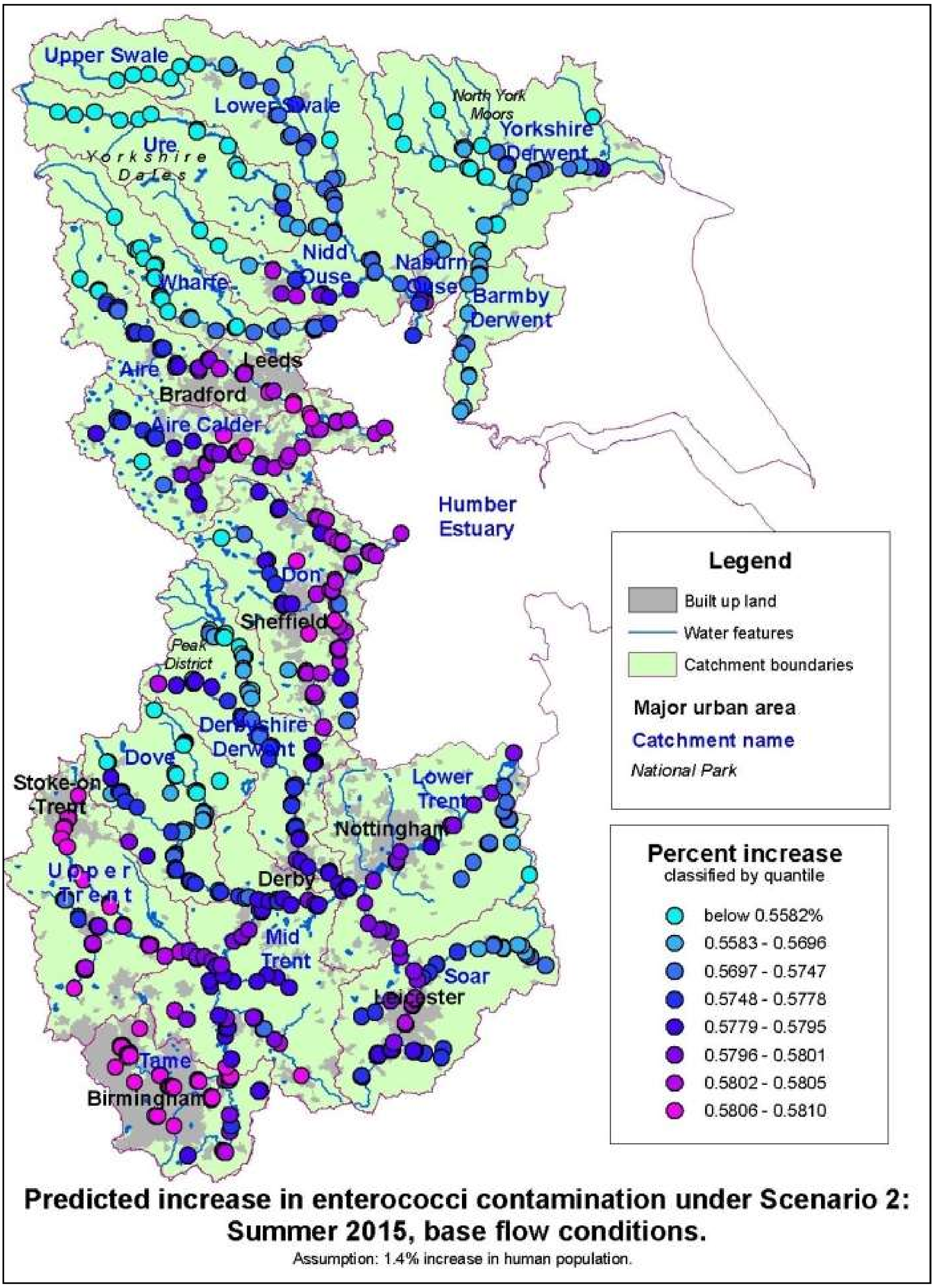
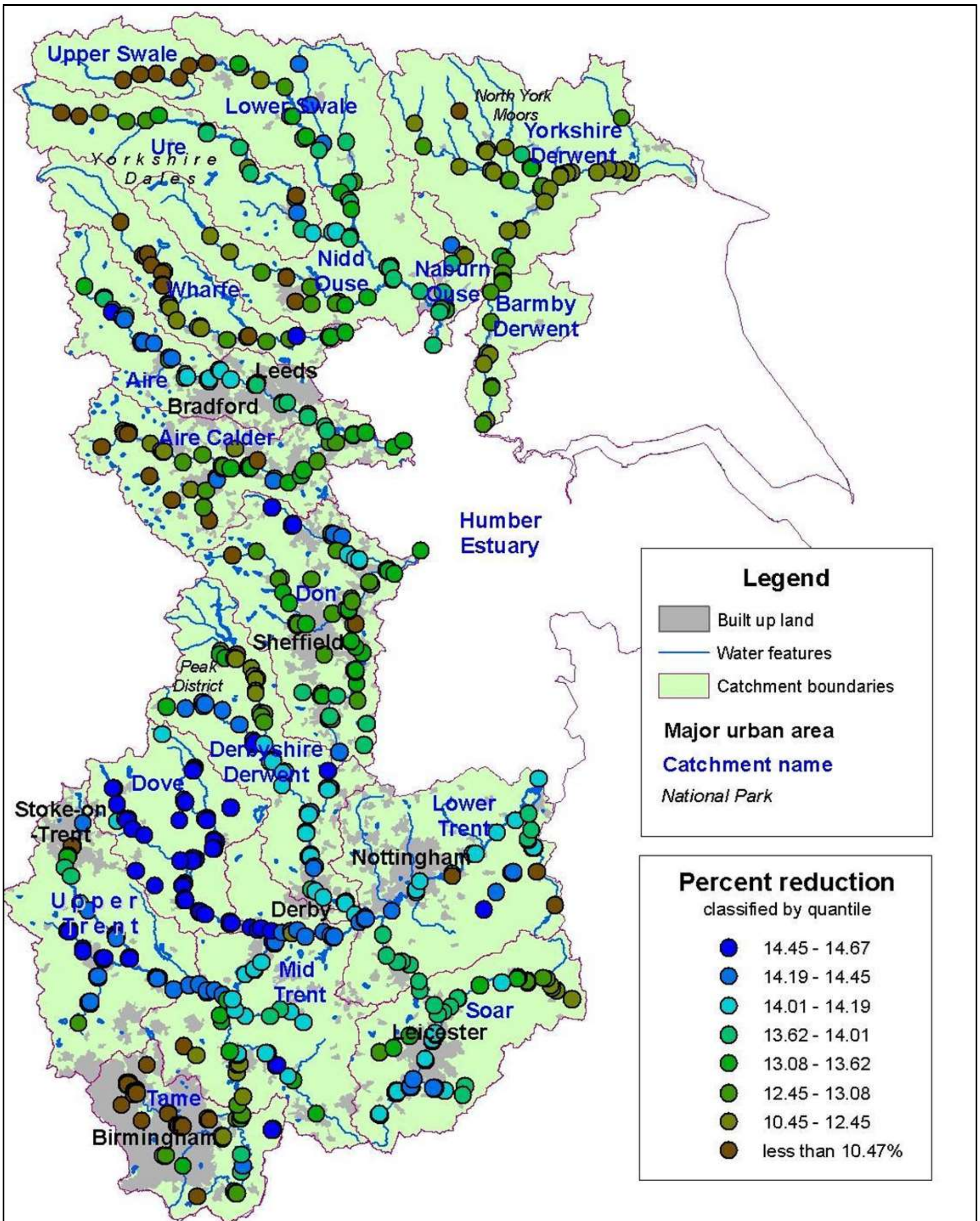


Figure 14: predicted decrease in FC concentrations in response to scenario 3, a 1.4% increase in human population, a 20% decrease in dairy livestock and a 5% improvement in WwTW efficiency within the Humber RBD, summer 2004, base-flow.



Predicted reduction in E. coli contamination under Scenario 3: Summer 2015, high flow conditions.

Assumptions: 1.4% human population increase, an improvement in wastewater treatment equivalent to a 5% reduction of FIOs from human sources, and a 20% decrease in dairy cattle.

The human population of the Humber RBD was predicted to grow by 1.4% between 2004 and 2015 (Appleton, 2008) and the impact on predicted riverine EN concentrations is shown in Figure 13. This scenario assumes no changes in WwTW efficiency, i.e. WwTWs process the increased quantities of human waste proportional to their current rate of efficiency.

Table 22 shows that the small human population increase results in a small increase in riverine FIO concentrations across the Humber RBD, typically around 0.5% for both FC and EN. There is a strong correlation between riverine FIO concentrations and human population density. Figure 13 shows that the highest EN increases occur in the urban conurbations and the lowest increases are in the remote upland areas. Increases in EN concentrations occur within a very narrow range: 90% of base-flow EN concentrations increase by 0.57 - 0.60%.

In addition to a 20% reduction of FIOs from dairy sources, the third scenario effectively reduces human inputs by 3.6%, as the increased FIO concentrations arising from human population increase (+1.4%) are cancelled out by a hypothetical improvement to WwTW efficiency (-5.0%).

The distribution of predicted FIO reductions in scenario 3, shown on Figure 14, is very similar to scenario 1, shown on Figure 12. Table 22 shows that the improved WwTW enables the scenario 3 FIO reductions to be approximately 1% higher than the reductions bought about by scenario 1 at high-flow, and 1.5% higher than at base-flow.

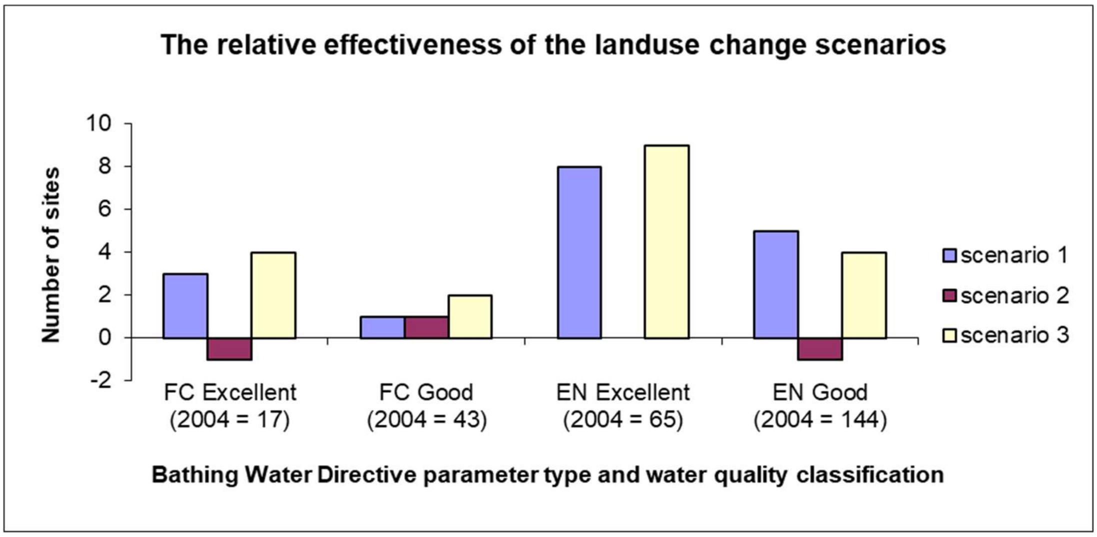
As the results of each of these three scenarios show, despite a blanket application of the scenarios across Humber, not all areas respond equally to the land use changes. This indicates that spatially differentiated policies could be implemented to maximise the return on different land use management strategies in different areas.

2.5.3 The effect of land use change on rBWD and WHO compliance

Figure 15 shows the number of change in the number of sample points that comply with each of the rBWD water quality classifications at base-flow following implementation of scenarios 1-3. Changes are compared against estimates of the number of FC and EN compliant sites in 2004 (shown in brackets in Figure 15).

Scenarios 1 and 3 produce slight improvements in the number of rBWD compliant sites. For example, the number of 'Excellent' EN sites is predicted to rise by 9 sites, from 65 sites in 2004 to 74 sites in 2015, if scenario 3 were to be implemented. The number of 'Excellent' FC sites rises from 17 sites in 2004 to 21 sites under scenario 3. Scenario 2 results in one less FC 'Excellent' site and one less EN 'Good' sites.

Figure 15: projected rBWD compliant sites in response to the proposed land use change scenarios



The implementation of the scenarios do little to improve the vast majority of sample points to rBWD standards of water quality. After implementing scenario 3 the mean EN concentration across the Humber RBD, at base-flow, only improves from 627 CFU 100 ml^{-1} to 581 CFU 100 ml^{-1} , still above the threshold of 400 CFU 100 ml^{-1} necessary for a site to be considered 'Good' under the rBWD.

So far the results have shown that water quality in the Humber RBD is generally poor and continues to be poor despite the three simple land use change scenarios. The next results assess the water quality of the Humber RBD in terms of its risk to human health.

Figure 16: results of a QMRA of ill health due to EN contamination in the Humber, base-flow, summer 2004

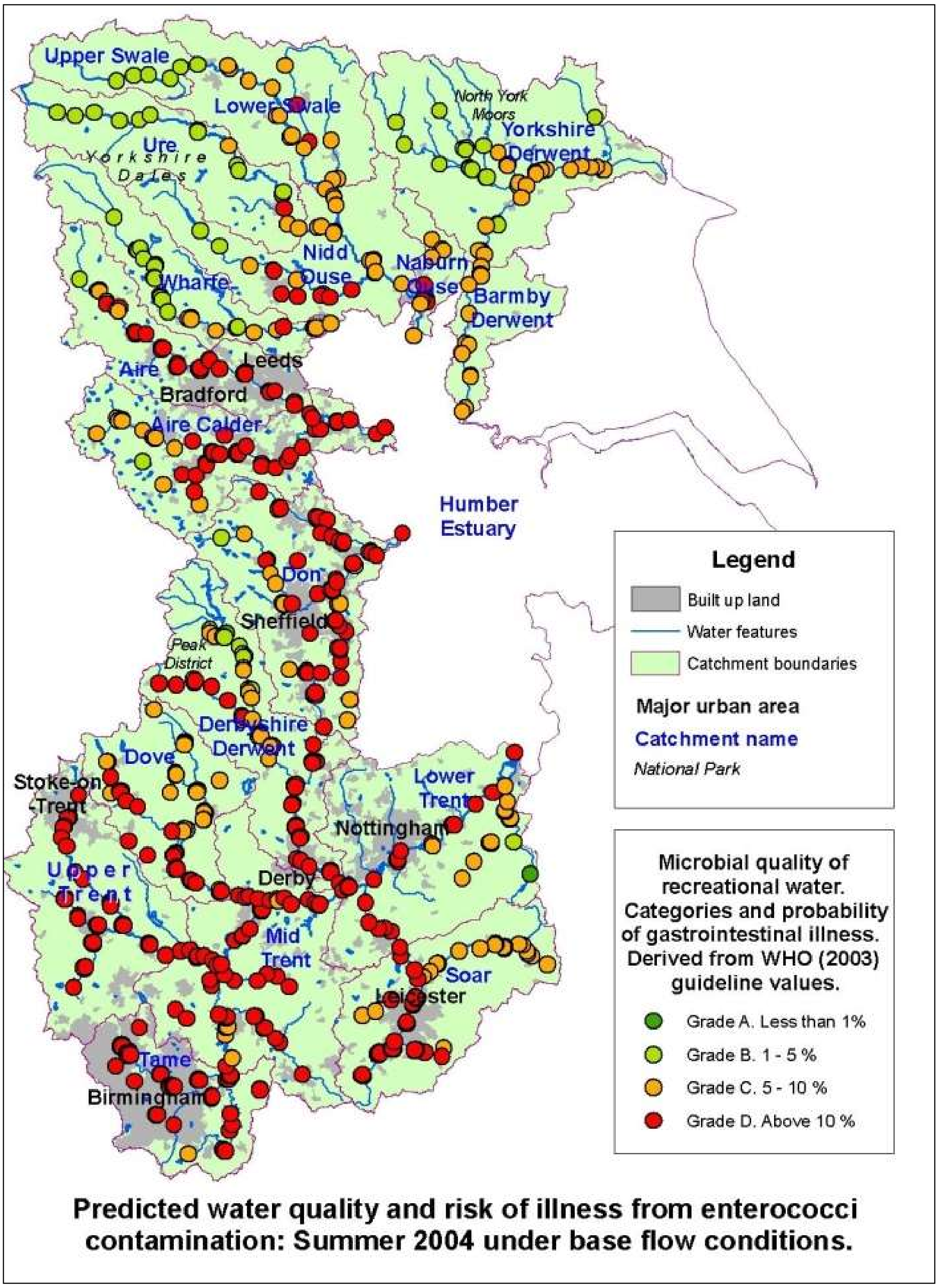
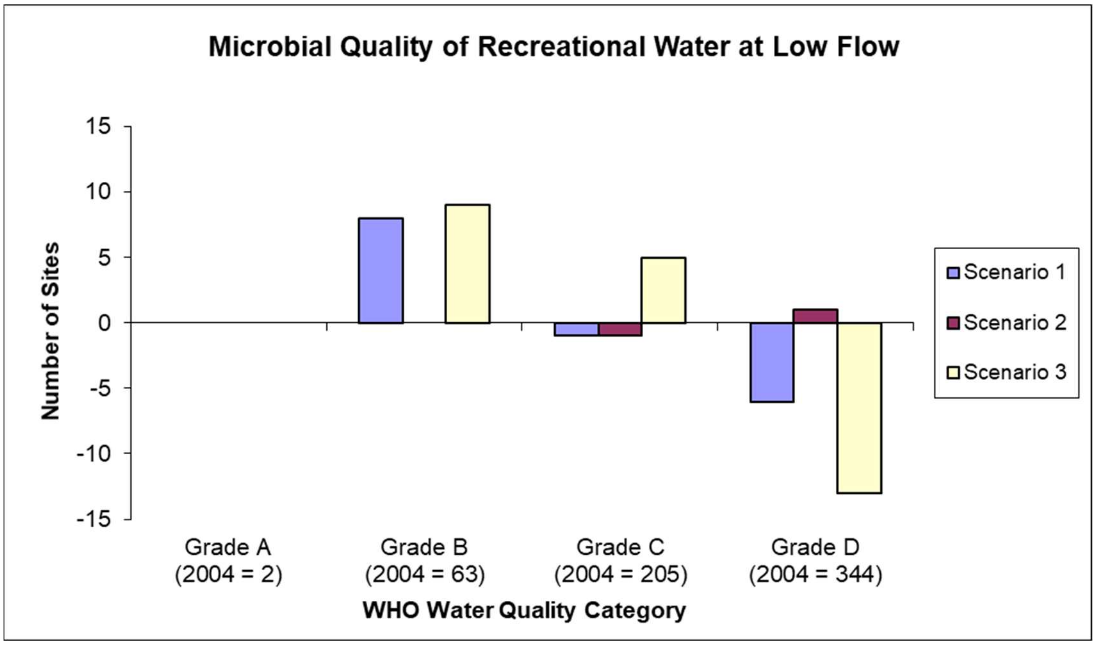


Figure 16 shows the predicted water quality and associated probability of gastrointestinal illness due to enterococci contamination during 2004, under base-flow conditions. Only the tributaries flowing from the Derwent and Knipton reservoirs are classed as Grade A. 63 sites have Grade B water quality. These sites are predominantly in the north-west of the catchment, along the upper reaches of the rivers Ure, Wharfe and Upper Swale. Long stretches of the Rye and Hodge Beck have ‘Good’ water quality before their confluence with the Yorkshire Derwent. 205 sampling points have Grade C water quality. The entire length of the River Wreake, in the Soar subcatchment, is in this category, as are long stretches of the Lower Swale, the Nidd, the Ouse and the Derwent in the Barmby Derwent subcatchment. Over half of all sampling points, 344, are in Grade D. These rivers have poor water quality and present the most serious risk of gastroenteritis. Almost the entire length of the River Trent has Poor water quality, as do virtually all urban rivers and rivers passing through areas of intensive dairy farming.

Figure 17 shows that the small improvements to water quality, bought about by implementing scenarios 1 or 3, result in small improvements to the risk of ill-health. Scenario 3 results in 8 more sites at Grade B, 5 more sites at Grade C and 13 fewer sites categorised as Grade D. Scenario 2 results in a slight increase in the risk of ill health, with one site reclassified from Grade C to Grade D.

Figure 17: comparing the effects of the land use change scenarios on microbial water quality at base-flow



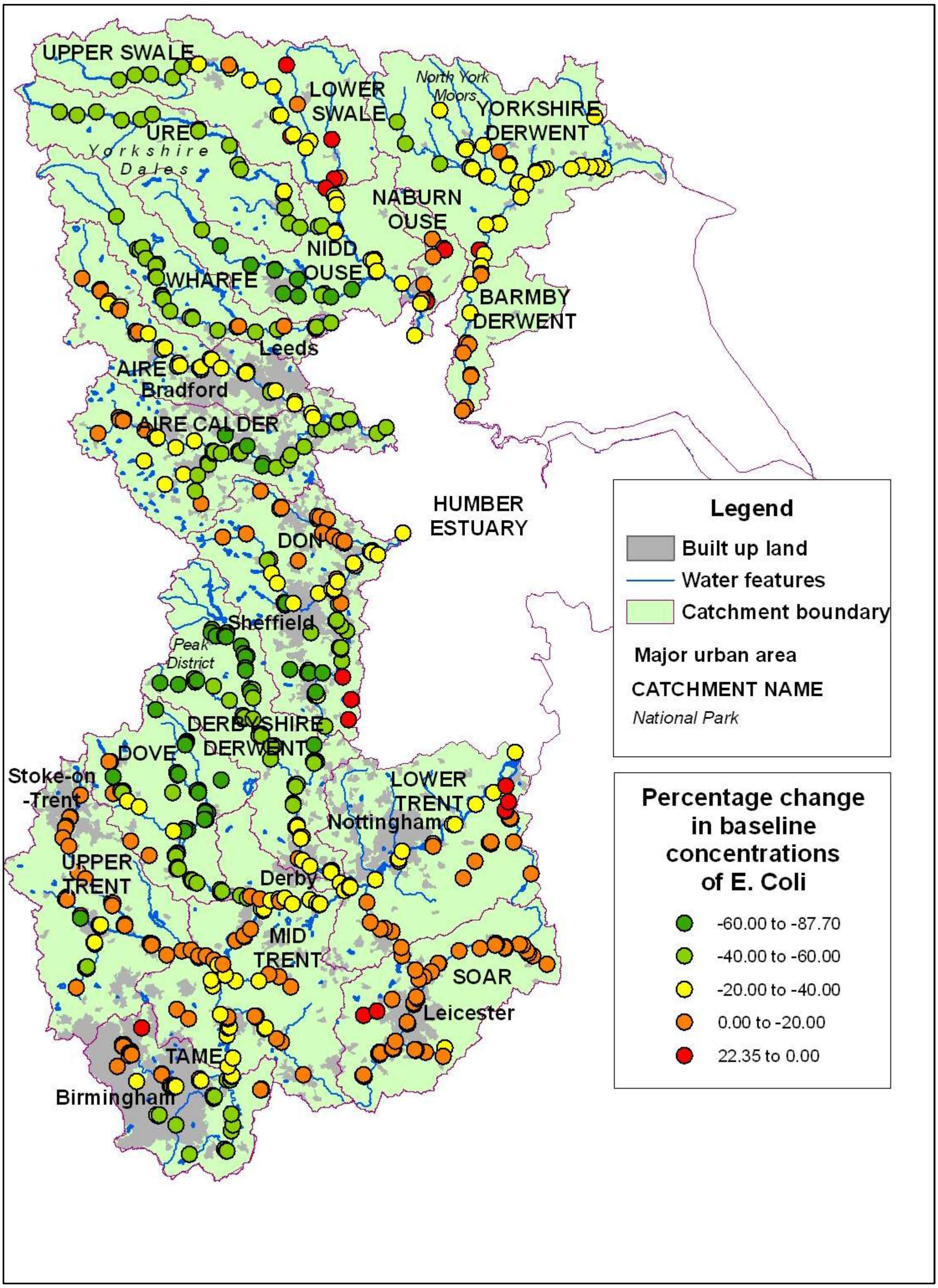
The microbial quality of water in the Humber RBD during high-flow is poor. Estimates for 2004 suggest that 613 of the 614 sample points were classified as Grade D, with one site, near the Knipton and Derwent reservoir, classified as Grade C. Scenarios 1 and 2 produce no change in classification. Implementing Scenario 3 results in one other sample point near the Knipton and Derwent reservoir being reclassified as Grade C. All of the remaining 612 sample points remain in Grade D, regardless of the water improvements brought about by land use changes.

2.5.4 Demand side constraint: adopting a nutrition driven food policy

The final assessment at RBD scale reported within this section is a demand side constraint on FIO inputs caused by the UK government adopting a nutrition driven food policy. Figure 18 shows that the largest reductions to riverine FC are predicted to occur in the upland areas (e.g. the Yorkshire Dales and the Pennines) to the west of the RBD, with dairy numbers (and consequent FIO concentrations) typically maintained or rising slightly in some HRUs (e.g. the Lower Trent, the River Soar and the River Don) in the east - possibly as production is transferred to lowland areas.

The LUAM also projects a drive to fewer producers rearing higher yielding dairy animals in larger herds. This is reflected in Figure 18, as areas of intensive milk production in the Upper Trent, Aire and Dove subcatchments experience relatively low reductions to herd size and riverine FC concentrations.

Figure 18: reduction in FC concentrations following adoption of the nutrition driven food policy



2.5.5 Section 3: the relative effectiveness of different remediation strategies within the Aire subcatchment

The remaining results focus almost exclusively on the results of the high-flow FC model, unless indicated otherwise. The section opens with an overview of the impacts of a range of remediation strategies at RBD scale, within the Humber RBD, before assessing their relative effectiveness at subcatchment scale within the Aire subcatchment (see Table 23, p.138).

2.5.6 Fiscal constraint: taxing fertilizer by £50/tonne

Although the econometric model (Equation 2) underpinning this scenario predicts that an increase in fertilizer price decreases the optimal shares of nutrient-intensive activities (which convert to low biomass yield land uses such as rough grazing), the increased price produces only a small predicted decrease in dairy livestock numbers, with average reductions in riverine FC concentrations (predicted in the high-flow FC model) of 1.9% distributed relatively uniformly across the Humber RBD.

2.5.7 Area intervention: designating the Humber RBD as an Environmentally Sensitive Area

The econometric modelling predicts important land use transformations, which lead to complex patterns of FIO reductions. ESA designation is predicted to significantly reduce the number of dairy cows, which are substituted by less intensive units, such as beef cows and, particularly, sheep. The high-flow FC model suggests that this shift leads to slight increases in sheep-derived FIOs in the upland HRUs of the Upper Swale and Ure subcatchments, but these increases can be offset against reduced dairy cow derived FIOs as economically marginal dairying operations are ceased. This area intervention produces average FC reductions in the Lower Swale, Nidd Ouse and Naburn Ouse of 25-30% and mean reductions in FC concentrations of 9.5% across Humber. Intensive dairying in the Upper Trent, Dove and Aire subcatchments see only modest improvements from implementation of this policy.

2.5.8 Cost intervention: raising the price of the EU milk lease quota by £20

The high intensity dairying regions of the Aire, Dove and Upper Trent subcatchments do not see high reductions in FIOs for a measure designed specifically to put pressure on milk production. The larger producers in these areas enjoy greater efficiencies of scale and are more likely to weather the increased transaction costs associated with the milk quota price rise. Economically marginal producers in the Lower Swale, Nidd Ouse, Naburn Ouse and Soar subcatchments appear most affected by this policy, which leads to relatively high reductions in dairy cow derived FC concentrations in these regions. Increasing the price of the milk quota achieves mean reductions in FC concentrations of 6.4% across Humber.

2.5.9 Production constraint: reducing dairy cattle stocking rates by 20%

A 20% decrease in dairy cattle numbers results in mean high-flow reductions of 11.7% for FC across Humber. Areas of intensive dairy farming, such as those found in the Aire, Upper Trent and Dove subcatchments, respond particularly well to compulsory destocking, with FC reductions along some sections of these rivers of c.13.5%.

2.5.10 Input constraint: reducing fertilizer application by 20%

The patterns of FC reductions predicted by the FIO model within this scenario are very similar to those achieved by dairy cattle destocking, described above, but the policy measure is less effective at reducing FIO concentrations: the measure is predicted to result in mean high-flow reductions of 7.9% for FC across the Humber RBD. However, this scenario does not fully account for potential real-world behaviour: Dairy farmers may increase manure applications to compensate for reduced fertilizer application, thus maintaining grassland productivity and FIO emissions, but this possibility is not considered here: the impact of the input constraint is applied uniformly across space.

The effects of the above land use intervention scenarios provide modest improvements in FC concentrations within the Aire subcatchment, described on Table 23, ranging from 1.66% to 11.58%, with a mean improvement of 7.6% immediately downstream of the high intensity dairy region.

Table 23: comparison of the effectiveness of the remediation strategies in the Aire subcatchment at high-flow

Remediation measure	Predicted FC concentration (1000 CFU 100ml^{-1}) exiting dairy region	Percent reduction from baseline concentrations exiting dairy region	Predicted FC concentration (1000 CFU 100ml^{-1}) at subcatchment outflow	Percent reduction from baseline concentrations at subcatchment outflow
Baseline prediction, high-flow, 2004	93	-	118	-
Taxing fertilizer by £50/tonne	91	1.66	117	1.4
ESA designation in Aire	84	8.99	108	8.6
Increase milk quota cost by £20	87	5.74	100	5.02
Reduce dairy stocking by 20%	82	11.58	105	11.23
Reduce fertilizer application by 20%	84	9.96	109	7.62
Installation of stream bank fencing	38	58.59	77	34.69

Taxing fertilizer produces small reductions in livestock numbers across the Aire subcatchment, resulting in a 1.4% reduction of FC concentrations at the subcatchment exit. Designating the Aire subcatchment as an ESA is predicted to reduces dairy livestock populations in the more economically marginal areas in the north-west of the subcatchment. This results in mean predicted reductions of FC concentrations of 8.6% at the subcatchment exit. This is below the mean reduction for the Humber as a whole (9.5%). Similarly, raising the price of the EU milk lease quota by £20, results in FC reductions of 5.02% at the subcatchment exit. This is also less than the mean reductions in FC concentrations of 6.4% within the Humber RBD.

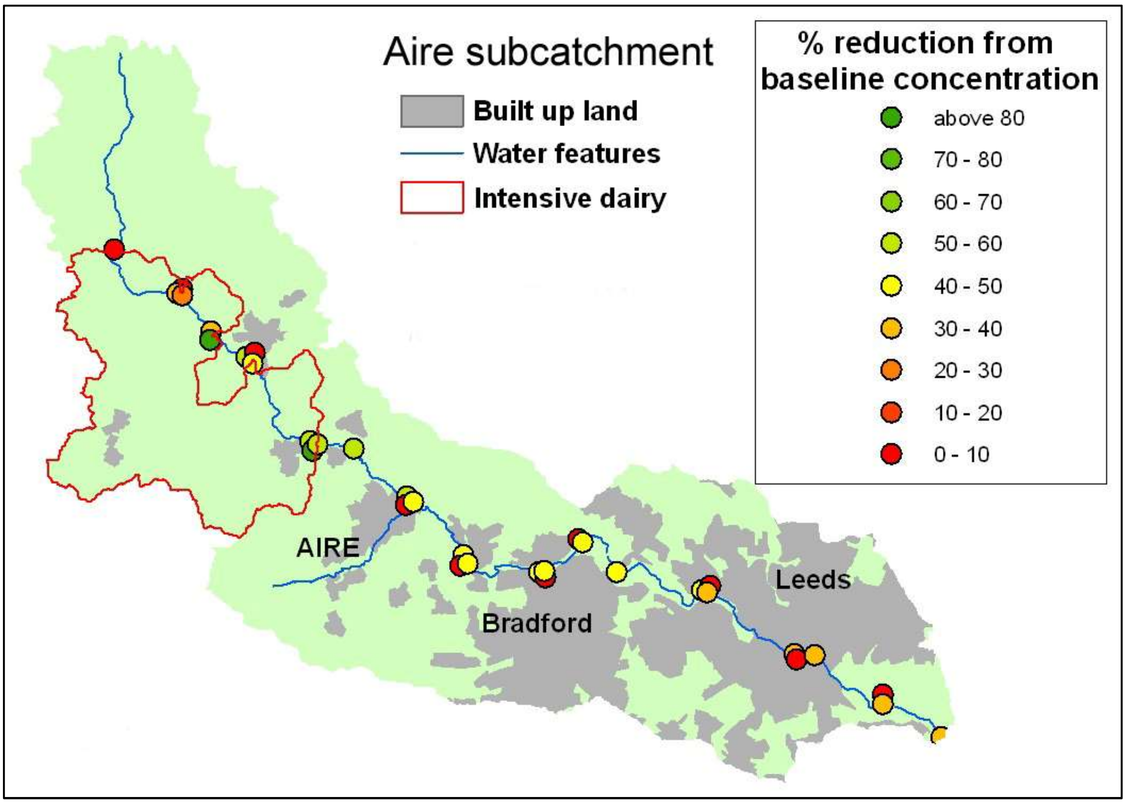
Blanket applications of policy interventions may not be the best approach towards reducing emissions of FC in the high intensity dairy regions within the west of the Aire subcatchment. The large-scale operations in that area enjoy greater efficiencies of scale, are better able to weather increased transaction costs, and, consequently, maintain the sizes of their dairy herds. For this reason, compulsory dairy livestock destocking applied uniformly across all areas of the Aire subcatchment, including the areas of intensive dairy operations, results in a predicted reduction of 11.23% in FC at the subcatchment exit. This intervention policy is more effective, but may be accompanied by a considerable loss of revenue to those intensive producers.

Rather than interventions designed to reduce FIO emissions via reductions in dairy herd size and dairy population densities, interventions designed to manage faecal waste (i.e. prevent waste from entering watercourses) may be more appropriate. The next scenario examines the impact of installing stream bank fencing within the areas of high intensity dairy operations in the Aire subcatchment.

2.5.11 Micro-level land use management: stream bank fencing

This micro-level scenario applies the high flow FC model (Table 16) to the Aire subcatchment only. With stream bank fencing, erected only within the intensive dairy regions (outlined in red in Figure 19), there are marked reductions in FC concentrations downstream (see Figure 19 and Table 23). There are predicted reductions in FC concentrations of 58.59% immediately below the improved area and 34.69% at the subcatchment outflow in the east. This targeted intervention is predicted to be far more effective at reducing riverine FIO concentrations than the less focussed blanket policies described above

Figure 19: reductions in high-flow FC concentrations resulting from stream bank fencing erected in intensive milk producing HRUs



2.6 Discussion

The majority of the results in Table 23 show trivial improvements in riverine water quality for measures which could potentially result in large reductions in farm income or destabilise fragile local communities in economically marginal upland areas. Some of the more aggressive macro policies outlined here (e.g. raising the price of milk quota or designating Humber as an ESA) may result in the concentration of milk production into large enterprises, exacerbating localised microbial pollution concentrations.

It is acknowledged that not all microbial burden to land is of equal mobility or persistence (Jones, 1999) and there is still uncertainty surrounding the effectiveness of riparian buffer strips (Kay et al., 2007a). In a recent review (Kay et al., 2012), vegetated buffer strips were found to have a median FIO attenuation rate of 90%, which provides some confidence in the 95% rate of attenuation attributed to streamside fencing within this research. Likewise, the results of empirical research conducted at Brighouse Bay found that the installation of FIO remediation measures, principally stream side fencing, helped to reduce riverine FC fluxes by 66.3% at high-flow (Kay et al., 2007b). Although the Brighouse study was vulnerable to the potential, but unquantified, effects of seasonality, the FC reduction achieved at Brighouse is comparable to the 58.59% reduction in FC concentrations, due to streamside fencing, predicted by this research (Table 23). Stream side fencing may be a low-cost high-yield strategy. This micro-level land management strategy greatly exceeds the reductions achieved by the other policy instruments at the subcatchment outflow (see Table 23, p.138).

In terms of improving the water quality of UK rivers the majority of the policy measures investigated here have relatively low impact on reducing FIO concentrations. The models predict that both faecal indicators greatly exceed EU guidelines for inland water quality, in Annex I of the Bathing Water Directive (2006a), at both flow rates. To reduce mean riverine FIO concentrations to mandatory levels (e.g. 400 CFU $100\ \mathrm{ml}^{-1}$) would require unattainable short-term improvements to WwTWs and politically impossible changes to the farming sector. Indeed, extreme measures would be required to make significant reductions to the risk of ill health at the theoretical Humber RBD sampling points. If all dairy farming ceased and the sewage infrastructure was improved to be

100% efficient, only then would base-flow water quality improve to Grade A. Even with these extreme measures, high-flow water quality could not improve beyond Grade C due to unexplained sources of FIOs contained within the intercept terms of the FIO models on Table 16.

Previous research estimated that 10% of WwTW plants would have to fit tertiary treatment systems by 2015 at enormous financial cost (Wither et al., 2005; Water Industry Network Wales, 2007). Even after investments on this scale, catchments have still been found to be non-compliant due to diffuse agricultural FIO sources (Crowther et al., 2001; Aitken, 2003). The 2001-02 foot and mouth outbreak produced statistically significant reductions in FIO concentrations (Stapleton et al., 2003), but farm incomes and rural communities were seriously jeopardised. Reduction in farm incomes will critically limit the adoption of widespread destocking (Cuttle et al., 2007).

This research highlights issues of spatial scale surrounding the delivery of land use policy measures. Although large scale integrated catchment management strategies have been successfully implemented abroad for several years (Heinz, 2003; US EPA, 2007) both the Rural Economy and Land Use Programme (RELU) and Defra are keen to stress local adaptation and innovation in land management policy implementation. Chadwick et al. (2008) highlight the need for farmers to target their efforts efficiently, per individual circumstances. There is also a requirement for the assessment of environmental trade-offs from different land management practices to enable farmers to develop effective micro-level mitigation strategies.

Clearly a balance between an optimal level of legislation, an efficient level of pollution and a spatial differentiation of land use policies needs to be devised. The results of the transfer exercise show that very few sites within the Humber RBD comply with EU standards or WHO recreational water criteria. The interpretation of the legislation is crucial to the implementation of pollution remediation strategies.

FIOs are typically extremely concentrated in rivers (as opposed to FIOs at coastal bathing sites, which are significantly diluted by large volumes of seawater), and as a consequence, there are very few inland sites designated as fit for bathing.

Only 11 of the 574 designated bathing water sites monitored by the EA are inland freshwater sites (Harley, 2008), none of which are in the Humber RBD (Defra, 2008b).

Annex I of the rBWD (EU, 2006a) allows deviation from compliance parameters during 'short-term pollution' events which typically last no longer than 72 hours. The vast majority of high-flow events fall into this category and could, in theory, be exempted.

With regard to non-compliance at base-flow, as no sites in the Humber RBD are designated as either bathing or shellfish harvesting sites (Defra, 2008b; Harley, 2008), it may be argued that, firstly, there is no legal requirement for improvements under the rBWD and SWD directives (EU 2006a and 2006b); and, secondly, to make improvements in Humber would not be the best use of scarce resources when there are officially designated sites elsewhere in need of remediation.

The WFD states that emissions should be as low as practicable, or rather, to comply with 'good' ecological status water quality should deviate only 'slightly' from normal conditions (EU 2000; Blacklocke et al., 2006). Furthermore, WFD compliance need not incur excessive, or disproportionate, implementation costs (Turner, 2007). The WFD provides no precise definitions of what constitutes 'slight deviation' or 'excessive costs'. Either are open to interpretation.

Efficient levels of pollution must be found. Economic theory tells us that the socially optimal level of pollution is achieved where the marginal external cost to society caused by each unit of pollution equals the per unit marginal abatement cost faced by the firm (Pearce and Turner, 1990). Pollution discharge rates, external costs and abatement costs are not evenly distributed throughout England and Wales. Different rivers are subject to spatially disparate concentrations of microbiological pollution arising from variable rates of inputs from human and livestock sources. Abatement costs are inconsistent due to a range of factors including variable costs for different remediation methods (e.g. purification lagoons, tertiary ultraviolet wastewater treatment). External costs vary depending on the type and severity of the pollution. Different types of recipient lose welfare in different ways (e.g. reduced profits for commercial shellfish harvesting companies or a loss of non-market welfare for recreational users such as anglers and rowers). Each of these factors impede the calculation of the socially optimal level of pollution or the calculation of costs and benefits arising from WFD compliance. Although cost effectiveness analysis is often used when compliance targets must be met, it neglects the value of the benefits that may arise from WFD implementation (Blacklocke et al., 2006; Lawlor et al., 2007). Pearce (1998) suggested that the directives on bathing water and drinking water would not pass a CBA. The WFD appears to fall into the same category. CBA studies are inconclusive and conflicting, in part due to the difficulties of quantifying non-market benefits (Whitelegg, 1993; Turner et al., 1994). For example, a report to the Scottish Executive (2002) estimated net benefits in Scotland arising from implementing remediation measures designed to comply with the WFD, whereas a similar study, conducted in England by Defra, estimated net costs of £530million per year, with over half of those costs falling on water companies (Water Industry Network Wales, 2007). The cost of pollution remediation is then passed onto household water bills (Haygarth et al., 2005): 2009 saw an average 5.8% increase in household water bills across England and Wales (BBC, 2008). It remains to be seen if proposed river improvements represent excessive or disproportionate costs to consumers.

WFD compliance seems likely to yield spatial variation not just in the distribution of the benefits of FIO reduction but also in the willingness to pay (WTP) for those improvements. The models produced in this research predict that improved water quality reduces the risk of ill health. Although many studies show an higher WTP to avoid increased ill health (EFTEC, 2002; Ready et al., 2004), the limited CBA evidence available suggests that preferences for health benefits alone may not justify WwTW investments (Kay et al., 1999). For example, water bills payers typically demonstrate higher preferences for more tangible benefits, such as odour reduction and a reduced risk of drains overflowing (Dwr Cymru, 2005; Scottish Water, 2004).

2.7 Conclusions

These are the first generic FIO models to be developed for the UK to incorporate direct measures of key FIO sources (namely human and livestock population data) as predictor variables (Crowther et al., 2011) and this research has pioneered the development of a transfer methodology which enables the models to predict FIO concentrations in unmonitored UK watercourses.

The following conclusions can be drawn from this research:

- The models used in this research provide a cost-effective diagnostic tool capable of identifying and predicting the sources and spatial distributions of microbial pollution.
- By incorporating human and livestock FIO sources as explanatory variables these models can be used to help apportion the responsibility for microbial pollution between the water industry and the agricultural sector.
- The regression modelling approach, by enabling spatially sensitive FIO function transfers, can inform integrated catchment management programmes, as required by the WFD, and offer insights into the optimal cost-effective mix of remediation strategies.
- The models can be used at a range of spatial scales and are capable of identifying non-compliant HRUs which may benefit from micro-scale BMPs.

Integrated basin-wide planning solutions must be developed to reduce the flux of microbial pollution to receiving waters, and an optimal mix of regional and site specific policy measures will be required in order to achieve the highest reductions. This chapter has demonstrated that transferable models of FIO concentrations may prove to be extremely cost effective diagnostic tools, helping to not just identify the spatial distribution of FIO sources (e.g. areas of high dairy livestock density) but to also make predictions of FIO concentrations in adjacent watercourses following a range of policy scenarios. The models offer real insights into the optimal cost-effective strategies for the delivery of WFD induced FIO remediation measures.

The different policy measures explored here may have unacceptable adverse consequences: compulsory destocking may have undesirable adverse impacts on farm incomes; farmers may circumvent restrictions on fertilizer application by applying manure to maintain grassland productivity (and herd size); adoption of the Nutrition Driven Food Policy may result in localised increases of FIO emissions as producers concentrate milk production into productive lowland areas. Results suggest that the installation of stream-side fencing may be the simplest and most practical policy measure. The installation of streamside fencing in areas of intensive milk production may be one of the most effective, targeted policy measures in reducing riverine FIO concentrations: The high-flow FC model predicts 58.6% reductions of FC immediately downstream of the high intensity dairy regions in the Aire subcatchment. BMPs have the potential to be effective and cost efficient. Dickson et al., (2005) observed a 40% reduction in high-flow FIO concentrations in the Brighouse catchment after the installation of BMPs. Meals (1996) found BMPs capable of reducing FIOs from dairy sources by up to 70%.

In conclusion, this work goes some way towards addressing an important knowledge gap which is of interest to a variety of stakeholders including other researchers, government agencies, the water industry, consumer groups and, most importantly, the general public. This research contributes to the emerging international debate on the use of farm best management practices and policy instruments to reduce FIOs and agricultural diffuse pollution e.g. Bateman et al. (2006a), Chadwick et al. (2008), Monaghan et al. (2008), Helming and Reinhard (2009), Hutchins et al. (2009), Maringanti et al. (2009) and Oliver et al. (2009).

2.7.1 Limitations and potential improvements to the research within Chapter 2

The River Tame catchment covers an area of 1,400 km^2 and includes the north east of the Birmingham conurbation, containing 1.8 million people (Crabtree et al., 1999). The Tame then flows through the Lea Marston purification lakes before its confluence with the Trent. These purification lakes are over 43 hectares in area and retain the water for up to 12 hours, allowing pollutants to settle out of the water. Even during storm conditions up to 90% of solids are retained (Martin and Brewin, 1994). As a result, downstream water quality is significantly improved under both base- and high-flow conditions (Environment Agency, 2004).

Engineered schemes are also extremely effective in reducing FIO concentrations at base-flow. Perkins and Hunter (2000) report the constructed wetland at Crow Edge sewage treatment plant as being 85-94% efficient at removing FIOs. Thurston et al. (2001) and Vrhovsek et al. (1996) claim over 98% efficiency for the constructed wetlands at the Pima County Constructed Ecosystems Research Facility and the Gradisce Constructed Wetland respectively. At high-flow the efficiency of reservoirs and purification lakes becomes much reduced. During wet weather the Tame continues to have a severe impact on water quality in the Trent (Crabtree et al., 1999).

Each impoundment situation is complex and detailed knowledge is required to calculate FIO concentrations. In line with the above literature, and based on the declared twelve-hour retention time, a predicted 90% reduction of FIOs at Lea Marston under base-flow conditions is not unreasonable (Kay, 2008c). The dominant mechanism of pathogen removal at Lea Marston is via particle-bound sedimentation. During high-flow conditions contaminated water is retained for shorter periods, sedimentation is reduced and the increased flow rate can cause resuspension and entrainment of pathogens (Crabtree et al., 1999; Upadhyay, 2002). For these reasons, a conservative estimate of the effectiveness of the purification lakes is 25% efficiency at high-flow (Kay, 2008c; Woods et al., 1984). Therefore, for the modelling exercises in Chapter 2, the PVM set the purification efficiency of the Lea Marston lakes at 90% efficiency during base-flow conditions and at 25% efficiency at high-flow.

However, the current method of using the models' predictor variable matrices (PVM), to estimate FIO concentrations emitted from reservoir catchments may be inaccurate. CREH's work has previously focussed on small rural catchments, with small proportions of land occupied by reservoir catchments. Errors in predictions of FIO concentrations may increase with reservoir catchment size or become unacceptably large in catchments dominated by reservoirs (cf. the rules governing the selection of study catchments in Chapter 1), or in catchments where engineered reservoirs are designed specifically to reduce riverine pollutants.

One solution to this issue will be to reconfigure the data matrix to more accurately reflect reduced FIO concentrations at reservoir outlets. This adjustment may be straightforward: it may be more accurate to use empirical data for pathogen concentrations at the reservoir exit (where such data exists), then adjust predictions of downstream FIO concentrations accordingly. The problem then becomes one of setting the rate of reduction at reservoir locations where empirical data are unavailable. Large reservoirs are very efficient (approaching 100%) at removing FIOs. Small reservoirs are less efficient. It may be possible to develop rules governing the efficiency of reservoirs based on factors such as reservoir size, river flow rate and the length of time the river is impounded.

Anomalies in the discharge routes of rivers affect the ability of the models to predict FIO concentrations. An example was encountered during this research: during base-flow the Upper Derwent catchment discharges via the Derwent but at high-flow excess water is artificially discharged to the coast via Sea Cut, an entirely different route (Crowther, 2008c). If anomalies such as this are uncorrected, the PVM would make inaccurate predictions. Further, more detailed knowledge of the vagaries of the river network is necessary to prevent this problem.

A natural extension to the research presented in this thesis would be to use actual river discharge data to enable the models to provide estimates of FIO export coefficients (number of organisms discharged per unit time), in addition to estimates of the concentrations of FIOs. River discharge data is available from the CEH National River Flow Archive for over 1300 gauging stations nationwide (2008b). The HRUs used within this research are consistent with those used within the ChREAM project and, as discussed in chapter 2, discharge data is only available for the main catchments within the Humber RBD. Discharge data is not available for all of the HRUs within the RBD. For those cases where river discharge data are unavailable, it would be possible to estimate base-flow discharges using the method described by Gustard et al. (1992).

The generic modelling approach is general enough to be implemented in a variety of empirical contexts. Chapter two demonstrated this by generating predictions of FIO concentrations using the datasets produced by the ChREAM (Bateman et al., 2006a) and LUAM (Jones and Trantor, 2008) research groups. Climate change will affect land use and livestock farming patterns and this in turn will redistribute microbial pollution sources. Preliminary work to model the effects of climate change on riverine pollution distributions in Humber has been undertaken (Fezzi et al., 2015) and the FIO models can be applied to those datasets to assess the impact of climate change on microbial pollution concentrations.

WFD compliance seems likely to yield spatial variation not just in the distribution of the implementation costs but also in the benefits of FIO reduction and in the WTP for river improvements. The natural extension to the research presented here is to produce a spatially explicit valuation of the non-market benefits and WTP arising from FIO reduction and, within a cost benefit framework, to assess the relative cost effectiveness of different remediation strategies, such as those discussed in Chapter 2. Analysis of this nature will be essential when assessing the optimum spatial differentiation and implementation of land use policies.

Spatial differentiation of land use policies are required to maximise benefits and minimise costs. The models presented here are capable of identifying non-compliant HRUs, such as the two HRUs situated along the upper Aire, discussed in Chapter 2. These areas may benefit from small scale, cost effective BMPs, such as installing fencing to prevent livestock accessing streams (Oliver et al., 2007). At a larger scale, integrated catchment management strategies have been successfully implemented overseas for many years (Heinz, 2003; US EPA, 2007). The increased uptake of holistic strategies is widely recommended and encouraged (UKTAG, 2005; Kay et al., 2006b; Turner, 2007). Effective catchment management schemes have been proven to prevent pollution to raw water *and* generate cost savings by minimising water treatment costs (Andrews and Zabel, 2003). The models produced in this research can be used to identify the optimal locations for engineered pollution remediation schemes capable of large scale improvements. With the modifications to the PVM reservoir calculation, discussed above, the models may provide improved predictions of downstream water quality improvements.

3 Chapter 3: River Water Quality: Who Cares, How Much and Why? Using Choice Experiment Methods to Elicit Preferences for River Water Quality

3.1 Introduction

The Water Framework Directive (WFD) requires substantial improvements to the quality of Europe's waters so that the 'good ecological status' of surface waters is achieved (EU, 2006a). One important motivation for the implementation of the WFD appears to be the creation of non-market social benefits, such as improved provision and opportunities for open-access recreation (see Articles 4, 9 and 11 of the WFD (Environment Directorate General, 2005)). For this reason, from a policy perspective, it is necessary to correctly identify and adequately measure the benefits of water quality improvements, as it is estimated that non-market values may represent significant components of water quality improvements (Hanley et al., 2006; Martin-Ortega and Berbel, 2010). The legislation recognises the crucial role of economics in its requirement that member states assess the social and non-market benefits of measures designed to achieve 'good ecological status'.

Previous research has demonstrated that poor water quality burdens society with substantial economic costs (Dodds et al., 2008; Pretty et al., 2003). Although early estimates suggest that the costs of implementing remediation programmes may be prohibitively expensive (Wither et al., 2005), the financial costs of remediation must be offset against the (often non-market) benefits of that remediation. Policymakers must also consider the spatial differentiation of pollution concentrations and costs of suitable land use policies and remediation strategies which could be implemented to deliver river quality improvements (Hampson et al., 2010). Even relatively inexpensive, cost-effective, remediation strategies to maintain sparsely populated headwaters at 'good ecological status', in subcatchments such as the River Ure in the Yorkshire Dales, may be disproportionate to the benefits created – particularly if those subcatchments are too remote to enable meaningful improvements in benefit values. Conversely, more heavily polluted rivers, such as the River Aire flowing through Bradford and Leeds, may require prohibitively expensive or technically infeasible remediation measures despite the potential for large benefit gains in terms of increased use values. The terminology within the WFD acknowledges such issues and allows derogations (article 4, par. 4, 5, 7) where remediation is either technically infeasible or where remediation costs are disproportionate to the benefits gained (EU, 2000). Under such circumstances technical infeasibility may justify extending the deadline for achieving 'good ecological status' up to 2027[11].

Disproportionate costs may trigger more achievable targets, such as a requirement for an 'acceptable ecological state'. Assessments of potential WFD investments, particularly for disproportionate cost extensions, require that costs and benefits are assessed within an cost-benefit framework (EFTEC, 2011; Görlach and Pielen, 2007). Interdisciplinary research on these highly contested issues is being undertaken across the EU (Balana et al., 2011; Galioto et al., 2013; Molinos-Senante et al., 2011; Postle et al., 2004).

In addition to the non-market benefit values of improved water quality there are other benefits which arise from improved water quality. These benefits, although quantifiable, are not necessarily easy to identify or calculate. The following paragraphs outline two examples.

Forced closures of recreational facilities, due to poor water quality, are commonplace nationwide. The UK's premier freshwater sports venue, the National Water Sports Centre on the River Trent, periodically suffers closures and revenue losses due to poor water quality (Robinson, 2015). Leisure activities are regularly curtailed on the Yare (the present case study area) due to poor water quality. During the summer of 2015 triathlon competitions and swimming activities at Whittingham Outdoor Education Centre had to be cancelled due to the health risks created by toxic blue-green algal blooms (Lines, 2014). These financial

[11] According to article 4.4.of the WFD, time extensions for achieving 'good status' or 'good potential' of water bodies shall not be longer than two planning cycles beyond 2015. As a consequence the year 2027, the deadline of the third WFD planning cycle, is currently the final date for achieving 'good status' or 'good potential'. The only current exemption is if natural conditions are the reason for not achieving the objective. This means that Member States will have to decide whether they can realistically expect to achieve 'good status' or 'good potential' for the respective water bodies by 2027 or, if not, to set less stringent objectives according to article 4.5 of the WFD (European Commission CIRCABC, 2016a). Discussions on post-2027 arrangements are ongoing (European Commission CIRCABC, 2016b).

losses due to poor water quality can seriously affect the viability of smaller recreational clubs and the social and cultural networks which they support.

The dose-response relationship pertaining to illness from contaminated freshwater is poorly understood (Fewtrell et al., 1992) and recreational users are frequently faced with a choice between an unquantified (but real) health risk or forgoing their recreational activities. Gastro-enteric illnesses such as the 'Trent Tummy' are recognised conditions among recreational users at the National Water Sports Centre. Unfortunately, due to necessity, the risk of ill health caused by microbiologically polluted water is often downplayed or regarded by rowers and swimmers, across the UK, as an unfortunate by-product of undertaking their recreational activities (Heron, 2014). Remediation of riverine pollution yields reductions in ill-health, the number of working days lost due to that ill health and reductions in the costs of medical treatment for those experiencing severe ill health.

Benefits arising from pollution remediation may be intangible and, consequently, overlooked. It may be argued that derogations on grounds of excessive costs should not proceed without comprehensive assessments of environmental and socio-economic benefits. Stated preference valuation techniques have a long history of use within environmental economics to model preferences for environmental change and previous research has explored the willingness to pay for pollution remediation within a cost-benefit framework (Bateman et al., 2006a; Glenk et al., 2011; Metcalfe et al., 2012; Van Houtven et al., 2007).

However, previous analyses of the benefits arising from pollution remediation systematically overlook the distinction between the ecological and microbial quality of river water. Within UK research, economic valuation studies have typically assessed WFD benefits in ways which conflate ecological improvements with the value of recreational gains and, therefore, assess water quality as a single attribute of preference (Bateman et al., 2011; Ferrini et al., 2014). This practice must be revised in order to correctly ascertain the true values people hold for different attributes of water quality. Most river visitors are not specialised users (e.g. anglers, swimmers or boaters), but those who walk along river banks for recreation. Although visitors directly benefit from ecological improvements (i.e.

improved odour, clearer water, less algae, a greater range of flora, fauna, birds and fish to enjoy, etc.), this research explores the hypothesis that respondents may hold values for the improved recreational opportunities arising from microbial pollution remediation, quite separate from their values for ecological water quality improvements.

To summarise, this research is motivated by the obligation imposed by the WFD to fully identify and understand the benefits of riverine pollution remediation. It aims to further the knowledge on non-market valuation of river water quality by examining the relative importance of ecological or recreational water quality improvements. This research has both academic and policy relevance. Decision makers require detailed information about the spatially differentiated non-market benefits obtainable via ecological, recreational or a mix of water quality remediation schemes. It is hypothesised that the non-market benefits relating to each type of remediation offer heterogeneous values and, in order to optimise the value of an investment in water quality, efficient investment must be informed by the various factors that influence value. There is little reason to invest if there is no little or no benefit to that investment, or if that investment would be better directed elsewhere. Improving the accuracy and availability of data pertaining to the relative values of the different types of non-market benefits is a strong motivation for this research. Academically, this work extends the research literature by disaggregating the values respondents hold for the ecological and recreational attributes of water quality using appropriate research methodologies. From a policy perspective this research enhances the ability of the policy-maker to more fully understand, calculate and incorporate potential benefits and, thus, produce more accurate cost-benefit analyses.

To begin, the literature is reviewed to provide an overview of the legislative imperatives to incorporate meaningful assessments of all costs and benefits into pollution remediation strategies. The economic valuation methods typically used to assess non-market benefits are discussed, with attention given to the advantages and disadvantages of stated preference techniques; in particular conditional logit (CL) and latent class (LC) model specifications, the modelling approaches used within this research. Typical socio-economic and cultural differences between respondents, which can produce preference heterogeneity, are identified and discussed. The literature on water quality information is discussed. More specifically, the way in which water quality information has previously been used to underpin water quality characteristics presented as choice information within valuation experiments. The literature review closes with an overview of the relevant findings from the Catchment hydrology, Resources, Economics and Management (ChREAM) project, the predecessor to this research.

An overview of the research challenges, the hypotheses and objectives of this research is then provided. This is followed by a summary of the case study area within which this research is conducted. The principles guiding the design of the survey and survey instruments are then discussed. The choice experiment design, choice attributes and attribute levels are presented, followed by the rationale for the recruitment of respondents. The modelling strategies used, namely CL and LC analyses, are discussed before the results of the research are presented.

The results section opens with an overview of the results from a pilot study and a discussion of the summary statistics of the main survey data. CL models, results and willingness to pay (WTP) estimates are presented and discussed. This is followed by the presentation and discussion of the results, WTP estimates and postestimation statistics obtained using a LC modelling approach.

The research findings are discussed followed by concluding comments set within the context of the academic interest and policy relevance of those findings. The chapter concludes with an overview of the limitations of this research and explores avenues for future research identified throughout the course of this work.

3.2 Literature review

Cost effectiveness analysis and cost benefit analysis are the two main methods adopted for economic assessment under the WFD. Cost benefit analyses help guide the planning and management communities into channelling resources into the remediation projects that yield the greatest gain in net benefits to society. This is particularly useful where alternative policies, or remediation strategies, exist or where financial resources are limited (Field, 2008; Turner et al., 1994).

Typically money is used as the unit of measurement, enabling the commodification of the services which the natural environment provides (Perman et al., 2003). Unfortunately it is often difficult to apply accurate values to some costs or benefits, particularly the external unpriced effects of environmental losses caused by pollution, or the non-market benefit values associated with projects designed to improve non-market public environmental assets (Bateman et al., 2011; Turner et al., 1994; Whitelegg, 1993). Furthermore, even where a market price does exist, it may often only represent an approximation of value (Boardman et al., 2014).

Official UK guidelines state that costs and benefits, however difficult to monetize, must not be ignored but must be "quantified where possible and meaningful" (H.M. Treasury, 2003). Without quantification, potential benefits can only be described qualitatively. Reliance on qualitative data may negatively influence policy relevance as policymakers tend to overlook the qualitative, preferring instead the (often) simpler quantitative forms of assessment and communication in order to address complicated questions about the nature and significance of the problem to be addressed (Gysen et al., 2006).

A range of quantitative economic valuation techniques have developed over time to monetise preferences regarding environmental costs and benefits (Bateman et al., 1993; De Groot et al., 2006; Hanley et al., 2009). Choosing the most appropriate valuation technique requires a consideration of the nature of the environmental goods to be measured (Bateman et al., 2011). The economic valuation of environmental resources typically takes two main approaches: revealed or stated preference. These valuation methods, their usefulness and limitations, and some of the issues arising from their use which may impact upon this present research, will now be discussed. There is a growing literature providing detailed analyses of environmental valuation techniques, e.g. Bateman et al. (2001), Birol et al. (2006), Carson (2000) or Champ et al. (2012).

One approach used to measure respondents' preferences for an environmental resource is to reveal their preferences from an analysis of their behaviour. Revealed preference techniques have a long history. Hotelling (1949) proposed an assessment of the non-market value of parks using an assessment of the respondent's travel costs. Hotelling's suggestion stimulated research in revealed preference valuation methods, leading to the development of hedonic pricing (Ridker and Henning, 1967) and travel cost (TC) (Clawson, 1959) valuation methods.

Within the context of environmental valuation, hedonic pricing is used to estimate the economic values of environmental goods that directly affect market prices. It is most commonly applied to variations in property prices which reflect the value of localised environmental attributes such as urban trees, or access to watercourses (Malpezzi, 2003; Tyrväinen, 1997). Hedonic pricing may inform us of the localised effects that river water quality may have on property values. Unfortunately, it is unsuitable for this present research because it cannot inform us of differing benefit values beyond close proximity to the river.

The basic premise of TC method is that the time and travel costs arising from a respondent's real world behaviour in order to visit a site represent the price of access to the site. Much use has been made of TC to estimate the non-market use values of environmental resources, such as recreational fishing (Shrestha et al., 2002), or the preservation of environmentally vulnerable freshwater environments (Fleming and Cook, 2008). These values can be used within a cost-benefit analysis (CBA) framework to assess the use values of the environmental resource. By definition, TC method cannot capture the values held by non-visitors (Parsons, 2003).

We know non-visitors hold values for environmental goods. So, to quantify the values held by visitors and non-visitors we must take a stated preference approach to valuation. There are two main types of stated preference methods used where direct quantification of value is not possible or where there is no observable respondent behaviour to measure: contingent valuation and choice experiments. These methods generate and quantify the non-market values held by respondents for environmental resources and these values can be used within a cost benefit framework to represent non-market benefit values. There is a considerable literature on stated preference techniques. For detailed critiques of the two methods please see, for example, Boxall et al. (1996) and Hanley et al. (2001). Here is an overview of the two techniques and their methodological advantages and disadvantages.

Stated preference techniques also have a long history of use, dating back to suggestions made by Ciriacy-Wantrup (1947). Contingent valuation method (CVM), one of the earliest stated preference techniques, estimates the values held by an individual by directly asking that individual a question which involves paying for an improved environmental good. The respondent's WTP to secure the improved good, or their willingness to accept (WTA) in compensation to forgo the improvement, can then be assessed from their answer (Pearce, 1998). CVM has a long history of use for deriving non-market values, which, as previously discussed, are essential components of the total economic value of environmental resources (Pearce et al., 2006). For example, Smith and Desvousges (1986) and Desvousges et al. (1987) carried out one of the seminal CVM studies to examine WTP for improved water quality in terms of its suitability for recreational use. Although the study was hypothetical, it produced meaningful and policy relevant results. CVM continues to have relevance and has been used on a wide range of large scale CBA exercises, e.g. to examine management issues surrounding water quality (Bateman et al., 2008) or losses from water pollution (Carson et al., 2003).

A problem with the CVM approach is its reliance on the accuracy of the survey instruments used to present precise changes to water quality within a single choice question. This issue of informational accuracy has challenged researchers, particularly in situations where precise data on water quality standards has been unavailable (Kay et al., 1994). The reliance on accurate information, relating to a specific definition of water quality, is mitigated within the choice experiment (CE) approach. Louviere (1996) explained that choice experiments rely less on the accuracy of a single description of water quality (as in CVM), but more on the comprehensive description of hypothetical situations. This is an important distinction, particularly when we consider the availability of accurate information relating to water quality standards, the respondent's comprehension of the environmental issue, or the accuracy and accessibility of the survey instruments, all of which will be discussed shortly.

Choice experiments typically use multiple choice sets comprised of different attributes, and different levels within those attributes, to present a series of hypothetical choice situations reflecting different states of the environment. The respondent is asked to choose their preferred alternative from each choice set and their choice reflects the trade-offs they make between the different attributes within each hypothetical scenario. When a price attribute is included within the choice set, probabilistic modelling enables the estimation of the marginal utility the respondent holds, in monetary terms, for each of the non-price attributes (Boxall et al., 1996; Hanley et al., 2006). This approach is of enormous value to policymakers as it allows the assessment of changes in the marginal values for different levels of the environmental good. The repeated sampling method central to choice experiments reduces the concerns of lower informational efficiency affecting the model structure, as is the case in CVM (Carson, 1991).

Presenting the respondent with multiple choice sets also has the distinct advantage of providing the economist with rich information on intra-respondent preferences for different attributes, levels and scenarios. This is grounded in Lancaster's (1966) conceptual framework, which assumes that respondent's utility for a good can be decomposed into the attributes of that good. The format of choice experiments are also thought to minimise the incidence of yea-saying, protest bids or other strategic behaviours which can be encountered when using CVM (Bergstrom et al., 1989; Cummings et al., 1994; Hanley et al., 2001). For example, research by Day et al. (2012) compared the rate of positive responses using CVM and CE methods and found a significantly lower rate of positive responses, and, consequently, lower levels of yea-saying when using the CE method.

These advantages over CVM have, in part, led to the increased adoption of choice experiment methods in water quality studies since the late 1990s

(Adamowicz, 2004). This trend has continued apace over the last ten years, with the method used to assess a diverse range of water quality issues including assessments of wetland conservation projects (Birol et al., 2006), multi-country assessments of benefit transfer in water conservation projects (Brouwer et al., 2015), adaptations to river use (Andreopoulos et al., 2014), river restoration (Bliem et al., 2012), and improvements in river ecology (Hanley et al., 2006).

Some of the issues (e.g. heterogeneous preferences) which may affect the accuracy of utility estimates, when using CE methods, are now discussed. The efforts to minimise the effects of these issues are discussed subsequently, within the section of this chapter which describes the experimental design.

The two stated preference methods (CVM and CE) outlined above, belong to the family of methods associated with random utility theory (RUT) (Bennett and Blamey, 2001). A central tenet of RUT is homogeneity in respondents' preferences. In reality, respondents are frequently imperfectly informed on, or differently motivated by, environmental issues. These differences introduce preference heterogeneity, inconsistent with RUT, into any subsequent probabilistic modelling of that choice data. There is considerable, and often heated (Weikard, 2002), debate surrounding the ability of CVM to accurately estimate non-market values as, due to their ethereal nature, the non-use component of non-market values can be highly subjective (Barbier, 1993). For example, Schultze et al. (1983) found the existence and preservation values held by respondents to be far higher than respondents' use values. These heterogeneous differences in personal motivations can prove to be difficult to identify, classify and quantify.

Another issue, particularly applicable to the calculation of the non-market benefits of river water quality improvements, is that of defining the spatial boundaries of a remediation project and estimating the number of potential beneficiaries (Bann et al., 2003; Brouwer et al., 2010). Errors in estimating the numbers of beneficiaries of an environmental change can compound errors in estimates of per-person WTP when aggregate values are calculated (Bateman et al., 2002). To minimise this problem researchers have increasing used geographical information systems to incorporate spatial variables into their research in order to assess and calculate distance decay effects (Bateman et al., 2006b; Brainard et al., 2002). The UK Aquamoney project (Bateman et al., 2008) found that respondents' WTP for river improvements decreased with distance. Georgiou et al. (2000) found a partial solution to the problem of defining spatial boundaries: they calculated both a distance-decay effect and a limited distance boundary for values relating to water quality improvements. By calculating the average WTP, adjusting for distance decay and spatial boundaries, then multiplying the averages by the number of people affected by the environmental change, analysts can then obtain an estimate of the total value placed on that environmental change (Moran, 1999; Turner et al., 1994). It is often the case that individual preferences and corresponding preference boundaries are heterogeneous across users and topographical locations. For example, Van Houten et al. (2007) identified significant heterogeneity across respondents' WTP for environmental improvements, due to geographical factors, in their meta-analysis of US stated preference studies.

There is a growing literature on non-market valuation methods to estimate respondent's willingness to pay to avoid environmental health risks. Previous research has found that individuals who perceive greater environmental health risks are generally more likely to be willing-to-pay (and willing-to-pay more) for a given reduction in that risk. For example, Sukharomana and Supalla (1998) found that that an individual's WTP for groundwater improvements increased if their perception of risk was greater. Similarly, Georgiou et al. (1998) found that an individual's WTP for improvements in bathing water quality was strongly correlated with that respondent's perception of the health risks associated with exposure to that polluted water.

Georgiou et al. also found that an individual's WTP for improvements in bathing water quality was not only dependent on their perceptions of health risks but was also strongly correlated with their socio-economic status. Hunter et al. (2012), found that socio-economic factors, such as income or the number of environmental memberships held by the respondent, significantly influenced that respondent's WTP for reducing toxic cyanobacterial blooms in recreational water. Hoyos, et al. (2009) found that age and cultural identity caused heterogeneous preferences with regard to recreational resources.

Within stated preference studies attribute non-attendance, where the respondent ignores an choice attribute, is found to occur frequently (Scarpa et al., 2009). Identifying and accurately treating process heterogeneity is necessary for accurate utility estimation as, left untreated, it significantly impairs the efficiency of coefficient estimates resulting in over- or under-estimates of the marginal WTP for specific attributes (Campbell et al., 2011).

McFadden (1973) proposed modelling utility in terms of the characteristics of the choice alternatives, interacted with the attributes of the respondent. McFadden's development, the CL model, is particularly suited to modelling choice behaviour, where the explanatory variables include attributes of the choice alternatives, e.g. water quality attributes or price, as well as respondents' socio-economic characteristics, e.g. age or income. CL models are also useful when the number of combinations of choice alternatives are large, as is frequently the case within choice experiments.

The assumptions of the CL model are that the error term is independent and identically distributed (IID) across observations, and is uncorrelated across options (Luce, 1959). As has been discussed, heterogeneity in the error term can frequently occur for a number of reasons within CL modelling. Researchers have sought methods to relax the strict assumptions of independence of irrelevant alternatives (IIA), such as the assumption of homogeneity. Accounting for heterogeneity within random utility based models is necessary in order to estimate efficient unbiased models of respondents' preferences (Boxall and Adamowicz, 2002; Yatchew and Griliches, 1985). Early approaches involved parameterizations of the scale factor in the random parameter logit (RPL) method (Layton, 1996; Train, 1998). Although these approaches incorporate heterogeneity, it has been argued that they are poorly suited to explaining the sources of that heterogeneity (Boxall and Adamowicz, 2002).

The differences between the mixed logit (MXL) family of models are sometimes unclear within the literature. McFadden and Train (2000) clarify that mixed logit is the family name of the different types of models (e.g. RPL, error component logit (ECL) and LC), used to reflect heterogeneity in preferences. Mixed logit reflects the fact that the choice probability is a mixture of logits with a specified mixing distribution and the term ‘mixed logit’ encompasses any interpretation that is consistent with the functional form (Train, 1999). There are different specifications and corresponding advantages across the family of MXL models; their commonality arises in the integration of the logit formula over the distribution of unobserved random parameters (Revelt and Train, 1998). MXL models are a generalization of standard logit that do not exhibit the restrictive IIA property and explicitly account for correlations in unobserved utility over repeated choices by each respondent (Revelt and Train, 1998).

The ECL interpretation is amenable to the analysis of complex substitution patterns. (Department for Transport, 2014). The RPL interpretation allows the data considerable freedom to directly reveal the form of any inherent taste variation, without recourse to any particular segmentation (Green and Hensher, 2003). In their most basic form, RPL models provide a mean with a distribution around that mean (i.e. estimates of the first and second moments, the mean and standard deviation) of tastes across the population of interest. The preferences of all respondents are variable along a continuous range. The mixing distribution, G, may come from a continuous parametric family, such as multivariate normal or log normal, or it may have a finite support. When G has finite support, MXL models are also called LC models (McFadden and Train, 2000). LC reveals allows segmentation for distinct classes (groups) of respondents. Although preferences vary across classes, the heterogeneity of preferences is smaller within classes; the segmentation provides a more precise estimation of the preferences held by class members. Although LC is less flexible, in that it approximates the underlying continuous distribution with a discrete one, it does not require the analyst to make specific assumptions about the distributions of individual heterogeneity (Green and Hensher, 2003).

Both RPL and LC offer alternative ways of capturing unobserved heterogeneity and other potential sources of variability in unobserved sources of utility[12]. The different methods are a matter of taste, but some specifications of MXL can produce clusters of results which can be difficult to interpret and it can be difficult to identify which error components to include in a MXL specification (Brownstone,

[12] For a more complete discussion on MXL please see McFadden and Train (2000) and for the technical differences between RPL and LC, please see Green and Hensher (2003).

2001; Department for Transport, 2014). LC analysis can be used to overcome, or minimise, these difficulties and is now discussed further.

LC analysis has a history of being used within market research. McFadden (1986) proposed the integration of choice information with socio-economic and attitudinal/psychological information to create latent variables in order to understand choice behaviour. Revealing latent attitudinal data in this way can provide significant opportunities to enrich economic analysis (Boxall and Adamowicz, 2002). The underlying assumption is that a respondent's behaviour within a CE is a manifestation of their underlying latent preferences (Morey et al., 2006). These latent attitudes can enable the identification of otherwise unobservable subgroups, or classes, within a sample of respondents and LC is used to explain the choice behaviour of those different classes.

Respondents within different classes will answer preference elicitation questions differently from one another due to their underlying latent characteristics. Although intra-class respondents may display relatively homogeneous preferences, the functional form of LC analysis places no restrictions on class membership probabilities, allowing for a wider range of preference heterogeneity within a class. This solves the limitations of IIA on the distribution of the preference parameters as there is no longer the assumption that parameters are normally distributed (Morey et al., 2006).

There are other advantages of using a LC framework. CL model structure does not control for intra-respondent panel data but instead treats all unobserved factors across observations as independent and unique. This is a serious misspecification when a respondent is presented with multiple choice tasks, as commonly used within a CE format, as we would expect that respondent to be influenced by that same (latent) bundle of unobserved factors throughout the CE sampling process (Train, 1998). LC modelling neutralises this difficulty by retaining the data's intra-respondent panel structure (Kemperman and Timmermans, 2006).

To mitigate against process heterogeneity by recreational water users, Campbell et al. (2011) modelled data with suspected attribute non-attendance using a LC framework, where classes were defined using LC rules that recognise the possibility of non-attendance to one or more attributes.

There has been a growing acceptance of LC analysis within environmental economic analysis. Provencher and Moore (2006) use the method with choice data to understand the preferences of recreational anglers. Boxall and Adamowicz (2002) estimate latent preferences for wilderness recreation using attitudinal and choice data. Shonkwiler and Shaw (2003) use socio-economic and choice data in a LC framework to assess reservoir recreation.

LC models are less computationally demanding than continuous mixture models and provide easy to interpret willingness to pay measure (Hess et al., 2011). So, from a policy perspective, LC analysis tends to be informative, yet simple to interpret (Scarpa et al., 2005) and it is conceptually appealing as it recognises that a population consists of subgroups distinguished from one another by their latent preferences which are shaped by their socio-economic characteristics and personal attitudes (Morey et al., 2006; Provencher and Moore, 2006). Knowing the attitudes of different groups helps environmental managers respond more appropriately to the preferences of those groups in relation to the environmental good being investigated. Moreover, LC method can identify subgroups which may be apathetic towards changes in the environmental good and, if such a group has been identified, LC method can estimate its size (Morey et al., 2006).

As mentioned previously, economic valuation studies of changes to the quality of the natural environment are only ever as valid as the natural science data upon which they are based. Here the completeness and accuracy of the natural science data is reviewed, starting with the measurement of pollutants.

This research focuses on the ecological and microbiological quality of the river's water. Both parameters exhibit a dose-response relationship, whereby the higher the concentration of the pollutant, the higher the risk of damage to both human health or the wellbeing of the components of the riverine ecosystem. The ecological quality of water is linked to the loadings of potentially harmful chemicals within the water. Concentrations of nutrients such as nitrogen or phosphorous have long been known to be associated with the overall ecological health of the aquatic environment, with excess concentrations negatively affecting the abundance of fish and other flora and fauna (Camargo and Alonso, 2006; Carpenter et al., 1998; Correll, 1998; Van Houtven et al. 2007). Raised concentrations of nutrients may also cause toxic blooms of cyanobacteria to develop, creating an hazard to human health (Chorus et al., 2000; Pilotto et al., 1997). Please see Chapter 2 for a discussion of the risk of ill health due to excess exposure to microbiological pollution.

These, and other facets of the overall water quality span a continuum, from that which can easily be perceived (e.g. turbidity or algal growth), to characteristics imperceptible to human senses (e.g. dissolved oxygen content or concentrations of microbial organisms). These different aspects can be problematic because respondents are often heterogeneously and imperfectly informed about ambient environmental risks (Konishi and Coggins, 2008). Respondents' perceptions of water quality may bear no resemblance to the actual water quality. Factors such as anecdotal evidence, which creates or reinforces a local reputation for water quality, may be more important determinants of respondents' preferences than the actual water quality (Binkley and Hanemann, 1978; Happs, 1986). Langford et al. (2000) suggest that public perceptions of risks from polluted recreational waters can be explained by cultural theory. They argue that social constructions shape and form an individual's worldview and influence their cognitive judgements about the magnitude and acceptability of risk. Without adequate information on water quality, personal, social or cultural misperceptions may bias respondents' preferences and WTP estimates for environmental improvements or health risk reductions.

Clearly, it is important to minimise respondents' misperceptions of risk in order to minimise experimental error. To aid experimental accuracy, researchers have sought to produce objective water quality indices based on different combinations of scientifically quantified parameters based on expert judgment (Bouwes and Schneider, 1979). Unfortunately, providing the respondent with accurate information, on its own, is not enough. Green and Tunstall (1999) found that face-value comprehension of the (often) complex choice information provided within non-market valuation experiments cannot always be taken for granted. Early stated preference experiments frequently presented the attributes of non-market goods to respondents as a table of values, but it has been found that respondents can have difficulty evaluating numerical or categorical data within choice experiment options. Hibbard et al. (2002) found that respondents chose inferior options 45% of the time when presented with tabulated numerical data. When visual representations of the same data was presented to those same respondents, error rates fell to 16%. Willingness to pay for quality improvements depends upon the respondent's ability to accurately perceive water quality changes. Researchers have sought to aid the comprehension of respondents by using water quality ladders to portray water quality information.

Early stated preference experiments, such as Mitchell and Carson's (1981) contingent valuation study of US water quality, have used derivatives of the Resources for the Future water quality ladder devised by Vaughan (1981). Vaughan consulted a number of sources, including the National Sanitation Foundation's water quality index (Booth et al., 1976), to devise a numerical index linking potential recreational water uses (swimming, game fishing, course fishing, boating) to minimum acceptable standards for five measurable water quality characteristics (faecal coliforms, dissolved oxygen, maximum biochemical oxygen demand, turbidity and pH).

Mapping water quality characteristics onto water quality ladders can be problematic. Primary water contact is typically defined by activities, such as swimming, which involve full or partial immersion into the water, with a high possibility of ingesting water. Secondary activities, such as boating or angling, are defined by their reduced contact with the water and lower risk of ingesting water (Dorevitch et al., 2011). Whilst research has been undertaken to define safe microbial concentrations for primary contact activities (Kay et al., 1994; Kay et al., 2004b; Prüss, 1998), there are difficulties mapping values for safe secondary recreational use due to an ongoing lack of available epidemiological information (Cumming et al., 1986; Fewtrell et al., 1992; Stoner, 1978; US EPA, 1986). Vaughan (1981) identified a tenuous link between 2000 faecal coliforms per 100ml as the upper limit at which water becomes unsuitable for secondary activities. The value appears to be based either on the National Technical Advisory Committee standards (1968) or linked to its historical use by several US states - even though there has been no agreement between states on the suitability of that figure (US EPA, 2003). The most recent guidelines issued by the US EPA (2012) and the WHO (2003) are still unable to satisfactorily quantify safe microbial limits for secondary contact (US EPA, 2015). Despite the paucity of microbiological water quality data, standardised water quality parameters are increasingly used within water quality ladders, particularly for nutrient concentrations, or other factors (e.g. pH or maximum biochemical oxygen demand) which affect the ecological quality of the water. This increased standardisation enables the assessment of benefit transfers of utility values within economic valuation studies (Van Houtven et al., 2007).

In the UK, the most recent DCE water quality studies include Hanley et al. (2005, 2006), Glenk et al. (2011), and Metcalfe et al. (2012). Also of particular interest is a DCE, employing similar variables to the present study (i.e. water quality disaggregated into a series of ecological quality attributes and a recreational attribute), conducted in the Republic of Ireland by Doherty et al. (2014).

Hanley et al. (2005) tested the impact of different price vectors on CE results for improvements from ‘fair’ to ‘good’ status in the ecological quality, aesthetic and bankside condition attributes of the River Wear in County Durham. Two CE designs were employed: Design A contained prices ranging from £2-24, Design B used lower prices, ranging from £0.67-£8. Although implicit prices were lower in the low-price sample than in the high-price sample (in some cases, by as much as 45% lower), they found that WTP did not vary significantly across the two price vectors suggesting robustness to the framing effects such vectors might induce in respondents. Further design details and estimates for pooled WTP are shown in Table 24.

Hanley et al. (2006) extend their previous research to embrace both the Wear in County Durham and the river Clyde in Central Scotland. This extension allowed the authors to test the transferability of value estimates between these two rivers. While benefits transfer tests were rejected with preferences and values differing significantly across the two studies, other results suggest that river ecology provides the main driver of values; a similar result to that found in my own study. Hanley et al., also found significant preference heterogeneity within and across samples, observing that those living near the Clyde valued improvements to their local river more highly than people in Durham valued identical improvements to their local river. Given this overall WTP was higher in the River Clyde sample compared to the River Wear sample (Table 24). For the River Wear, Hanley et al found that people placed insignificantly different values on the three attributes of river quality, whereas for the River Clyde, larger differences were found in attribute values, with aesthetic improvements being valued appreciably lower than either river ecology or bankside conditions.

Glenk et al. (2011) describe the state of Scottish rivers and lochs and assess respondents' preferences for the potential future status of these waterbodies. The attributes used in their DCE are descriptions of the potential status of the rivers and lochs. The levels for the attributes are varying quantities of the water bodies that will be at the achieved environmental standard by the end of the given time frame. They find WTP per household per year of £1.05 and £0.89 for rivers and lochs, respectively, for a 1% marginal improvement in water quality. These estimates for 1% marginal improvements are similar to those observed by Metcalfe et al., (see Table 24).

Metcalfe et al. (2012) carried out a large-scale investigation of the value of the implementation of the WFD for all water bodies in the UK, employing a similar approach to that used by Glenk et al., e.g., the attributes used in their DCE are descriptions of the potential status of the water body in a number of years' time, with the levels for the attributes representing varying qualities of the rivers that will be achieved by the end of the time period. They assessed respondents' preferences for the potential future status of rivers using multiple elicitation methods (a DCE and two forms of CV). Their DCE estimated WTP per household per year of between £0.36 and £0.95 for a 1% marginal improvement in river water quality. The value of a scenario in which 95% of rivers are brought from 'fair' to at least 'good' ecological status by 2015 was found to vary from £30.70 to £76.20 per household per year via the DCE responses (see Table 24).

Table 24: prevailing DCE estimates of WTP for UK river water quality studies

	Sample size	Attributes	Design setting	WTP (£)
Hanley et al. 2005	210 design 1, 120 design 2 (River Wear – Durham)	Ecology (good, fair) Aesthetics/litter (good, fair) Bankside condition (good, fair) Price (water bill): (design 1): 2, 5, 11,15, 24 (design 2): 0.67, 1.67, 3.67, 5, 8	CNL-RPL 2 choice task +status quo (8 choices)	Pooled sample: Ecology £11.12 Aesthetics £11.10 Bankside £11.94 (£/hh/yr/period to improve from fair to good)
Hanley et al. 2006	210 per each river - general population quota sampling (River Wear and River Clyde)	Ecology (good, fair) Aesthetics/litter (good, fair) Bankside condition (good, fair) Price (water bill): 2, 5, 11,15, 24	CNL-RPL 2 choice task +status quo (8 choices)	River Wear: Ecology £12.19 Aesthetics £12.07 Bankside £12.67 River Clyde: Ecology £38.70 Aesthetics £28.57 Bankside £42.99 (£/hh/yr/period to improve from fair to good)
Glenk et al. 2011	144 for each location - general population quota sampling (Scottish lakes and lochs)	Rivers in 7 years: percentage of low, medium and high Lochs in 7 years: percentage of low, medium and high Price (water bill): 5,10,20,40,50,75,100	CNL-RPL 2 choice task +status quo (8 choices)	Rivers in 7 years: £1.05 Lochs in 7 years: £0.89 (£/hh/yr/ for 1% marginal improvement)
Metcalfe et al. 2012	1487 across 50 sites general population quota sampling (Nationally)	HighL8: proportion at high quality in local area in 8 years HighN8: proportion at high quality nationally in 8 years High20: proportion at high quality in local and national areas in 20 years Price (water bill): 5,10,20,30,50,100, 200 (levels were pivoted around the respondent local status quo levels)	CNL-RPL 2 choice task +status quo (7 choices)	HighL8: £66.40 HighN8 £76.20 High20: £30.70 (£/hh/yr/period to improve from fair to good) HighL8: £0.77 HighN8 £0.95 High20: £0.36 (£ hh/yr for 1% marginal improvement)

A common characteristic of the studies reported in Table 24 is the inclusion of a status quo option in the choice task. None of those studies sought to separate the microbiological/recreational component of water quality from the ecological attribute.

Not undertaken within the UK, but of relevance to this study, is the DCE conducted by Doherty et al. (2014) in the Republic of Ireland. They observe that 'a consequence of focusing on just the ecological status of the water bodies being analysed is that the marginal value of a specific characteristic of a waterbody (e.g. the marginal value of a change in the recreational or aesthetic attribute) cannot be estimated.' Within their research, they disentangle ecological water quality characteristics into aquatic ecosystem health, water clarity, bankside condition and odour attributes. Although the study uses a status quo option, it does provide a disaggregated attribute to describe recreational water quality. The lowest valued attribute was associated with recreational access, a finding which echoes the results reported here. Their results suggest aggregate compensating surplus WTP per person per year of €129 for 'good' water quality. Their survey was generalised to apply to all water bodies (including rivers, lakes and the sea) in Ireland, rather than specific rivers. (i.e. including good levels of ecosystem health, clarity and odour, suitability for recreational use and having good bank condition).

This literature review now closes with an overview of the predecessor to this present research, the ChREAM project (Bateman et al, 2006a). ChREAM used the water quality ladder proposed by Hime et al. (2009), which was developed using UK Technical Advisory Group ecological guidelines (2008). ChREAM also developed a stated preference choice experiment design to determine respondents' preferences for different water quality levels. The CE attributes used were price and four water quality levels; Red as the baseline then, in ascending order of quality, Yellow, Green and Blue. When defining the four water quality levels the ecological and recreational components of water quality were conflated. The Red attribute level defined water with Low ecological and Low recreational quality: capable of supporting only a limited range of wildlife species, unable to support fish and unsuitable for swimming or boating. The Yellow water quality level defines water able to support a limited range of coarse fish (but no game fish) and an improved range of birds and other wildlife. Yellow water quality is suitable for boating, but continues to be unsuitable for swimming. Green water quality is suitable for a wide range of coarse fish (but still unsuitable for game fish), a wider range of other wildlife and is suitable for both swimming and boating. The Blue water quality level is composed of water with High ecological and High recreational quality attributes. It is suitable for the most complete range of wildlife, including course fish and pollution-sensitive game fish. Blue water is suitable for all recreational activities, including swimming and boating. For a comprehensive discussion of the water quality levels used in ChREAM please see Hime et al. (2009).

Table 25: coefficients derived from a CL model of data from a ChREAM survey of 1100 respondents in Leeds, UK.

Choice	Coefficient	Standard Error	P>z	95% Confidence Intervals	
Price	-0.021	0.001	0	-0.022	-0.019
Yellow	0.171	0.012	0	0.147	0.194
Green	0.334	0.011	0	0.312	0.355
Blue	0.439	0.012	0	0.416	0.461
LL -7271.68					
Pseudo r^2 0.12					

Table 25 shows the complete and transitive results of a CL analysis using the ChREAM choice data collected from 1100 respondents in Leeds, UK[13]. With Red water quality as the baseline, respondents were more likely to choose an option with higher water quality and less likely to choose an option with increased price. Because the ChREAM CE design conflated the ecological and recreational aspects of water quality, it was not able to estimate which aspect of water quality respondents preferred, or provide WTP estimates for ecological or recreational water quality improvements independently of one another.

3.2.1 The present research challenge

This research will add to the academic literature by disentangling and identifying respondents' preferences for the ecological and recreational aspects of river water quality, as the trade-offs between respondents' preferences in this area are not well understood. To achieve this, the research builds upon the foundations laid by ChREAM. A new water quality ladder which separates ecological and recreational water quality into separate attributes is developed and applied. To the extent that people care about water quality, it is hypothesised that there are differences between the two water quality attributes that affect respondents' values.

This research also identifies and samples different types of recreational users to ascertain how their attitudes towards recreational and ecological water quality may differ. The impact different attitudes have on WTP values for water quality

[13] The parameters reported in Table 25 were obtained from a cursory exploration of a subsample of the ChREAM CE data. Although they illustrate the likely relationships between water quality parameters, their coefficient values may deviate from the final output arising from a full analysis of ChREAM data.

attributes is assessed. Socio-economic variables are used to explore the hypotheses that people from different economic, educational and cultural backgrounds have heterogeneous environmental preferences which affect their WTP values. Previous research indicates that respondents' values for spatially fixed environmental goods fall as the distance to the location of the proposed improvement increases. This research incorporates spatially referenced data within a geographical information systems (GIS) framework (ESRI Inc., 2012) to help assess any differences in respondents' WTP which may be attributable to distance decay.

Respondents' perceptions can be at odds with objective reality. This doesn't mean that we should abandon any attempt to use public preferences within decision-making. Respondents are reasonably proficient at judging the more obvious aspects of water quality, e.g. turbidity, but are less able to assess the more obscure aspects of water quality such as microbial or nutrient loadings. Within this research uncomplicated and accessible survey instruments are developed, to provide respondents with information to assist their understanding of the water quality issues that are not immediately transparent to them, thus enabling them to make reasonable judgements of the same.

CL and LC methods are used to analyse the data collected within the choice experiment. The next section of the chapter provides an overview of the aims and objectives. This is followed by the methods used, including the case study area, the development of the survey instruments, the experimental design and the modelling strategies.

3.3 Aims and objectives

This chapter aims to further the knowledge on non-market valuation of river water quality by setting the question: "River water quality: who cares, how much and why?" To explore this question, there are a number of objectives to be met:

- Fully examine the nature of the research problem by unpicking the various components of the question, e.g. Which aspect of water quality? What type of respondent? What valuation methods can be used to elicit respondents' preferences (and ascribe monetary value to those preferences) for non-market goods? What primary data may best categorise respondents' motivations?
- Design survey instruments compatible with the above, which also perform the secondary function of collecting a robust dataset suitable for future quantitative and qualitative reanalyses, potentially integrated within other aspects of ChREAM econometric analysis.
- Develop a sampling strategy to reflect variation in preferences across key groups including non-visitors, non-specific visitors (e.g. those who use the areas around rivers for walking, picnicking, etc.) and specialist visitors (e.g. anglers, rowers, swimmers, etc.).Devise an efficient and parsimonious choice experiment.
- Conduct a pilot survey and assess the viability of the survey design and survey instruments.
- Conduct main survey interviews with identified respondents at a variety of locations and/or distances from the survey river in order to model spatial relationships.
- Assemble data within a geographical systems framework.
- Use CL and LC analytical techniques to identify and generate spatially explicit parsimonious models examining the relative importance to respondents of ecological or recreational water quality improvements on the River Yare in Norfolk, UK.
- In comparing preferences for ecological or recreational water quality improvements, identify socio-economic variables which significantly affect respondents' choices.

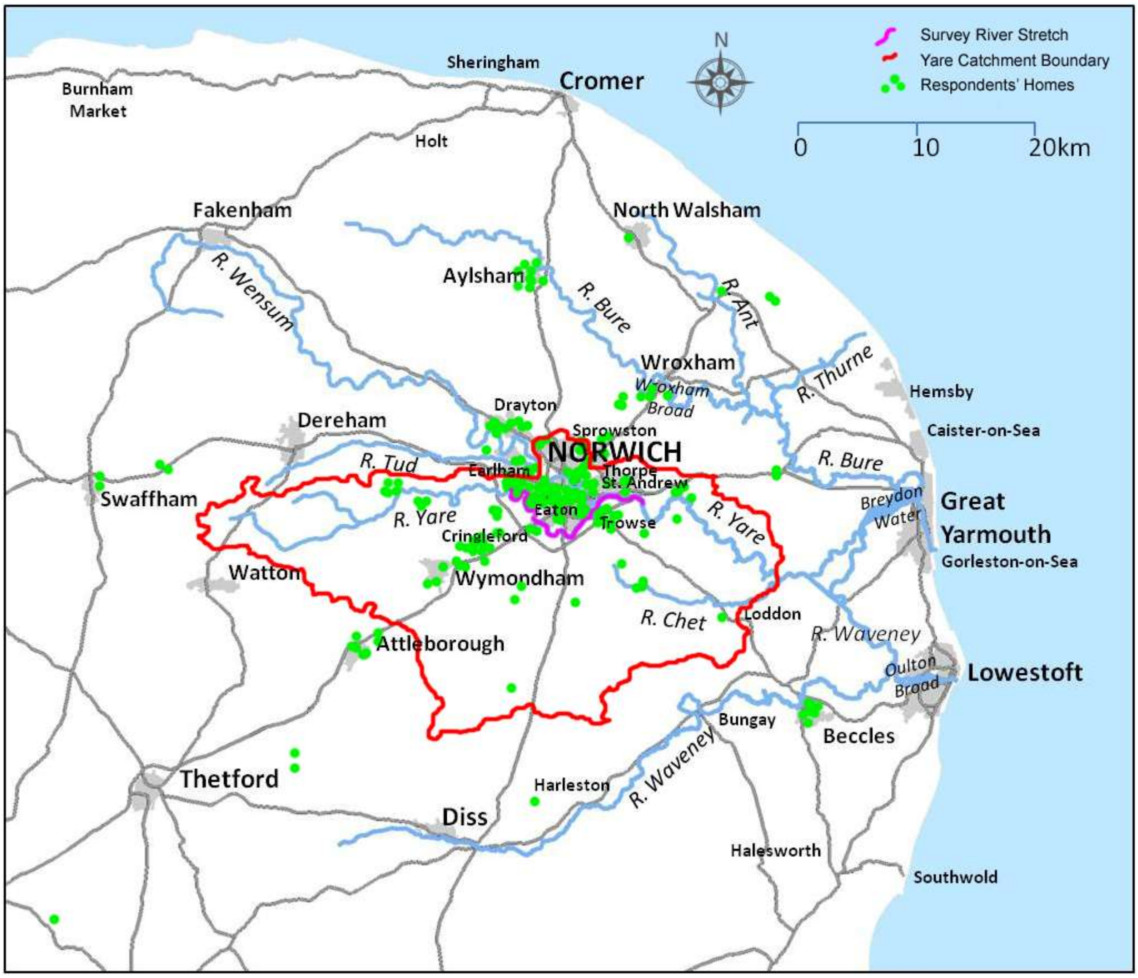
- Derive respondents' marginal willingness to pay (WTP) for ecological and recreational improvements for both CL and LC models.
- Refer to primary data concerning respondents' socio-economic characteristics to plausibly characterise respondents' preference motivations.

3.4 Methods

3.4.1 Case study area and catchment

The survey was conducted in and around Norwich and East Anglia, in the United Kingdom. Figure 20 shows the survey area, the locations of respondents' homes, the 20km survey stretch of the River Yare and the boundary of the Yare catchment. The River Yare was selected as the survey river, and as the case study for studying ecological and recreational values, for several reasons: the Yare catchment is predominantly rural, supporting agriculture and horticulture, but is prone to diffuse agricultural pollution from nutrients (primarily nitrates and phosphates). The catchment, which has a total area of 845km^2, was designated as a Catchment Sensitive Farming priority catchment in 2006 (Natural England, 2009b) and has been identified as a High Water Quality Priority Area for Catchment Sensitive Farming advice and Countryside Stewardship, as it fails to meet WFD targets (Natural England, 2016). Because of the difficulties of containing diffuse agricultural pollution on the Yare, the water quality is variable throughout the year. This is partly due to the farming cycle but other factors, such as rainfall, also contribute. The water quality is regularly monitored to ensure its compliance with water quality parameters and its suitability for recreational use (Lines, 2014). The Yare catchment also has difficulties meeting WFD quality targets due to point source pollution discharges from the water industry (Environment Agency, 2015a). The population of the Yare catchment is concentrated within the city of Norwich, a city of 210,000 inhabitants. Wastewater from the Norwich area is processed by the Whitlingham Sewage Treatment Works, near the village of Postwick, two miles to the south-east of the city, and discharged into the river downstream of the survey river stretch.

Figure 20: the survey area, survey river stretch and spatial distribution of respondents



The River Yare rises to the east of the village of Shipdham in Norfolk, near the town of Dereham, and flows eastwards towards Norwich. The survey stretch of the river, which measures 20km (12.4miles) in length (highlighted in purple on Figure 20), flows to the south of Earlham, through Eaton and Cringleford villages before skirting Norwich to the south, flowing through Trowse before ending at the south of the village of Thorpe St. Andrew. Beyond Postwick the Yare becomes tidal, draining into the area known as the Broads. The Broads have the equivalent status of a national park, due to their unique habitats and over 90 sites have been designated as Sites of Special Scientific Interest (SSSI). Economically, the Broads are a popular destination for tourists, particularly for boating, as there are over 280km of navigable waterways. The Yare's confluence with the North Sea is at Gorleston-on-Sea, near Great Yarmouth.

3.4.2 Survey instruments and choice experiment design

The goal of stated preference research design is to obtain the most accurate parameter estimates of the variables of interest (Bennett and Blamey, 2001). To achieve this goal, this research is guided by the research design framework proposed by Perman et al. (2003) and the excellent advice on CE design in Bateman et al. (2002) Hensher et al. (2005) and Hess and Daly (2014). To follow best practice (Johnston et al., 2017), care was taken to reflectively develop all aspects of the survey instruments and design the survey's implementation procedures to best maximize the validity and reliability of the resulting utility estimates. These issues are crucial in obtaining unbiased estimates which can add to the growing literature on non-market benefit valuation.

Firstly, hypothetical future water quality scenarios were developed based upon the requirement for the UK to achieve compliance with the WFD. Increased domestic water bills were proposed as an hypothetical means of payment. The survey instrument was then used with a sample of 200 respondents, to adequately populate the possible combinations of the survey (Grafton et al., 2004). The survey responses were analysed and used to generate compensative variation (CV)[14] WTP estimates using CL and LC modelling strategies.

However, before the discussion of the modelling strategies or results, this section discusses the following: the purpose of the survey and the questionnaire design; the CE design (including attribute selection, attribute levels and experimental design); the framing of the choice task; the graphical instruments (i.e. water quality ladders) used to help respondents compare the ecological and recreational attributes; the orthogonality inherent within the choice design (and choice sets providing examples of orthogonality); the rationale for respondent selection and the strategies used to recruit respondents.

3.4.3 Questionnaire design

The questionnaire was designed to disentangle and identify respondents' preferences for ecological and recreational river water quality. Within the survey

[14] CV is the adjustment in income that returns the consumer to the original utility after an economic change has occurred. In the case of a positive economic change, CV is often referred to as the maximum a consumer is willing to pay in order to have the economic change happen. When there is a negative economic change, CV is the minimum the consumer needs in order to accept the economic change (Bateman and Turner, 1993).

instrument design the goal was to use simple language and graphics to portray complex information relating to ecological and microbial pollution, and provide respondents with accurate, unbiased information on different water options, choice outcomes and associated costs (the survey questionnaire and showcards are provided in Appendix III). Terms that respondents could easily comprehend were used throughout, as the majority of respondents were expected to have little or no prior knowledge of river water quality issues.

These approaches were used to help minimise some of the limitations of stated preference experiments such as hypothetical bias, where respondents give misleading or poorly thought out answers because they believe the choice situation to be entirely hypothetical, and question framing bias, where the way in which the question is stated affects or influences the respondents' responses. For a full discussion on these limitations of stated preference design, please see Bennett and Blamey (2001). A CE design was used to minimise 'yea-saying', where respondents try to make themselves look good by claiming they would pay at a higher rate than they actually would, and prevent other strategic behaviours, where respondents deliberately misrepresent their preferences, in order to influence the experiment (Bennett and Blamey, 2001; Bergstrom et al., 1989; Hanley et al., 2001).

To reduce informational bias, each respondent was provided with almost exactly the same information from which to aid their choice decisions and that information was carefully considered to prevent excess information disclosure which can lead to inflated utility estimates (Samples et al., 1986). As far as possible the interviewer did not deviate in appearance or demeanour, to minimise interviewer variability and bias (Bailar et al., 1977).

In addition to the considerations described above, the survey instruments were also designed to collect data which may be compatible with, and used to integrate into, the ChREAM dataset, enabling a future amalgamation and reanalysis of the two datasets. For this purpose, the survey included questions to capture data which could be used within travel cost and contingent valuation analyses (Ferrini et al., 2014). Within this present analysis the information collected from the TC and CVM questions were used to generate socio-economic variables to produce comprehensive summary statistics and postestimation results.

3.4.4 CE design

Arguably, the most crucial part of this research was the design and execution of the choice experiment. To minimise the cognitive load on the respondents three attributes for the water quality scenarios were developed. The attributes and levels are shown in Table 26. Two attributes described the ecological and recreational aspects of river water quality and a third attribute, price, provided a metric from which respondents' utility and willingness to pay could be assessed in monetary terms. Ecological quality had four levels ranging from Blue (the highest) to Red (the lowest). Recreational quality had three levels; High, Medium and Low. Price was composed of eight levels ranging from £0 to £100.

Table 26: the attributes and levels used in the choice experiment design

3 Attributes														
Ecological Quality				Recreational quality			Price (£ per household, per year)							
4 levels				3 levels			8 levels							
Blue	Green	Yellow	Red	High	Medium	Low	0	10	20	30	40	60	75	100

The experimental design was devised by Professor Dan Rigby using NGene 1.1 (ChoiceMetrics PTY Ltd, 2012a). The combination of attributes and levels used on choice cards (shown in Table 26) was derived following the D-efficient design strategy[15]. This strategy ensures that the choice cards' combinations are orthogonal, balanced[16], and maximize the parameter precision of a CL model[17]. In total 48 combinations were produced, which were arranged into four blocks of twelve choices, with each block presented to 50 respondents, for a total of 200

[15] The D-optimality criterion seeks to maximize the determinant of the information matrix or to minimize its inverse, the determinant of the variance–covariance matrix of the parameter estimators. D-optimal design has dominated the design literature for CEs because it performs well in parameter estimation prediction and it is easy to obtain (Kessels et al., 2006). Further details of D-efficient designs can be found in Ferrini and Scarpa (2007) and in the Ngene v.1.1 manual (ChoiceMetrics PTY Ltd, 2012b).

[16] Orthogonality is a desirable property of experimental designs that requires strictly independent variation of levels across attributes, in which each attribute level appears an equal number of times in combination with all other attribute levels. Balance is a related property that requires each level within an attribute to appear an equal number of times (Johnson et al., 2013).

[17] D error is the determinant of the variance covariance matrix of the conditional model and is directly linked to parameters precision. D error was 0.306965. An alternative efficiency measure is the A error, which considers the trace of the variance-covariance method. A error in this design was 1.631721. The parameters precision is higher when these efficiency measures are closer to 0. The resulting design yields a higher efficiency level than those typically observed in the wider literature.

respondents. Each block was chosen at random and answered by 50 respondents.

3.4.5 Forced choice design with non-defined baseline water quality

Presenting a baseline water quality and providing the respondent with a ‘status quo’ choice alternative were deliberately avoided for experimental reasons.

A common characteristic of CEs is the inclusion of a constant level of environmental quality (the ‘status quo’) within the choice task and, within such studies, respondents could prefer the status quo to any proposed change. Defining the status quo in water quality studies is problematic as attributes vary according to river morphology, season, geographical location, etc. and encapsulating these variable components into a single fixed state may be overly restrictive. Furthermore, respondents are often heterogeneously and imperfectly informed and their perceptions of the ‘status quo’ may bear little or no resemblance to reality (Happs, 1986; Konishi and Coggins, 2008). This divergence can reduce the accuracy of welfare estimates: Poor et al. (2001) demonstrate the adverse impact of objective vs. perceived measures in valuing the water clarity of lakes in Maine. Given the potential discordance between objective status quo levels for water quality and respondents’ perceptions of the baseline water quality (discussed further below), and the variability in water quality along the survey river stretch, it was felt prudent to avoid rigidly defined ‘business as usual’ water quality levels.

More importantly, offering a status quo within the context of this study is not a real world option because changes to river water quality are necessary, there must be improvements. The WFD is compulsory and is forcing change via WFD programmes of measures. The absence of a status quo option reflects the reality of water management. This reality negates two of the main advantages of offering a status quo: that a status quo mimics a real-world market setting, wherein everyone is free not to buy and that it enables people to simply opt out (Krosnick et al., 2002).

Forced choice designs have been used in other water CE studies where the status quo is no longer an option (Hensher et al., 2005; Rigby et al., 2010; Train et al., 2005). Welfare analysis is still possible as long as the current levels of the attributes are included within the experimental design. The status (health) of the water environment of the River Yare, and its tributary rivers, was assessed by the Environment Agency in 2013 as being generally ‘moderate’ (Environment Agency, 2014). That assessment corresponds with the Green ecological and Medium recreational water quality attributes described below.

Furthermore, research by Krosnick et al. (2002) has shown that a status quo option can enable respondents to provide ill-considered answers to survey questions. They suggest that, as appealing as offering a status quo response may be, doing so may lead researchers to collect less valid and informative data than could be done by omitting it; offering status-quo options may discourage some respondents from doing the cognitive work necessary to report the true opinions they do have.

For the above reasons it was felt more appropriate, and more academically informative, to force respondents to choose between different levels of attributes, rather than allow them to opt into a status quo.

Where the quality of a good varies, as is the case for river water quality, there is often a difference from the objective baseline water quality and the respondent’s perceived baseline for water quality (Ferrini and Scarpa, 2007). Previous research (e.g. Georgiou et al., 1998; Sukharomana and Supalla, 1998) has demonstrated that a respondent’s perceptions of water quality are correlated with that respondent’s WTP for water quality improvements. Different respondents have different perceptions of water quality. High frequency visitors may be better informed, whereas infrequent visitors may have differing perceptions of water quality, and, for those infrequent visitors, factors such as anecdotal evidence may be more important determinants of respondents’ preferences than the actual water quality (Binkley and Hanemann, 1978; Happs, 1986).

For these reasons, an objective baseline for water quality was avoided so that a bias between the objective and perceived levels of water quality attributes was not introduced. Instead, within the survey, data on respondents’ perceptions of the ecological and recreational water quality at the survey river site, and rivers more generally, were collected (see Appendix III). If respondents queried the current water quality, the annual variability, discussed previously, was explained to them and they were asked to reflect on their own experiences and perceptions of the river. Differences in perceptions across respondents are highlighted in Table 32.

The approaches described above enable the analyst to have the freedom to calculate welfare in several ways and for various different scenarios: using the lowest water quality levels, by setting low ecological water quality and low recreational water quality as the baseline levels, to estimate preferences for sites where the current situation is objectively poor; the objective water quality, by setting Green ecological or Medium recreational water quality levels as the baselines, to correspond with EA water quality data (Environment Agency, 2014); or, as in Hynes et al. (2008), by using respondents' perceptions of water quality as these provide a natural baseline from which to estimate welfare. It is hypothesised that estimates of WTP using respondents' perceptions of water quality may differ from estimates using low water quality levels as the baseline.

3.4.6 Ex-ante design measures to reduce hypothetical bias

In stated preference valuation surveys, hypothetical bias can be defined as the difference between what a person indicates they would pay in the survey and what that person would actually pay. There are multiple ex-ante procedures that have been suggested to reduce hypothetical bias and enhance the validity of stated preference value estimates (Loomis, 2014). Cheap talk refers to the process of explaining hypothetical bias, and the tendency of respondents to inflate value estimates, prior to asking respondents valuation questions (Farrell and Gibbons, 1989; Lusk, 2003). However, the incentive properties of cheap talk are not clear and cheap talk does not always reduce value estimates (Murphy et al., 2005; Loomis 2014). Another ex-ante procedure involves the use of oath scripts, in which respondents are asked to sign a truth-telling oath. However, the ways in which oaths affect behavior are unclear. It has been suggested by Carlsson et al. (2013) that the primary function of the oath script is to increase respondents' commitment and attention.

Cheap talk and oath scripts were not used here. Johnston et al. (2017) believe that the most promising ex-ante approach to reducing hypothetical bias is a consequential design with a binding payment; Vossler et al. (2012) and Carson et al. (2014) find that consequential choice alternatives encourage truthful preference revelation. The scenario presented to respondents here is not hypothetical: real-world consequential changes to water quality are coming via the WFD. The payment vehicle was presented as a hypothetical binding increase to the respondents' annual domestic water bill, in response to the same. The payment vehicle corresponded with those used in previous studies on water quality improvements in public areas (Ferrini et al., 2014; Glenk et al., 2011). This approach was believed to be most appropriate as, in the UK, domestic water bills also include a sum towards improving wastewater services that, in turn, leads to improved river water quality. It was felt that respondents would view the policy scenario and the payment vehicle as consequential, thus aiding incentive compatibility (Herriges et al., 2010).

3.4.7 The framing of the choice task scenario

Respondents were introduced to the choice task with a scenario that new laws have been introduced to improve the quality of UK rivers and that any improvements to river water quality would incur costs. Respondents were reassured that the costs of improvements would be distributed equitably among all water users, including domestic, industrial and agricultural users. Respondents were told that an additional sum may be added to their annual domestic water bill as a contribution towards river water improvements.

Each respondent was asked to choose between two different hypothetical future water quality states, each containing an ecological, a recreational and a payment option. To prompt the respondent to consider the distance, use and cost issues implicit within their choice decision, they were asked to consider how close the river is in relation to their home, whether they would benefit from any improvements and they were reminded that any additional money they would be willing to spend on their water bill could not be spent on other goods or services.

3.4.8 Graphical instruments

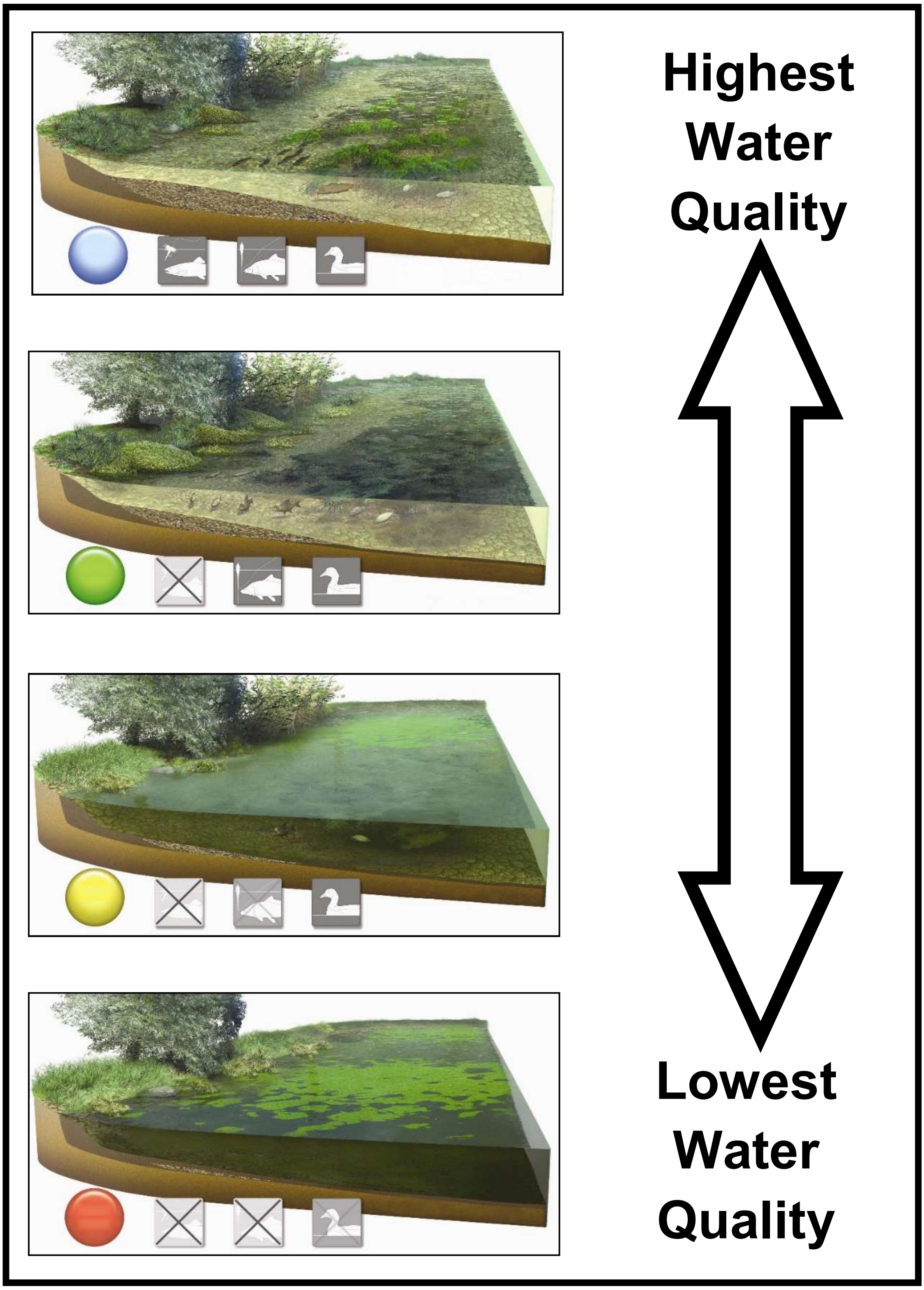
Technical descriptions were kept to a minimum. Instead, easily understandable graphics and images (water quality ladders) were used, where possible, to help respondents visualise and compare the ecological and recreational attributes. The four ecological water quality levels arranged in a water quality ladder, shown in Figure 21, were described to respondents. The picture marked with a blue circle depicts a river of the highest ecological quality. The first symbol shows that Blue rivers are suitable for pollution sensitive game fish such as salmon and trout. The second symbol shows that the river is suitable for coarse fish, such as carp and chub while the last symbol shows that the river is suitable for all bird species. The image shows a wide variety of plants in and around the river, which had very clear water.

The picture marked with a green circle represents a river of the second highest ecological quality which contains some ecological pollution. Within this scenario, pollution sensitive game fish cannot survive in the river (shown by a cross within the icon), but there is no reduction in coarse fish or birds. The variety of plants in and around the river is slightly lower but the water is still quite clear.

The image with the yellow circle shows higher levels of ecological pollution, with no game fish and significantly fewer coarse fish. The variety of plants is lower and algae has substantially reduced the water clarity. There are still a number of birds.

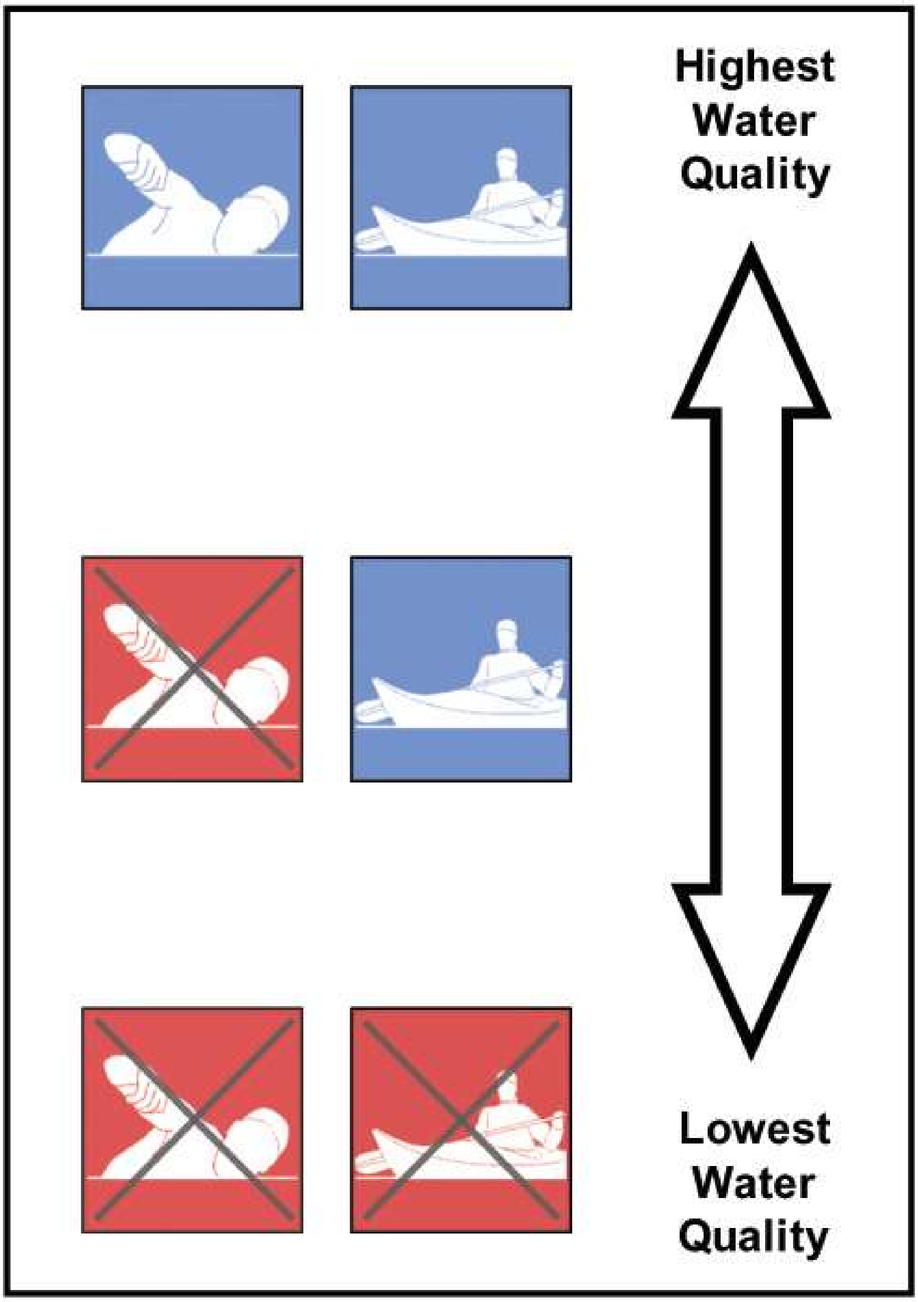
The final image, marked with a red circle, shows a river subject to the highest level of ecological pollution. It has no fish, few birds or water plants. There are large algal mats on the surface of the water and the water is very cloudy.

Figure 21: graphic used to depict the ecological water quality ladder



Given the paucity of empirical epidemiological data, the survey did not include specific numerical illness thresholds. Instead, the types of illness and the risk of illness due to biological pollution was stated using simple language. Respondents were told that the more contact a person has with biologically polluted water, the more likely it would be that they would get ill. For example, that someone swimming in the water has a higher risk of illness than a person in a canoe that only gets splashed with the water and that a person on the river bank, who has no contact with the water, has no increased risk of getting ill. Three attribute levels were generated using simple icons to define the three water quality levels in the recreational water quality ladder. These are shown in Figure 22.

Figure 22: graphic used to depict the recreational water quality ladder



The blue icons describe river water of a quality sufficient to be suitable for a given activity. Respondents were told that a river of the highest quality was suitable for swimming and boating and the risk of illness was low. The second river has higher levels of microbiological pollution affecting recreational quality. This type of river has a higher risk of illness, and, although it is still suitable for boating, the red icon indicates that it is no longer suitable for swimming. The river with the lowest recreational quality has the highest risk of illness and isn't suitable for swimming or boating.

3.4.9 Orthogonality across choice attributes

The survey script uses very simple language to portray the concept that living (microbiological) and non-living (e.g. chemical) pollutants have different and variable impacts on humans and the ecology of the environment. A table of pollution types and vectors (Showcard 5, Appendix III), simple descriptions of the impacts of faecal pollution on human health, the impacts of chemical pollution on ecological health (survey script, Appendix III) and graphical depictions of ecological and recreational water quality (e.g. the water quality ladders on showcards 6a and 6b, Appendix III), ensured that respondents understood that chemical and microbiological pollution can have independent sources and independent consequences.

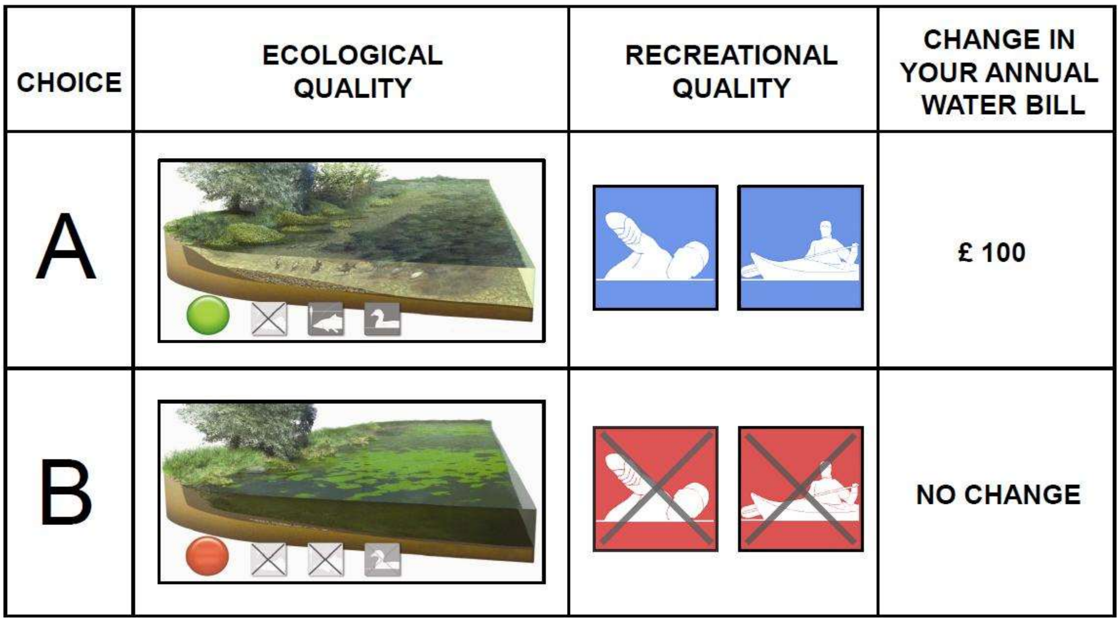
Microbiological pollution is not present in diffuse agricultural pollution from chemical fertilizers (e.g. nitrates or phosphates) or from phosphates from detergents. Similarly, aquatic flora and fauna are largely unaffected by faecal microorganisms derived from humans and livestock (e.g. shellfish are, themselves, largely unaffected by faecal organisms. If bacteria are present in the water they will build up in the shellfish tissue over time: bacteria can be 100 times more concentrated within the shellfish tissue than in the surrounding water (State of Maine Department of Marine Resources, 2016). The Shellfish Water Directive is not designed to protect the shellfish, per se, but rather to prevent the shellfish from causing ill-health to humans once harvested (EU, 2006b).

It is acknowledged that some pollutants can have the same source but duel impacts, e.g. faecal waste can contain both microbiological and ecological pollutants (e.g. nitrogenous matter and phosphates) and both can be emitted from the same point source at wastewater treatment works. Indeed, phosphates emitted from wastewater treatment works contribute to WFD non-compliance within the Yare Catchment (Environment Agency, 2015b).

It is important to note that the research in this chapter primarily concerns itself with the disaggregation of the non-market benefits arising from recreational (with the emphasis on the reduced risk of ill-health from microbiological contamination) and ecological improvements, and not the identification of the sources of pollutants. To ensure a parsimonious experimental design and reduce the possibility of cognitive difficulty, there were also other pollutants and facets of water quality that were not included within the analysis (e.g. heavy metals or other industrial contaminants, pesticides, the impact of climate change, etc.).

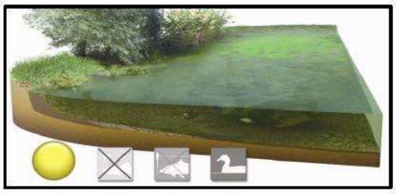
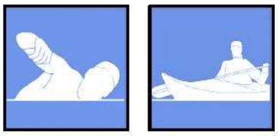
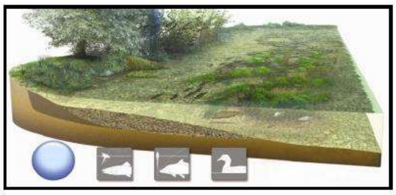
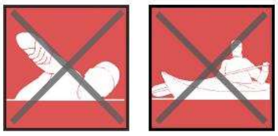
To illustrate the choice task, two of the choice sets used during the survey will now be examined.

Figure 23: choice set 11 from block 3



Choice set 11 from block 3, shown in Figure 23, provides the respondent with a choice between a future alternative, option A, which has High levels of ecological and recreational water quality at an additional cost to their annual water bill of £100, or option B, which has the lowest levels of water quality with no change to their annual water bill. Within this choice set, respondents are effectively asked to choose whether they are willing to pay for water quality improvements.

Figure 24: choice set 5 from block 3

CHOICE	ECOLOGICAL QUALITY	RECREATIONAL QUALITY	CHANGE IN YOUR ANNUAL WATER BILL
A			NO CHANGE
B			£ 10

Choice set 5 from block 3, shown in Figure 24, demonstrates the orthogonality inherent within the CE design. This choice set presents the respondent with future water quality scenarios which contain little correlation between ecological and recreational water quality attributes. Within option A the respondent can choose the second lowest ecological quality with the highest recreational quality, with no change to their annual water bill. Option B enables the respondent to choose the highest ecological quality with the lowest recreational quality with an additional £10 on their water bill. Given that the cost attribute of this choice set has only a small difference between options, the respondent is effectively faced with a choice between high ecological quality or high recreational quality.

3.4.10 Respondent selection

The main aim of this research is to further the knowledge on non-market valuation of river water by disaggregating the values people derive from ecological and microbial aspects of river water quality. Given the link between microbial quality and recreational river use, respondents were carefully targeted to investigate how the values for these distinct attributes of river water quality differ over people who (i) engage with the river in different ways (e.g. rowers, swimmers, anglers) and (ii) who live at different distances from the river. There are a variety of sampling strategies available to the researcher. An online survey was not considered, as that sampling method may have created problems of self-selection bias, preventing sampling objectives from being achieved. It was felt that face-to-face interviews, conducted by the PhD candidate, would be most appropriate to meet sampling objectives and also provide the candidate with a more complete understanding of the survey data (e.g. to better identify protesting respondents or identify (and assist) respondents who had cognitive difficulty with the questionnaire or survey tasks). In addition, face-to-face interviews can be flexible, enabling respondents to fully explore their responses to open-ended questions (e.g. Questions 22e, 23g in Appendix III) or provide other insights to the researcher. A single interviewer carried out all surveys to minimize informational bias caused by using different interviewers (Bailar et al., 1977). Table 27 shows the breakdown of the different types of respondents interviewed, before a brief discussion of why, how and where these respondents were recruited.

Table 27: respondent type, recruitment method and respondent numbers

Respondent type	Recruitment Method	Number completing CE
General public	door to door	139
Rowers	by appointment	10
Swimmers	by appointment	5
Experts	by appointment	7
General public	at Whitlingham visitor centre	39
Total number of respondents		**200**

139 respondents were drawn from the general public, with care taken to sample from a range of different socio-economic backgrounds. The locations of respondents' homes, within Norwich and East Anglia, are shown on Figure 20. It was particularly desirable to interview 'hard to reach' respondents (e.g. members of the general public who do not use or visit rivers for recreation, or the elderly and people on low incomes, who may not use or have access to the internet) and to interview such respondents at a range of distances from the survey river. It was felt that the most appropriate method of doing this was to use census data to identify the areas in which respondents from the general public (with a range of socio-economic characteristics) may be found, and then to canvas for those respondents door-to-door. Interviews were conducted morning, afternoon and early evening during weekdays and weekends, to collect data from respondents with different temporal availability and socio-economic circumstances. In each sampling location random survey routes were taken. This strategy was devised

to explore the hypotheses that people from different economic, educational and cultural backgrounds have differing viewpoints and preferences on how our rivers should be managed.

Respondents were interviewed at a range of distances, 0.1-79.4km, from the survey river stretch to test the hypothesis that respondents who lived further from the river would be less willing to pay for water quality improvements. It was expected that the door to door respondents would primarily be a mix of casual or non-visitors.

Recreational river users, with high levels of contact with the river water, were recruited from local recreation clubs. Rowers from Norwich Rowing Club and University of East Anglia Rowing Club and swimmers from Tri-Anglia Triathlon club were interviewed by appointment at times and locations convenient to them. Despite invitations being placed in their respective club newsletters, these respondents proved to be somewhat reticent about volunteering for the survey.

Seven experts[18], working locally in specialisms relevant to the different areas of interest[19], were identified and kindly agreed to participate in the survey and to take part in semi-structured interviews to express their views on river water quality management. Importantly, this method of interviewing can enable the respondent to focus on topics of individual importance, rather than the researcher introduce (and promote discussion of) their own biases (Christie et al., 2008). The format also enables the researcher to form additional questions that occur to them during each individual interview, based entirely on each particular respondent's engagement with aspects of the research topic (Watts and Stenner, 2005). Such interviews enabled the researcher to gain unique insights in the localised issues relating to river water quality and river recreation. Experts were interviewed by appointment either at their workplace or at a neutral location. Expert participants

---

[18] i.e. those, defined by Thompson (1966), who have expert knowledge due to their formal training or engagement, or who have a position of authority within society.

[19] These included people of theoretical interest from both public (e.g. local and national government) and private organisations, professionally engaged with maintaining and improving river water quality or directly implementing programmes of measures to improve local WFD compliance. Chairmen from local organisations representing the key recreational communities (i.e. angling, rowing and river swimming) were also interviewed.

were expected to be very familiar with the river stretch with a tendency, due to the nature of their work, to be highly knowledgeable high frequency visitors.

Respondents from the general public, who do not use the survey river for rowing or swimming but instead enjoy non-contact activities such as walking or picnicking, were canvassed at the Whitlingham Country Park Visitors Centre, adjacent to the survey river. Whitlingham Country Park, to the south of Norwich, covers 35 hectares and is managed in a partnership between Whitlingham Charitable Trust and the Broads Authority (Bentley, 2014). Whitlingham offers recreation opportunities ranging from woodland walks to sailing, swimming and kayaking. The site also has a purpose built education centre. Respondents were selected at random and interviewed on-site during the park's opening hours on weekdays and weekends. Respondents interviewed at Whitlingham were, by the very location of the interview site, expected to visit the river location for, at the least, non-water-contact riverside recreation activities.

3.4.11 Modelling strategies

The data obtained from the CE was analysed using CL and LC modelling approaches, both of which are now discussed.

The CE approach relies on the assumptions that respondents have preferences that conform to standard expectations of rational behaviour, that they prefer options that maximise their utility and that their preferences are complete and transitive across the choice sets offered to them (Gale and Mas-Colell, 1978).

A set of discrete choice models were estimated using CL models, as a starting point to ensure that the data were clean and that reasonable results were obtained (Boxall et al., 1996; Brey et al., 2007; Hausman et al., 1995).

Within the CL framework, the indirect utility function for each respondent i (U) is comprised of two parts: an objective or deterministic element (V), which is usually specified as a linear index of the attributes (X) of the j different alternatives contained in the choice set, and a stochastic element (e), which represents the random error or unobservable influences on the respondent's choice, shown in Equation 4 (Louviere et al., 2000).

$$Uij = Vij(Xij) + eij = bXij + eij$$

Equation 4: the indirect utility function for each respondent within the CL model

It is important to note that within a single choice occasion ($T=1$), the respondent's utility is a function of alternative characteristics and individual characteristics. Within this dataset there are repeated observations per respondent. For simplicity, within the CL modelling T has been omitted. Subsequent LC modelling takes the panel structure of the data into account.

The probability that respondent i prefers option g to the alternative option h, is expressed as the probability that the utility obtained from option g exceeds that of option h, as in Equation 5.

$$P[(Uig > Uih) \forall h \neq g] = P[(Vig - Vih) > (eih - eig)]$$

Equation 5: the probability of the utility of option g exceeding option h

Error terms (e_{ij}) are assumed to be independently and identically distributed with a Weibull distribution (Nakagawa and Osaki, 1975), Equation 6.

$$P(e_{ij} \leq t) = F(t) = \exp(-\exp(-t))$$

Equation 6: the Weibull distribution of the error terms

The distribution of the error term in equation 6 implies that the probability of the alternative g being preferred is expressed in terms of the conditional logistic distribution (McFadden, 1973), shown in Equation 7.

$$P(Uig > Uih, \forall h \neq g) = \frac{\exp(\mu Vig)}{\sum_j \exp(\mu Vij)}$$

Equation 7: McFadden's conditional logistic distribution

μ is a scale parameter, inversely proportional to the standard distribution of the error distribution. As this parameter could not be separately identified it was assumed to be one. The model was then estimated using maximum likelihood procedures, with the log-likelihood function shown in Equation 8. The indicator variable, y_{ij} ,was 1 if respondent i chose option j and 0 otherwise (Hanley et al., 2001).

$$\log L = \sum_{i=1}^{N}\sum_{j=1}^{J} yij \log\left[\frac{\exp(Vij)}{\sum_{j=1}^{J}\exp(Vij)}\right]$$

Equation 8: the log likelihood function of the CL model

Socio-economic variables were included with choice set attributes in the X terms in equation 4, but as these are constant across choice occasions for each individual i, they were treated as interaction terms and interacted with the choice specific attributes.

Within the linear utility model, marginal utility estimates were converted into WTP estimates for changes in attribute levels and welfare estimates were obtained for combinations of attribute changes. After parameter estimates were obtained, a WTP compensating variation welfare measure, conforming to demand theory, was derived for each attribute using the formula shown in Equation 9 (Hanemann, 1984). Where V^0 represents the utility of the initial state and V^1 represents the utility of the alternative state, the coefficient of the price attribute, b_y, gives the marginal utility of price.

$$WTP_i = b_y^{-1}\ln\left\{\frac{\sum_i \exp(V_i^1)}{\sum_i \exp(V_i^0)}\right\}$$

Equation 9: the WTP compensating variation measure for each attribute

The above formula can be simplified to the ratio of coefficients given in Equation 10 where b_c is the coefficient on any of the attributes, where the marginal WTP estimates are the negative of the ratio between the mean coefficients for each attribute and the mean coefficient of the payment attribute.

$$WTP_{ic} = \frac{-b_c}{b_y}$$

Equation 10: WTP derived from the ratio of coefficients

Unfortunately, the CL model has several drawbacks (Luce, 1959). Repeated observations by the same respondent cannot be accommodated by the model, heterogeneity in preference cannot be properly addressed and correlation among alternatives cannot be estimated. Non-random effects were examined using the Lagrange Multiplier (LM) test (Breusch and Pagan, 1980; McFadden and Train, 2000). The LM test uses artificial variables to verify heterogeneity in preferences and verify whether the distributional assumption on the error components is supported by data. Dropping the t-index for simplicity, the artificial variables can be obtained as shown in Equation 11.

$$z_{in} = (x_{in} - x_{Cn})^2, \text{with } x_{Cn} = \sum\nolimits_j x_{jn}\, P_{jn}$$

Equation 11: constructed artificial variables within the Lagrange Multiplier (LM) test

where P_{jn} is the CL choice probability. The CL model in Equation 7 is re-estimated including the artificial variables and the null hypothesis of no random coefficients on attributes x is rejected if the coefficients for the artificial variables, tested using a Wald or Likelihood Ratio test, are significantly different from zero. (Brownstone, 2001; Hensher and Greene, 2003). When the test fails to reject the null, the implication is that the CL assumption on the error term is inappropriate and other assumptions must be tested. An alternative model specification is LC, which overcomes CL limitations in addressing preference heterogeneity (Morey et al., 2006) and repeated choices (Kemperman and Timmermans, 2006).

Formally a LC model uses a probabilistic class allocation model and a CL model for the alternatives choice. Each respondent i belongs to class s with probability π_{is}, with $\pi_{is} \in (0,1)$ and $\sum_{s=1}^{S} \pi_{is} = 1$. The probability of respondent i belonging to class s is shown in equation 12,

$$\pi_{is} = \frac{\exp(\delta_s + \gamma_s z_i)}{\sum_{l=1}^{S} \exp(\delta_l + \gamma_l z_i)}$$

Equation 12: the LC probabilistic class allocation model

where δ_s is a class specific constant, z_i is a vector of individual socio-economic characteristics and γ_s is a vector of the parameters to be estimated, which determines the probability of respondents belonging in class s. This term models observable respondent characteristics which may help to explain preference heterogeneity.

Conditional on the probability of being in class s the probability of choosing option j among the J alternatives is equivalent to Equation 13, where b is the vector of parameters and x represents the attributes:

$$P(y_{it} = j | b_i = b) = \frac{\exp(b\, x_{ijt})}{\sum_{q=1}^{J} \exp(b\, x_{iqt})}$$

Equation 13: the probability of choosing option j among the J alternatives

The unconditional probability of choosing option j for respondent i for choice situation $t=1$ is shown in Equation 14.

$$P_i\ (j | b_{1,} \ldots, b_s) = \sum_{s=1}^{S} \pi_{is} P(y_i = j | b_i = b_s)$$

Equation 14: the unconditional probability of choosing option j for respondent i for choice situation $t=1$

Unlike continuous mixture MIXL models, LC does not require simulation techniques to estimate the model parameters. In common with CL, maximum likelihood procedures can be used. In cases of multiple choices ($t > 1$) the log-likelihood function is shown in Equation 15.

$$ln\, L = \sum_{i=1}^{N} lnP_i\, \left(j_{1,} \dots , j_T \middle| b_{1,} \dots , b_s\right) = \sum_{i=1}^{N} ln\left[\sum_{s=1}^{S} \pi_{is} \left(\prod_{t=1}^{T} P_{it|s}\right)\right]$$

Equation 15: the log-likelihood function within the LC model, in cases of multiple choices

Once the model parameters are obtained the welfare estimates can be obtained as combinations of parameters. Following the welfare theory (Hanemann, 1984), the ratio of marginal utility of each attribute (k) and that of price (p) provides the willingness to pay measure of Equation 16.

$$WTP_s = -\frac{b_{k|s}}{b_{p|s}}$$

Equation 16: the WTP measure of the LC model

The Krinsky and Robb (1986) method can be implemented to provide the confidence interval of the WTP measures and respondents' individual class membership probabilities can be calculated using the method described in Morey and Thacher (2012). This method enables each respondent to be allocated to their most likely class using that individual's conditional class-membership probabilities, which are based on their responses to the choice questions. From that data, postestimation results can be defined, informed by class members' socio-economic characteristics.

The choice setting deliberately avoided defining the current water quality level (status quo) and the survey collected information on perceived water quality. WTP was estimated in two ways: either Low ecological or recreational quality as initial water quality states, or, as in Hynes et al. (2008), by using respondents' perceptions of water quality. At the individual level, the perceived water quality was considered and, where that respondent's perception was equal to (or higher than) the proposed improvement, WTP was set to zero. The hypothesis of compensation was ignored (e.g. should we wish to improve the river from Low to Medium water quality, if the respondent's perception was already Medium (or High), the improvement did not produce any benefit). The marginal WTP was registered for the other cases.

3.4.12 Pre-survey testing

In order to assess the design of the survey instruments, confirm appropriate levels for the water quality attributes and ensure reasonable future water quality scenarios, a focus group and a pilot survey were performed.

The focus group was used to help select reasonable levels for choice attribute and assess the overall clarity of the survey instruments. Minor changes to the survey were made where necessary, primarily to improve the clarity of the survey instruments. For example, following the focus group it was clear that some participants had difficulty understanding the verbal description of the different types of pollutants (contained within the ‘Pollution Information’ section of the survey script, see Appendix III). To remedy this, the script was altered slightly and a table (see Showcard 5, Appendix III) was created so that respondents had a graphical aid to assist their understanding of the different pollution types and sources.

The survey design was further tested in a pilot survey, in which 20 respondents (10 male, 10 female, with a mean age of 50) participated. For sample representativeness, these participants were recruited from the primary target population (i.e. the general public) (Johnston et al., 2017). Their understanding of the subject and their responses to the questionnaire were assessed by the interviewer for signs of cognitive burden, fatigue or misinterpretation of questions, as these have been identified as potential issues in obtaining reliable CE data (Mazzotta and Opaluch, 1995; Swait and Adamowicz, 1996). The survey also included a question to enable respondents to report the level of cognitive difficulty they experienced while undertaking the survey tasks (see Question 21 of Appendix III). The pilot survey’s respondents reported that they were not overly taxed by the survey’s complexity: 8 respondents thought the choice task was fairly easy and 7 found it to be easy.

A simple CL model on the pilot data was estimated using Stata 13.1 (StataCorp L. P., 2013). This model (presented in Appendix IV) was used to assess the general correctness of the pilot respondents’ responses and to highlight any need for adjustments to the experimental design before undertaking the main survey.

A priori expectations were that the results should be broadly similar to the ChREAM survey results, reported in Table 25.

The CL analysis of the pilot data is now briefly discussed. Three coefficients, measuring respondents preferences for ecological quality (EQ), recreational quality (RQ) and price were generated. Coefficients for EQ and RQ were both positive, meaning that respondents were more likely to choose options containing higher levels of EQ and RQ. The strength of the coefficients relative to one another (0.725 against 0.646 respectively) suggested that respondents preferred higher ecological quality. Both water quality coefficients were highly significant ($p=0.000$). The coefficient for price was negative, meaning that respondents were less likely to choose an option containing increased price. The coefficient for price was statistically insignificant ($p>0.102$). Given the size limitations and exploratory nature of a pilot study, conventional 95% confidence intervals may be unrealistically stringent, i.e., 90% confidence intervals may be more appropriate (Hertzog, 2008).As the significance of price was very close to the 90% confidence interval, and as the confidence intervals in the dataset were wide (typical of a small dataset), it was felt that with a larger dataset the price coefficient would naturally become significant as the confidence interval became narrower.

It is important to note that due to the small sample used within the pilot study, there is the possibility of making inaccurate predictions or assumptions on the basis of pilot data; although pilot study findings may offer some indication of the likely outcome of the main survey, they cannot guarantee this because, being based on small numbers, they do not have an accurate statistical foundation (Johnston et al., 2017; van Teijlingen and Hundley, 2001). Other than the insignificance of the coefficient for price, there were no other anomalies apparent within the pilot results. The pilot experiment produced results consistent with a priori expectations.

The analysis of the accessibility of the survey and the complexity of the choice task, along with the results of the CL analysis, suggested that the survey instruments and CE design were fit for purpose within the main survey.

To avoid problems arising from selection bias or data contamination (Lancaster et al., 2004), respondents who participated in the pilot study were excluded from participating in the main survey.

3.5 Results

3.5.1 Sample Representativeness

The representativeness of the sample (in terms of income, age and gender) was compared against census data. The mean income of the sample was £28,400. This is slightly lower than the mean income (£30,000) for Norfolk inhabitants (Norfolk County Council, 2015a). The mean age of respondents was 51. This corresponds closely with the mean age (51.3) for Norfolk residents aged over 18[20] (Norfolk County Council, 2015b). The sample contains proportionally less males: 44% of the sample were male, compared with 49% for the region (Office for National Statistics, 2016). 8% of the sample were anglers, which corresponds closely with official estimates (9%) of the proportion of people who go freshwater fishing (Environment Agency, 2010b). The sample used here cannot be said to be entirely representative of the wider population. In terms of the purpose of the survey[21], the sampling strategy necessarily oversampled river visitors (e.g. respondents interviewed at Whitlingham Visitor Centre), and also oversampled rowers and swimmers (7.5%), who account for less than 1% of the wider population (British Rowing, 2017). These types of respondents were deliberately oversampled to more accurately capture their preferences (i.e. generate data to adequately populate choice alternatives) and to test the theory that such respondents may hold subjective preferences distinct from those held by the public more generally.

3.5.2 Fatigue and learning effects

Criticisms against the repeated choice format employed within choice experiments are that respondents can become less engaged, or fatigued, by the number of choice tasks (Bradley and Daly, 1994) or that learning effects lead to changes in the respondent’s overall preference structure (Braga and Starmer, 2005). The effects of learning or fatigue can become pronounced, particularly where many repeated choices are present. This CE employed 12 repeated choices. To test for learning or fatigue effects the results of a CL model on the

[20] Only people aged over 18 and who have responsibility for paying a domestic water bill were interviewed.

[21] e.g. to disaggregate the values different types of respondents, including a mix of river recreational users (such as rowers, swimmers, anglers), river visitors (who use the river for bankside activities such as walking, picnicking, wildlife watching), experts, and respondents from the general public, hold for the ecological and recreational aspects of water quality.

first 6 choices were compared against the same model on the last 6 choices, using a Chow test which tests whether the true coefficients in two linear regressions on different data sets are equal. Fatigue or learning effects are said to be insignificant if the parameters are statistically similar (Hess et al., 2012). A likelihood-ratio test, Prob. > chi2 = 0.11, reveals that at the 0.05 significance level the hypothesis that the parameters differ can be rejected. The Chow test suggests that there appear to be no fundamental issues in the data attributable to learning or fatigue effects.

3.5.3 Dominance and consistency checks

Each of the 4 choice blocks contained a dominated choice set (e.g. one option clearly dominated the other on all dimensions by containing superior levels for ecological and recreational water quality and lower price). Dominated options were deliberately included within the choice sets to enable consistency checks to assess rationality, engagement and protest behaviour (Burge and Rohr, 2004).

Table 28: number of respondents making rational responses to dominated choice options

Choice block	Number of respondents	Rational responses
1	50	50
2	50	48
3	50	50
4	50	49

Of the 200 opportunities shown on Table 28, the dominant option was chosen on 197 occasions. In interviews conducted after the completion of the survey, the three respondents who chose the dominated option gave reasonable explanations for their choice behaviour. All three essentially refused to choose recreational water quality suitable for swimming. One respondent didn't want people swimming in rivers because they felt that people should use swimming pools if they want to swim. The second respondent felt that swimming in rivers is too dangerous (e.g. risk of accident or drowning). The third respondent was a volunteer conservation ranger at Whitlingham and expressed a preference to prevent swimming where possible. All three individuals exhibited good understanding of the purpose of the survey and gave the choice questions careful consideration. They were all retained within the sample to prevent selection bias or to prevent the reduction of the statistical efficiency and power of the estimated choice models (Lancsar and Louviere, 2006).

Non-trading choice behaviour occurs, especially in the case of labelled choice experiments, when a respondent always chooses the same alternative across choice sets (Hess et al., 2010). To detect for any place bias, the result of a Mann Whitney test (p>0.049) confirmed that the hypothesis of no significant difference in the choice alternative chosen by respondents cannot be rejected when the confidence level is 1% or lower than 5%.

3.5.4 Summary statistics of attribute level selection

The following three tables report the number (and proportion) of times each level of the three choice attributes were chosen.

Table 29: distribution of ecological water quality level selection

Ecological water quality level	Red	Yellow	Green	Blue	Total
Rejected	985 (82)	563 (47)	515 (43)	337 (28)	2400 (50)
Accepted	215 (18)	637 (53)	685 (57)	863 (72)	2400 (50)
Total for each level	1200 (100)	1200 (100)	1200 (100)	1200 (100)	4800 (100)

Proportion (%) in brackets

Within the ecological water quality attribute we see that respondents preferred to choose choice alternatives that contained higher levels of ecological water quality. For example, choice alternatives containing Blue ecological quality were selected 72% of the time when offered as an option within a choice set. Choice alternatives containing Red ecological quality were rejected on 82% of choice occasions.

Table 30: distribution of recreational water quality level selection

Recreational quality level	Low	Medium	High	Total
Rejected	1047 (65)	749 (47)	604 (38)	2400 (50)
Accepted	553 (35)	851 (53)	996 (62)	2400 (50)
Total for each level	1600 (100)	1600 (100)	1600 (100)	4800 (100)

Proportion (%) in brackets

We see a similar distribution of preferences for selecting the recreational water quality attribute. Choice alternatives containing High recreational quality were chosen 62% of the time when offered as an option within a choice set, whereas alternatives containing Low recreational quality were only selected 35% of the time.

Table 31: distribution of price level selection

Price (£) level	0	10	20	30	40	60	75	100	Total
Rejected	216 (36)	310 (52)	282 (47)	314 (52)	292 (49)	299 (50)	294 (49)	393 (66)	2400 (50)
Accepted	384 (64)	290 (48)	318 (53)	286 (48)	308 (51)	301 (50)	306 (51)	207 (35)	2400 (50)
Total for each level	600 (100)	600 (100)	600 (100)	600 (100)	600 (100)	600 (100)	600 (100)	600 (100)	4800 (100)

Proportion (%) in brackets

Although the distributions for ecological and recreational water quality levels show consistent increases in the number of times progressively improved levels of each attribute are selected, the pattern of price level selection differs. Although, as we would expect, the lowest price level (£0) is consistently preferred and the upper level (£100) consistently rejected, we see a different pattern within the remaining range of price levels, where the different levels have been selected in roughly equal amounts. At face value it would appear that respondents took the opportunity to select improved levels of water quality attributes and felt the mid-range of price levels contained within the bundle of attributes to be acceptable in obtaining improved water quality. This issue is discussed further within the limitations section of this chapter.

3.5.5 Sample summary statistics

Two hundred respondents were interviewed during the main survey using the sampling scheme described above. Of the 200 respondents there were 16 anglers across the five main respondent categories. Descriptive statistics of the main socio economic characteristics of the sub-groups within the whole sample are reported in Table 32 on the following page.

Table 32: summary statistics of the main survey data

	Respondent type					
	Public recruited door to door	Rowers	Swimmers	Experts/ Management Community	Public recruited at Whitlingham Park	Total
Number of respondents	139	10	5	7	39	200
Class share (%)	69	5	3	4	19	100
Mean age	51.7	31	49.4	48.6	53.5	51
Gender (% male)	45.30	20	60	85.70	35.90	44
Employment, income, education and environmental affiliation						
Employed (%)	40.30	50	80	85.70	46.10	44.5
Environmental occupation (%)	2.90	40	40	100	0	8.50
Mean income (£1000's)	24.1	33	57	52.7	35	28.4
Degree level education or higher (%)	42	60	100	57	46	46
Mean number of environmental memberships	0.35	1.4	1.6	0.86	0.59	0.51
Anglers (%)	8	10	0	42	3	8
Distance and trip information						
Mean distance the respondent lives from the Yare (kilometres)	7.75 (8.61)	4.86 (3.77)	1.83 (1.41)	3.57 (3.02)	11.24 (16.4)	7.99
% respondents who visited the Yare in the last year	39.7	100	100	100	94.9	57%
Mean number of Yare trips in the last year	9.8 (30.9)	171.6 (57.3)	94.4 (68.2)	49.4 (39.1)	33.9 (69.0)	26.1 (57.6)
Mean number of different activities undertaken at the Yare in the last year	0.9 (1.3)	2.3 (0.8)	2.8 (0.7)	2.6 (1.9)	2.4 (1.2)	1.4 (1.4)
Mean number of Yare trips in next year if water quality improvements made	20.7 (48.7)	171.6 (57.3)	130 (70.0)	92.7 (115.2)	43.2 (86.3)	37.9 (72.6)
Respondents' perceptions of current water quality (% of respondents)						
EQ blue	21	20	20	42.90	16	20.9
EQ green	53	50	80	57.10	57	55.6
EQ yellow	21	30	0	0	21.60	20
EQ red	4	0	0	0	5.40	3.5
RQ high	41	30	20	57	13	31
RQ med	55	70	80	43	81	65
RQ low	4	0	0	0	5	3
Importance of issues when making choice decisions (% of respondents)						
Bill is important	62.5	40	40	57.1	69.2	62
Distance is important	48	40	20	86	26	44
Ecological quality is important	94.9	100	100	100	94.9	95.5
Recreational quality is important	53.9	90	100	85.7	51.3	57.5

Standard deviation in parenthesis

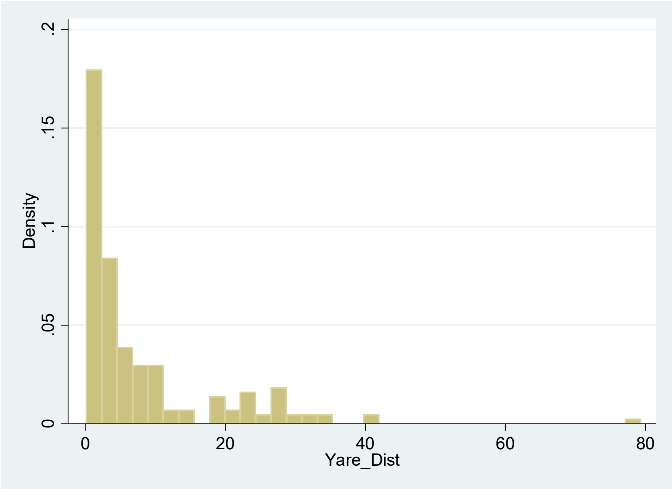
The mean age of all respondents is 51. There is little variation from this mean, except within the rowers category, which has a mean age of 31. This average is lower because approximately half of the rowers were students sampled from the local university rowing club. Forty four percent of the sample were male. Age and gender were insignificant determinants of choice preferences.

Forty four percent of all respondents were employed full-time. The highest rates of full-time employment were in the management and swimmers categories, which had 85% and 80% respectively. 50% of rowers were employed full-time. Respondents recruited door to door had the lowest rate of full-time employment (40.3%). Forty six percent of respondents recruited at Whitlingham were employed full-time. Across the sample 8.5% of respondents were employed in an environmental occupation. All of the respondents in the management community had an environmental occupation, as did 40% of rowers and swimmers. In contrast, less than 3% of the respondents recruited door to door and none of the respondents recruited at Whitlingham had environmental occupations. Occupation and employment type were insignificant determinants of choice preferences.

In addition to having the highest rates of employment, the expert and swimmer groups had the highest mean incomes at 52.7 and 57.0 thousands, respectively. These were far higher than the mean income for all respondents, which was 28.4 thousands. The lowest mean income was held by the respondents interviewed door to door (24.1 thousands). The average income of the public interviewed at Whitlingham was 35.0 thousands and the average income of the rowers was 33.0 thousands. As with age, it is probable that the mean income for the rowers was suppressed, as the majority of the student rowers had relatively low incomes.

An average of 46% of all respondents had a degree level education. Whitlingham respondents (46%) and door to door respondents (42%) were close to the mean. Experts (57%), rowers (60%) and swimmers (100%) had higher than average numbers of respondents with degree level education. Educational type or academic attainment was not a significant determinant of respondents choice decisions. There was also no correlation (p=0.0) between educational type and income.

Figure 25: distribution of the distance respondents live from the closest point of the Yare



Sample mean distance= 7.99km

Figure 25 shows the skewed distribution of the distance from respondents' homes to the nearest point of the River Yare. The effect of distance decay on willingness to pay was of considerable interest within this research. Despite being recruited at the survey stretch, the Whitlingham respondent group lived the highest mean distance from the Yare (11.24km). This average distance is positively skewed as several respondents (outliers) had travelled considerable distances to visit Whitlingham as their preferred recreation site. Respondents recruited door to door had the second highest average distance from the Yare (7.75km). Rowers (4.86km), swimmers (1.83km) and experts (3.57km) all lived less than the mean distance (7.99km) from the Yare.

Within the survey questionnaire, respondents were asked a series of questions relating to their trip behaviour over the year prior to the survey being administered (e.g. number of trips and types of activities undertaken on those trips). One year was chosen as a suitable period, as is standard with similar studies (e.g. Bateman et al., 2006a), to minimise recall bias[22].

[22] Recall bias is a systematic error caused by differences in the accuracy or completeness of the recollections retrieved ("recalled") by study participants regarding events or experiences from the past (Raphael, 1987).

Self-reported trip frequency and trip activity data indicates widely different rates of visitation and numbers of activities undertaken by the different respondent groups over the last year. All of the swimmers, rowers and experts had visited the Yare in the last year. All but one of the respondents interviewed at Whitlingham visited the Yare (the sole exception was attending a meeting at the park's cafe). The lowest rate of visitation was found in the door to door respondent group, of which only 39.7% had visited in the last year. In addition to containing a large proportion of non-visitors, respondents interviewed door to door took the lowest number of trips (9.8), on average ,during the last year. Whitlingham respondents visited an average of 33.9 times. Experts visited an average of 49.4 times, just under once per week, although their trips were often work related. Rowers visited most frequently, 171.6 times, or just over three times a week. Swimmers took an average of 94.4 trips, visiting just under twice a week. Door to door respondents did the lowest number of different activities, 0.9 activities, on average, when they visited. The average number of activities undertaken by all respondents on their visits was 1.4. Rowers (2.8), swimmers (2.3), experts (2.4) and respondents recruited at Whitlingham (2.4), all participated in an higher than average number of different activities when they visited.

Respondents provided estimates of the number of future trips they would take over the next year, if ecological quality and recreational quality were guaranteed to be high. 198 of the 200 respondents reported they would either visit more often or visit the same number of times. The two respondents who reported that they would visit less often provided rational reasons for their choices. Despite increased trip frequency across the sample as a whole, respondents’ estimated future use corresponds closely with their current patterns of use. Door to door respondents would still visit relatively infrequently, they estimated that they would visit 20.7 times over the next year, an increase of just under 11 trips per year. Whitlingham respondents stated they would visit 43.2 times in the next year, an increase of just 3 trips. This is a very small increase but it is worth considering that these respondents typically visit the Yare to enjoy non-contact activities (e.g. socialising, dog walking), which can be enjoyed irrespective of water quality. This suggests that water quality is not a strong determinant for their reason for visiting. Experts reported that they would visit an additional 43 times per year. In interview it was revealed that the majority of this increase would be to undertake more leisure activities. Swimmers said they would visit an additional 36 times in the next year if improvements were made to guarantee high water quality. The majority of this increase would be to swim more frequently, as, at present, there are periods when blue-green algae concentrations are unacceptably high, preventing swimming at the site. Rowers stated that they would visit the same number of times. Their visitation rates appear to be saturated: they visit so frequently at present (c. 3-4 times a week), that it would not be realistic for them to visit more often.

The differences in respondents' perceptions of water quality at the Yare is interesting. Experts tend to have greater access to accurate information on water quality so it is unsurprising that the experts had the highest perceptions of water quality. 42.9% of experts felt the ecological water quality was Blue and the rest felt it was Green. Swimmers and rowers were the next best informed on ecological water quality as their recreation clubs regularly provide information on water quality. Twenty percent of swimmers felt the ecological water quality was Blue, the rest thought it was Green. Twenty percent of rowers felt it was Blue, 50% Green and 30% thought Yellow. Non-visitors had the lowest perceptions of ecological quality. Door to door and Whitlingham respondents, the two non-expert, non-contact respondent groups, had very similar perceptions and were the only categories to contain respondents who perceived the ecological quality to be Red (approximately 4.5%). 21% felt the quality was Yellow, 53% felt it was Green and a 21% thought the quality was Blue.

There was a very similar pattern in respondents' perceptions of recreational water quality. Again, experts had the highest perceptions with 57% perceiving the recreational quality to be High and the remainder believing the water quality was Medium. Again, non-experts and non-visitors had the lowest perceptions of the recreational water quality: approximately 4.5% felt it was low, approximately 68% felt it was Medium and the remainder, 28.5%, felt the recreational quality was High. Swimmers and rowers had very similar perceptions. 30% of rowers and 20% of swimmers thought the recreational quality was High and the remainder from both groups, 70% of rowers and 80% of swimmers, felt the recreational quality was Medium.

After completing the choice task the respondents were asked to rate the importance of four key issues when making their choices. These were the ecological and recreational qualities of the water, the size of water bill increases and the distance from where they lived to where the quality improvements would happen. These four issues were rated by respondents as very important, important, neither, unimportant or completely unimportant.

The size of increases to the water bill (price) was important to an average of 62% of the respondents. Price was of most importance to the public interviewed at Whitlingham (69.2%) closely followed by the public interviewed door to door (62.5%). Forty percent of swimmers and rowers and 57.1% of the experts felt the size of bill was important to them when making their choice decisions.

Swimmers have the lowest proportion of respondents (20%) who felt that the issue of the distance from their home to where the improvement would happen is important. There are several reasons swimmers feel distance is unimportant: swimmers, on average, live closer to the Yare than any other respondent group and their recreation club, Tri-Anglia Triathlon Club, offers highly specialised recreation facilities and training opportunities at the survey river stretch. Forty percent of rowers felt the distance was important to them. This was just under the sample mean of 44%. As with the swimmers, rowers' club facilities are located on the survey river stretch at predefined, unchanging locations, without suitable substitutes nearby.

Respondents recruited at Whitlingham account for 19% of the total sample, and, of these, 74% felt that the distance was unimportant to them when making their choice decisions - despite this group living the highest mean distance from the Yare. This represents a sizable proportion of the whole sample for whom distance is unimportant. It may be that suitable alternative substitute sites are available for these respondents to enjoy less specialised activities. However, it appears that respondents in this category regularly visit the Yare as their preferred destination for dog-walking, exercise and other bankside activities, and they may also have an increased preference for visiting the survey river site due to the enhanced potential for socialising afforded by the indoor visitor centre and cafeteria.

Just under half of the respondents interviewed door to door felt distance was important; however, distance was particularly important to clusters of respondents who lived at greater distances from the Yare. For example, several respondents in Bungay, Beccles and Wroxham stated that they would rather pay to improve alternative rivers, e.g. the rivers Bure or Waveney, which are closer and more convenient for them.

All but one of the experts felt that the distance was an important factor during their decisions in the choice task. This figure seems rather high given that the experts also tended to live quite close to the Yare. It may be that distance is important to these respondents because they are professionally invested in the survey river stretch: improvements at the survey site benefit them both personally and professionally, whereas they have no professional responsibility for rivers outside of their jurisdiction.

Almost all respondents felt that the ecological quality of the water was important. All of the experts, rowers and swimmers thought it was important and 94.9% of respondents interviewed door to door and at Whitlingham thought it was important. There were large differences of opinion regarding the recreational quality of the water. All of the swimmers, 90% of the rowers and 85.7% of the experts felt that recreational quality was important when making their choice decisions. In marked contrast, only slightly more than half of those interviewed at either Whitlingham (51.3%) or door to door (53.9%) felt that the recreational quality of the water was important. In face to face interviews, several door to door respondents stated they would prefer it if people were prevented from using rivers for swimming. Reasons for this opposition were that they felt that rivers are too physically dangerous for use and that there are public swimming pools for people to use, if people want to swim.

3.5.6 CL modelling

This section reports two CL models. As with the pilot data, the CL modelling of the main survey data was undertaken using Stata 13.1 (StataCorp L. P., 2013). For the interested reader, a chronology of CL modelling development (CL models 1-4) is included in Appendix V. Table 33 describes the variables used within the two CL models.

Table 33: variables used within econometric modelling

Dependent variables	
Price	Respondents' response to cost of water quality, expressed as a continuous variable.
Medium ecological quality	Composite variable composed of Yellow and Green ecological quality categories combined, expressed as a categorical variable. 1=Yellow and Green (Medium) ecological water levels, 0=other ecological levels.
High ecological quality	High ecological water quality, expressed as a categorical variable. 1=blue/high ecological water level, 0=other ecological levels
Medium recreational quality	Medium recreational water quality, expressed as a categorical variable. 1=medium recreational water level, 0=other recreational levels.
High recreational quality	High recreational water quality, expressed as a categorical variable. 1=high recreational water level, 0=other recreational levels.
RQ*EQ	Variable describing the interaction between recreational and ecological water quality, expressed as a continuous variable.
Independent variables	
Rowers and Swimmers	Composite variable composed of rowers and swimmers. Swimmers recruited via Tri-Anglia Triathlon Club, rowers recruited via local rowing clubs. Binary variable: 0=respondent is not a rower or swimmer, 1=respondent is a rower or swimmer.
Anglers	Respondents who are anglers. Binary variable expressed as 0=respondent is not an angler, 1=respondent is an angler.
Environmental memberships	Respondents who have environmental memberships. Binary variable expressed as 0=respondent does not have environmental memberships, 1=respondent has environmental memberships.
EnvMemberCont	The total number of environmental organisation memberships held by the respondent, expressed as a continuous variable.
Distance	The distance the respondent lives from the closest point of the Yare. Inverse multiplicative, expressed as 1/Distance.
DistanceBin	The distance the respondent lives from the closest point of the Yare. Binary variable: 0=respondent lives <8km, 1=the respondent lives >8km (Mean distance=7.99km).
Income	The respondents' gross household annual income. Inverse square, expressed as 1/Income^2.

Model 5 incorporates socio-economic variables and distinguishes between different respondent types. Model 6 represents a more parsimonious modelling solution. The 7 experts, having professional interest, are excluded from these analyses. Each model is now discussed in turn.

The CL models, shown on Table 34, report main effects interactions in the naïve way (interaction of attributes and socio economic variables) to provide insights into preference heterogeneity. Respondents' preferences for Yellow and Green ecological quality levels were insignificantly different from one another so those levels were collapsed into one intermediate variable called Medium ecological quality (see Wald test on p.310, Appendix V). For clarity Blue ecological quality is renamed High ecological quality. A definition of Model 5 is

$$U = V + \sum_{c=1}^{C} z_c * (\mathrm{V})$$

Where

$$V = \mathrm{b_p} * \mathrm{Price} + \mathrm{b_{em}} * \mathrm{MediumEQ} + \mathrm{b_{eh}} * \mathrm{HighEQ} + \mathrm{b_{rm}} * \mathrm{MediumRQ} + \mathrm{b_{rh}} * \mathrm{HighRQ}$$

Equation 17: definition of CL Model 5

In this case the c=[1,...,5] are the socio-economic characteristics specified in the model (e.g. rowers or swimmers, environmental memberships, etc.).

The first section of Model 5 displays estimated marginal utilities of the general public (i.e. have not used the river for swimming, boating or fishing over the last year, although they may have visited for other purposes) and who are not members of environmental groups. Coefficients for Medium and High ecological, and Medium and High recreational water quality levels are complete and transitive. The strength of the coefficients relative to one another suggests that such respondents, on average, value improvements in ecological quality more than they do improvements in recreational/microbial water quality. Respondents dislike options containing higher prices, *ceteris paribus*.

An interaction term, RQxEQ, describes a highly significant positive interaction for all respondents: improvements in one dimension of water quality (whether it be ecological quality or recreational quality) are valued more highly the higher the quality level of the other dimension of water quality (The function RQxEQ is discussed further on p.316, Appendix V).

Table 34: CL models

	Model 5		Model 6	
Variable	**Coef.**	**s.e.**	**Coef.**	**s.e.**
Baseline coefficients				
Price	-0.024***	0.003	-0.021***	0.002
Medium ecological quality	1.213***	0.152	1.860***	0.090
High ecological quality	2.194***	0.228	3.111***	0.128
Medium recreational quality	0.533***	0.145	0.771***	0.081
High recreational quality	0.739***	0.188	1.286***	0.086
RQ*EQ	0.157***	0.045		
Socio-Economic Coefficients				
Distance from river (1/distance)				
DistancexPrice	0.005**	0.002	0.006***	0.002
Income (1/income^2)				
IncomexPrice	-145842.3**	72882.52	-18460.1***	63811.14
Rowers and Swimmers (n=15)				
Rowers and swimmersxPrice	-0.003	0.011		
Rowers and swimmersxMedium ecological quality	-1.190***	0.390		
Rowers and swimmersxHigh ecological quality	-1.880***	0.617		
Rowers and swimmersxMedium recreational quality	1.181***	0.415		
Rowers and swimmersxHigh recreational quality	1.842***	0.532		
Anglers (n=16)				
AnglersxPrice	-0.0003	0.009		
AnglersxMedium ecological quality	2.364***	0.791		
AnglersxHigh ecological quality	4.011***	1.070		
AnglersxMedium recreational quality	0.101	0.635		
AnglersxHigh recreational quality	-1.035*	0.573		
Environmental Memberships (Binary variable)				
Environmental membershipsxPrice	0.004	0.005		
Environmental membershipsxMedium ecological quality	1.003***	0.238		
Environmental membershipsxHigh ecological quality	1.320***	0.334		
Environmental membershipsxMedium recreational quality	-0.022	0.199		
Environmental membershipsxHigh recreational quality	0.215	0.220		
Number of observations = 4632				
Pseudo R^2	0.403		0.360	
Log Likelihood	-957.700		-1026.820	

Model 5: CL model of price, categorical water quality levels, EQ*RQ interaction term, with distance, income, swimmers and rowers, anglers, and environmental membership as interaction terms. Model 6: CL model of price, categorical water quality levels with distance and income interaction terms. *, ** and *** = significance at 10%, 5% and 1% levels

Within the sample, we find a significant distance decay in respondents' choice preferences; the further from the survey river the respondent lives, the less likely they are to choose choice alternatives containing higher prices. Distance is expressed as the multiplicative inverse ($1/x$). The effect of distance on respondents' WTP is discussed below.

Income has a significant, positive impact on respondents' willingness to choose choice alternatives containing higher prices. Income is expressed as the inverse square ($1/x^2$). The effect of income on respondents' WTP is discussed below.

Swimmers and rowers are significantly more likely to value improved recreational water quality and significantly less likely to choose options containing higher ecological quality. This is reasonable given that improved levels of recreational water quality is important in order for them to enjoy their activities safely.

Anglers are significantly more likely to value improved levels of both Medium and High ecological water quality. These preferences make sense when viewed from an angler's perspective. Anglers require rivers with Medium ecological quality for coarse fishing and High ecological quality for game fishing. Anglers have significantly lower preferences (relative to the other respondents) for High recreational water quality. This is reasonable if lower recreational quality reduces the number of people using the river and disturbing the angler and the fish. In interview, several anglers stated preferences for quiet undisturbed locations.

Membership of environmental organisations is typically used as a surrogate variable to positively identify respondents who would be expected to care more highly about the environment. During the survey, respondents were asked if they held personal memberships for any environmental organisations[23]. Membership of environmental organisations are expressed as a binary variable (0 = not a member, 1 = a member of one or more environmental organisations). Within this sample, members of environmental organisations have highly significant preferences for higher levels of ecological water quality.

Model 6 is now discussed. Although this parsimonious model has slightly lower explained variance (R^2), it is, arguably, more policy relevant as it contains important, relevant, interaction terms (i.e. distance and respondents' income) which, unlike the highly specific socio-economic variables (e.g. recreational use

[23] Such organisations include walking clubs, ramblers associations, angling clubs, river recreation (e.g. rowing/canoeing/sailing) clubs, National Trust, RSPB, English Nature, Norfolk Wildlife Trust or similar, Greenpeace, Friends of the Earth, WWF, other environmental groups, outdoor swimming/triathlon clubs.

type, environmental memberships), can be readily obtained from secondary data sources without recourse to further survey work.

As with Model 5, respondents' dislike options containing higher prices, *ceteris paribus*. Coefficients for Medium and High ecological, and Medium and High recreational water quality levels are complete and transitive. Again, the strength of the water quality coefficients relative to one another suggests that respondents, on average, value improvements in ecological quality more than they do improvements in recreational water quality. Model 6 contains two interaction terms. One describes a significant distance decay in respondents' willingness to choose choice alternatives containing higher prices. The other shows that income has a significant, positive impact on respondents' willingness to choose choice alternatives containing higher prices.

3.5.7 Marginal WTP estimates derived from Model 6

This section reports the marginal WTP estimates for changes in attribute levels in CL Model 6. The marginal WTP estimates are the negative of the ratio between the mean coefficients for each attribute and the mean coefficient of the payment attribute (please see Equation 16, p.197).

Monetary values for WTP are derived by assessing changes in utility from V^0, the initial water quality state, and V^1, the alternative state. As discussed previously, using the correct value for V^0 is crucial, as, if incorrect, the resulting WTP estimates will also be incorrect. For example; if ecological quality is consistently low it would be correct to set V^0 for ecological quality to low. However, water quality on the Yare isn't low, but variable throughout the year, which complicates our attempts to accurately define V^0.

Tables 35 - 37 and Figures 26 and 27 report the welfare estimates for marginal changes in river water quality improvements, derived from Model 6. V^0 is, for this analysis, defined by Low ecological and Low recreational water quality levels. The data will subsequently be reanalysed using respondents' perceptions of water quality as V^0.

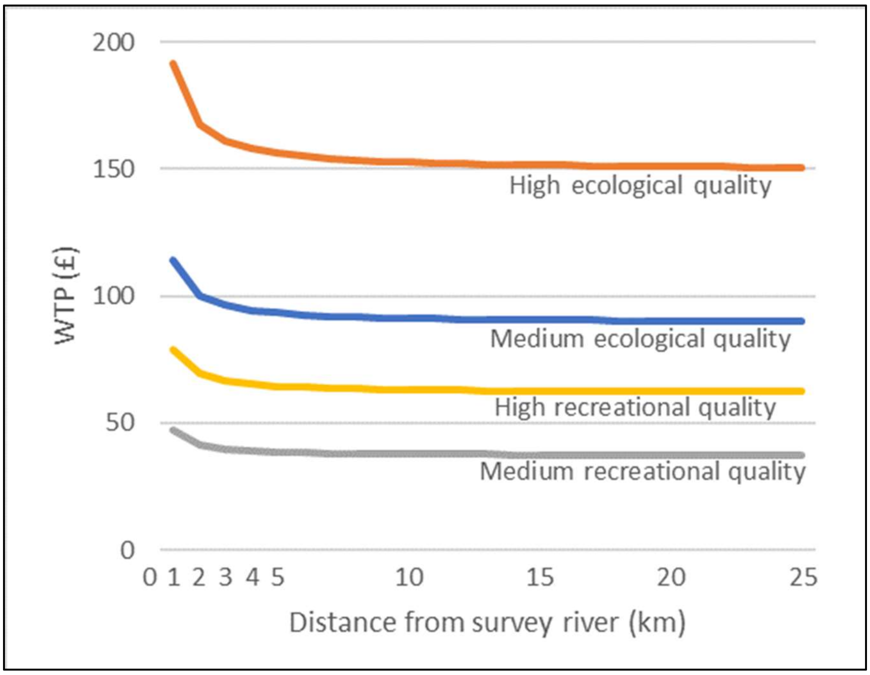
Table 35: marginal WTP estimates derived from Model 6 using Low water quality as the baseline

	Medium ecological quality	High ecological quality	Medium recreational quality	High recreational quality
WTP (£)	£89.23***	£149.39***	£37.00***	£61.74***
95% confidence intervals				
Lower limit	£70.51	£119.73	£26.59	£48.53
Upper limit	£108.08	£179.04	£47.42	£74.96

Note *, ** and *** = significance at 10%, 5% and 1% levels. WTP=£, per household, per year for the 20km survey river stretch

We now consider the impact of the distance and income interaction terms. Figure 26 shows the impact of distance on respondents' WTP for water quality. Figure 27 shows the impact of income on respondents' WTP for water quality. The effect of distance and income on respondents' WTP values is described numerically in Tables 36 and 37.

Figure 26: distance decay in respondents' WTP for water quality

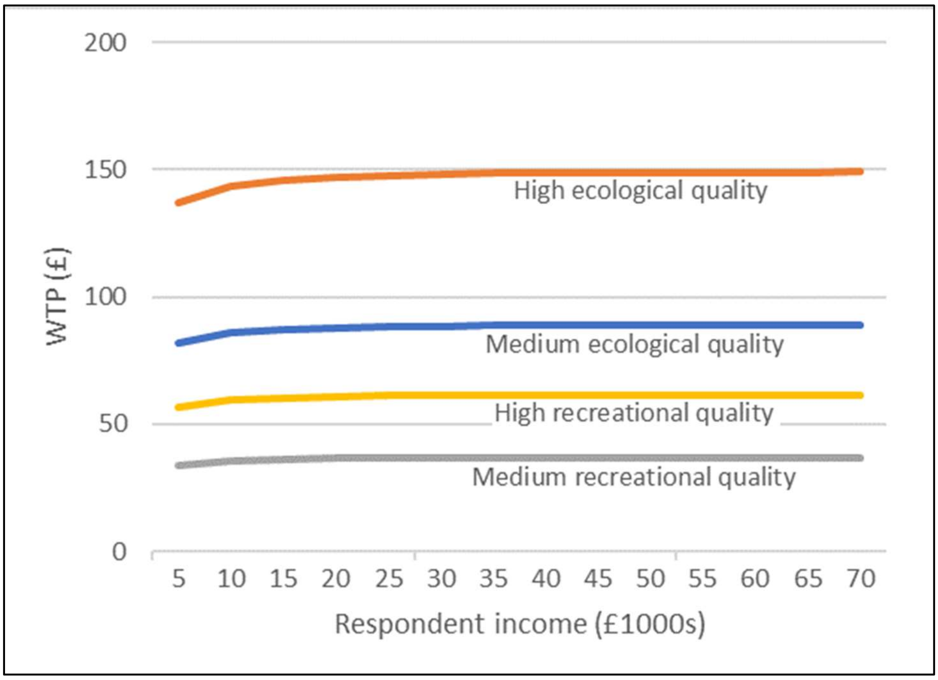


WTP = £ per household, per year. Distance is measured from the respondent's home to the closest point on the survey river stretch. Low water quality defines V^0, the baseline.

Figure 26 and Table 36 show that as respondents live further from the river, their estimated WTP for water quality decreases. Among respondents living closer to the river we see significantly higher WTP to secure each level of both water quality attributes. For example, a respondent who lives 2km from the river is estimated to be willing to pay £167.83 for High ecological water quality and £69.37 for High recreational water quality, whereas a respondent who lives 25km from the river has estimated WTP of £150.71 and £62.79, respectively.

The income interaction term describes a highly significant positive relationship between income and WTP. For example, the model predicts that a respondent with a household income of £10,000 is willing to pay £143.72 for High ecological water quality and £59.40 for High recreational water quality, whereas a respondent with a household income of £60,000 is estimated to be willing to pay £149.07 and £61.62, respectively (Figure 27 and Table 37).

Figure 27: impact of respondents' income on WTP



WTP = £ per household, per year. Respondent income is defined as the respondent's annual gross household income. Low water quality defines V^0, the baseline.

Table 36: the effect of distance on respondents' WTP

	Distance from respondent's home to the closest point on the survey river stretch (km)													
Distance (km)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**15**	**20**	**25**	**30**
Water quality type	**WTP (£ per household, per year for the 20km survey river stretch)**													
High ecological	191.47	167.83	161.20	158.07	156.25	155.07	154.23	153.61	153.13	152.74	151.61	151.05	150.71	150.49
Medium ecological	114.45	100.32	96.35	94.49	93.40	92.69	92.19	91.82	91.53	91.30	90.62	90.29	90.09	89.95
High recreational	79.14	69.37	66.63	65.34	64.58	64.09	63.75	63.49	63.29	63.13	62.66	62.43	62.29	62.20
Medium recreational	47.43	41.57	39.93	39.16	38.71	38.41	38.20	38.05	37.93	37.84	37.55	37.41	37.33	37.28

Low water quality defines V^0, the baseline.

Table 37: the effect of income on respondents' WTP

	Respondent's gross annual household income (£1000s)													
Income (£1000s)	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**55**	**60**	**65**	**70**
Water quality type	**WTP (£ per household, per year for the 20km survey river stretch)**													
High ecological	137.22	143.72	146.15	147.30	147.93	148.31	148.56	148.73	148.86	148.95	149.02	149.07	149.12	149.15
Medium ecological	82.02	85.91	87.36	88.04	88.42	88.65	88.80	88.90	88.98	89.03	89.07	89.11	89.13	89.15
High recreational	56.72	59.40	60.41	60.88	61.14	61.30	61.40	61.48	61.53	61.56	61.59	61.62	61.63	61.65
Medium recreational	33.99	35.60	36.20	36.49	36.64	36.74	36.80	36.84	36.87	36.90	36.91	36.93	36.94	36.95

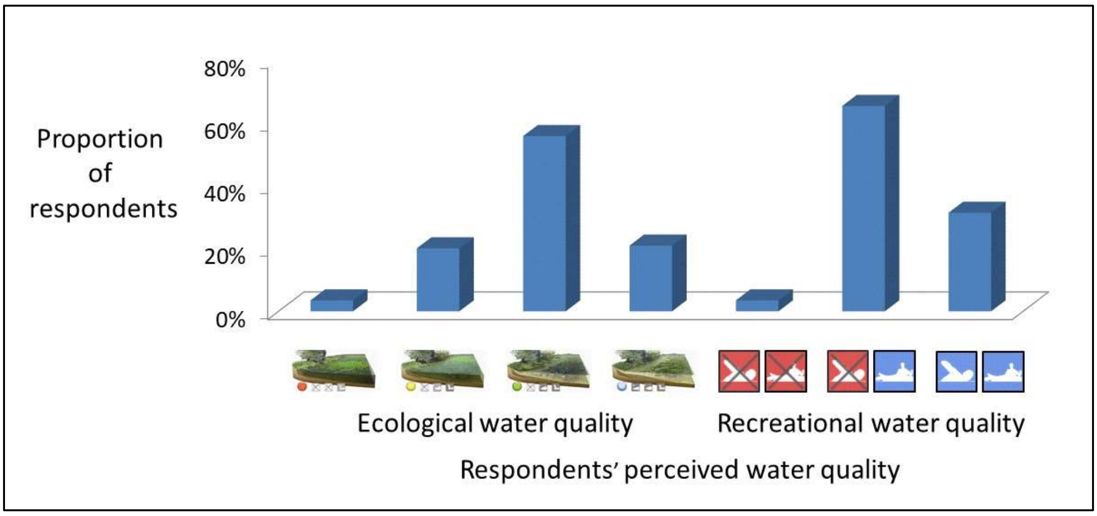
Low water quality defines V^0, the baseline.

3.5.8 WTP estimates based on respondents' perceptions of water quality

It is important that we use the correct level for V^0 to produce meaningful valuations: if V^0 is systematically set to the lowest water quality level WTP estimates are potentially overestimated. Perhaps the most important factor influencing the correct level of V^0 in situations where V^0 is variable, or where there is no correct level of V^0, is the respondents' perceptions of existing water quality. For this reason, WTP is adjusted for each individual, with V^0 set to the level of water quality perceived by that individual. Individual values are aggregated to produce estimates of WTP for each level of water quality attributes.

Within the survey the current state of the water quality was not fixed and was intentionally overlooked in the CE setting. Instead, respondents were asked what quality they thought ecological and recreational water quality was at the Yare. 153 respondents visited the Yare in the year prior to the survey. Their perceptions of water quality are shown on Figure 28.

Figure 28: Yare visitors' perceptions of water quality



Yare visitors' perceptions of water quality correspond relatively closely to the Environment Agency's estimates of Yare catchment water quality characteristics (2014)[24]. We see a small minority who believe that the current ecological water quality is Low, while the remaining respondents think that the water quality is higher, either Yellow, Green or Blue. The majority of respondents believe the

[24] The status of the River Yare, and its tributary rivers, was assessed by the Environment Agency in 2013 as being generally 'moderate' which broadly corresponds to the Green ecological and Medium recreational water quality attribute levels used in this study.

current recreational water quality is Medium with a sizeable proportion who think it is High. A small proportion think the recreational quality is Low. Based on this data, respondents would, on average, receive a disutility should the level of future ecological or recreational water quality be lower than Medium, and would wish to be compensated for that reduction in water quality. Within these results the hypothesis of compensation is disregarded as it is not a real world option. Non-visitors, unable to provide perceptions data, are not excluded from this analysis. Instead, their perceptions of water quality are set to the modal values (i.e. Medium ecological quality and Medium recreational quality).

Table 38: marginal WTP (£ per household, per year), derived from CL Model 6, based on respondents' perceptions of water quality.

Improvement type	If V^0 = low water quality	if V^0 = respondent's perception
Medium ecological quality	£89.23	£1.85
High ecological quality	£149.39	£55.46
Medium recreational quality	£37.00	£0.74
High recreational quality	£61.74	£21.03

WTP (£ per household, per year) for the 20km survey river stretch. Estimates of WTP where V^0 = low water quality are transcribed from Table 35.

Table 38 reports estimates of WTP derived from Model 6 (Table 34). V^0 is set to respondents' perception of ecological or recreational water quality. Within Table 38 we see that where V^0 is set to the lowest water quality level, WTP estimates are much higher than if we use respondents' perceptions of water quality as the baseline upon which to calculate welfare estimates. We see that estimates of utility for medium levels of ecological and recreational water quality at the survey river stretch are very low. This is because the majority of Yare visitors (Figure 28) believe that the water quality at the site is already medium[25], therefore they do not receive any additional benefit if the future water quality were to remain at the same level. On average, respondents have higher WTP for High ecological, rather than High recreational water quality.

3.5.9 Shortcomings of using a CL model specification on the survey data

As discussed previously, there was a possibility that the assumptions underpinning the CL modelling may be violated. A z-test of the variance was

[25] Perceptions of non-visitors were also set to medium, the model value of perceived water quality across the sample.

performed on the main variables in Model 6 and it was found that Price and Medium recreational quality have heterogeneous variance[26]. This is shown in Table 39.

Table 39: results of z tests on the main variables of Model 6

Variable	Probability	Decision
Price	0.046	Reject
Medium recreational quality	0.000	Reject
High recreational quality	0.869	Accept
Medium ecological quality	0.087	Accept
High ecological quality	0.267	Accept

This lack of homogeneity within the baseline coefficients (Morey et al., 2006), in conjunction with the CL modelling failing to control for intra-respondent variation (Kemperman and Timmermans, 2006), indicates that a CL specification of the choice data is not the best specification.

3.5.10 LC modelling

As CL models were not the best modelling solution, LC models were explored. LC models of the dataset were generated using Latent Gold version 5 (Statistical Innovations, 2014). In estimating the LC models 2, 3, 4 and 5 class solutions were investigated. Preliminary investigation of the optimal number of classes for LC models was informed using the Bayesian Information Criteria (BIC), the Akaike Information Criteria (AIC) and the consistent Akaike Information Criteria (CAIC) (Allenby, 1990; Ben-Akiva and Swait, 1986). Swait (1994) explains how these criteria should be used to determine the optimum number of latent classes. The criteria are based, in part, on the likelihood function and compare the relative plausibility of different models. Under the assumption that we have no prior preference for one model over another, the criteria identify the model that is more likely to have generated the observed data: the smaller the value of the statistic, the better the fit of the model. Using Swait's scheme, the three class LC model appears to be the optimal solution for the data used in this study.

The minimum BIC statistic, shown on Table 40, suggests a 3 class solution, but only marginally so, as the BIC value for a 4 class solution is only slightly higher. The CAIC statistic corroborates this finding. The AIC value continues to decrease

[26] Heterogeneous variance was also present in CL Model 4.

beyond a 3 class model and suggests a 5 class model. However, McLachlan and Peel (2004) warn us that the AIC tends to overestimate the most efficient number of classes due to a failure in the regularity conditions affecting the likelihood ratio test statistic. Although the BIC also fails certain regularity conditions, Leroux (1992) has shown that the BIC does not asymptomatically overestimate the optimal number of classes. Furthermore, the BIC has been described as the most conservative of the various indicators as it penalises additional parameters (Morey and Thacher, 2012). Wedel and Kamakura (2012) warn us that the various indicators are at best suggestive. With the above in mind, further investigation of the optimal number of classes was performed by hand, referring to the significance of the individual variables within the classes. A 2 class models was immediately dismissed as lacking in explanatory power. Although the BIC of a 4 class models is similar to that of a 3 class model, 4 class models resulted in large numbers of insignificant class variables and insignificant class membership covariates. For example, within a 4 class solution it was found that all but one of the variables in class 3 were insignificant.

Table 40: information criteria values determining the optimum number of classes when modelling all respondents

Number of classes	BIC	AIC	CAIC
2	1928.67	1885.80	1941.67
3	1782.68	1716.71	1802.68
4	1783.18	1694.13	1810.18
5	1804.51	1692.37	1838.51

Having decided on a three class solution, a number of 3 class models were generated to explore the effectiveness of different combinations of socio-economic variables as class membership covariates. Most of the socio-economic variables proved insignificant in explaining class membership and, in common with the CL modelling, the most significant socio-economic variables were the number of environmental memberships held by respondents and the distance respondents lived from the survey river. As a Wald test on the coefficients of the Yellow and Green ecological water quality parameters in the CL modelling (see p.310) demonstrated that they were insignificantly different from one another, it was prudent, and parsimonious, to use the Medium ecological quality variable in the LC modelling.

The most informative LC models have three classes using price, Medium and High ecological quality and Medium and High recreational quality as explanatory choice attribute variables and environmental memberships (EnvMemberCont) and distance (DistanceBin) as class membership covariates. For both of the LC models reported below the interaction variable RQ*EQ was insignificant.

The first of the LC models, Model 7, defines three distinct classes of respondents. All 200 respondents are included within this LC model. The discussion of the results this model makes reference to salient points from the model's postestimation results, reported in Table 43. Following this, the marginal WTP estimates, using the coefficients of the three class LC model, and the remaining postestimation results are reported.

Table 41: Model 7, a 3 class LC model with distance and environmental membership as class membership covariates

	Class 1	Class 2	Class 3
Class membership covariate coefficients			
Intercept	-0.269 (0.245)	0.460** (0.205)	-0.191 (0.236)
Environmental memberships (EnvMemberCont)	0.476** (0.229)	0.307 (0.218)	-0.783** (0.388)
Distance (DistanceBin)	-0.488* (0.272)	-0.371 (0.238)	0.859*** (0.331)
Choice attribute coefficients			
Price	-0.014*** (0.005)	-0.023*** (0.003)	-0.285*** (0.071)
Medium ecological quality	1.211*** (0.171)	3.546*** (0.281)	1.040** (0.448)
High ecological quality	1.647*** (0.262)	5.790*** (0.371)	1.903*** (0.727)
Medium recreational quality	2.006*** (0.208)	0.480*** (0.181)	0.984 (0.768)
High recreational quality	2.941*** (0.252)	0.993*** (0.175)	2.507* (1.406)
Class probability	0.31	0.585	0.105
Predicted number of respondents	62	117	21
Total respondents = 200			
Total observations = 4800			

*, ** and *** = significance at 10%, 5% and 1% levels. Standard errors in parenthesis

Class 1 has a class probability of 0.31. This means that 62 of the 200 respondents are estimated to be in this class. The number of environmental memberships held by these respondents is a significant determinant of Class 1 membership and the positive coefficient, 0.476** (s.e. 0.229), predicts that as the number of environmental memberships held by respondents increases, the more likely it is that they will be assigned to Class 1. Postestimation results, Table 43, predict that Class 1 respondents have the highest average number of environmental memberships, 0.61 (s.e. 0.83), compared with an average of 0.51 (s.e. 0.72) memberships for all respondents. The negative coefficient for the distance covariate, -0.488* (s.e. 0.272), predicts that respondents who live closer to the Yare are more likely to be assigned to Class 1. Their mean distance to the Yare is 7.8km (s.e. 10.3), against the sample mean of 8.0km (s.e. 10.5).

Of the three classes, members of Class 1 are least averse to increased price. They have the smallest of the price coefficients, 0.014*** (s.e. 0.005), and are only slightly less likely to choose an option with increased price. This is despite the class having the smallest percentage of respondents who identify as being employed, 35.5%, and respondents who earn slightly less than average income of £28,300 (s.e. 21.3). Class 1 has the highest proportion of respondents, 11.36%, who have an environmental occupation and 45.2% of these respondents have a degree level education.

Respondents in Class 1 are more likely than other respondents to choose choice alternatives containing higher levels of recreational quality. Postestimation results suggest that Class 1 contains 80% of the swimmers and 70% of the rowers, conforming to our *a priori* expectations that these user groups prefer, and value, recreational water quality highly.

Class 2 has the largest class probability, 0.585, and is estimated to contain 117 respondents. Environmental memberships and distance are not significant class membership covariates for Class 2.

Class 2 respondents are only slightly more averse to price increases, -0.023*** (s.e. 0.003), than Class one respondents, -0.014*** (s.e. 0.005). Class 2 has the highest proportion of respondents who identified as employed full time, 48.7%. This is above the sample mean of 44.5%. Class 2 respondents have the highest mean income, £29.5k (s.e. 18.4), of the three classes.

Class 2 respondents are more likely to choose choice alternatives containing higher levels of ecological quality than members of the other two classes. The postestimation results suggest the same, as 14 of the 16 anglers and 4 of the 7 experts are predicted to be in Class 2. As we saw from the CL modelling, anglers have significantly higher utility from ecological, rather than recreational water quality.

Class 3 has the smallest class probability, 0.105, and is estimated to contain 21 respondents. The number of environmental memberships held, and the distance Class 3 respondents live from the survey river, are significant class membership covariates. As the number of environmental memberships held by respondents increases, respondents are less likely to be assigned to Class 3. The postestimation results support this coefficient. They predict Class 3 respondents to have, on average, 0.14 (s.e. 0.35) environmental memberships, far lower than the sample mean of 0.51 (s.e.0.72). As distance from the river increases, respondents are significantly more likely, 0.8589*** (s.e. 0.331), to be assigned to Class 3. Again, the postestimation results support the distance coefficient: the mean distance from Class 3 respondents' homes to the survey river is 9.7km (s.e. 8.5km), almost 2km further, on average, than other respondents.

In common with respondents from other classes, respondents in Class 3 have complete and transitive preferences for water quality. Postestimation results suggest that this class is composed largely of respondents who rarely use the survey river for recreation. It is estimated that the class contains no swimmers or rowers, one angler, one expert and is largely composed of non-visitors. This class contains a high proportion (15 of the 21 class members) of the general public recruited door to door.

In marked contrast to respondents assigned to the other two classes, respondents in Class 3 have a much higher sensitivity to increased price and are significantly less likely to choose a choice option with higher price, -0.285*** (s.e. 0.071), against -0.014*** (s.e. 0.005) and -0.023*** (s.e. 0.003) for Class 1 and Class 2 respondents. Postestimation results hint at possible reasons for their aversion to price increases. Although an estimated 47.6% of this class identify as employed full-time (above the average of 44.5% of all respondents), their mean income, at £24,900 (s.e. £21,900) is almost £4,000 below the average income of all respondents. Although income was not a statistically significant determinant of LC class membership, CL Model 5 does find Income to be significantly positively correlated with respondents' willingness to choose choice alternatives containing higher prices. On face value, it seems that these respondents are, on average, in low paid employment and may prefer their disposable income to be directed elsewhere. Postestimation results also predict that members of this class have the lowest proportion of respondents, 28.6%, holding a degree level education. Class 3 respondents' sensitivity to choices with increased price impacts dramatically on their WTP for water quality, as will now be reported.

3.5.11 Marginal WTP estimates derived from the 3 class LC model

WTP derived from the 3 class LC model are now discussed. Changes in utility are assessed using Low water quality to define V^0, the baseline.

The rejection by Class 3 respondents of choice options with higher prices, results in them having the lowest WTP for both water quality attributes (Table 42).

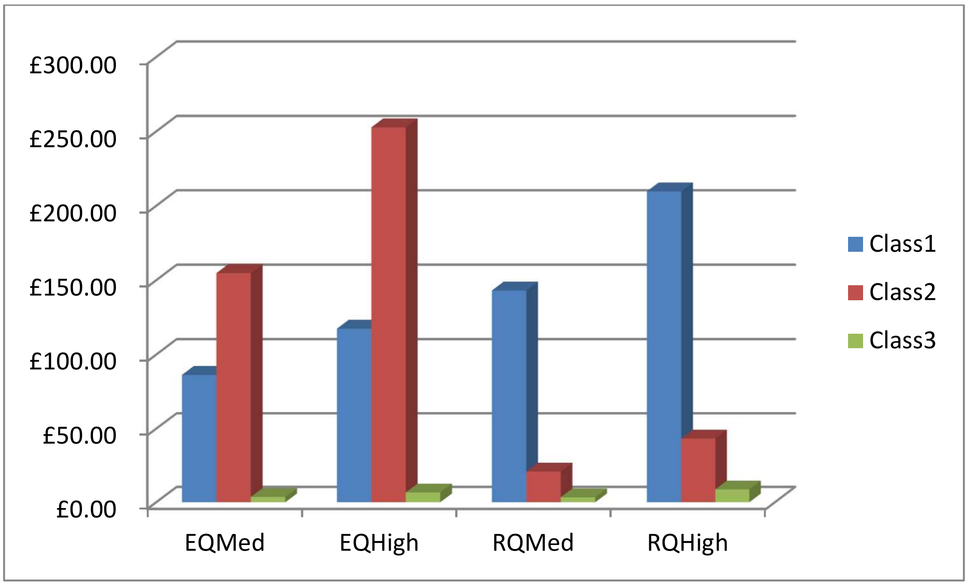
Table 42: marginal WTP estimates derived from the 3 class LC model

Water quality type	Class1		Class2		Class3	
	WTP (£)	s.e. (£)	WTP (£)	s.e. (£)	WTP (£)	s.e. (£)
Medium ecological quality	£86.38	£30.45	£154.81	£24.54	£3.65	£1.80
High ecological quality	£117.45	£41.79	£252.75	£36.89	£6.68	£3.01
Medium recreational quality	£143.10	£48.77	£20.94	£9.14	£3.45	£2.09
High recreational quality	£209.72	£68.24	£43.34	£9.13	£8.79	£3.24

Standard errors represent the 95% confidence intervals for the WTP estimates. WTP=£, per household, per year for the 20km survey river stretch. Low water quality defines V^0, the baseline

Class 1 respondents have a clear preference for the highest possible recreational quality. Despite their preference for recreational water quality, Class 1 respondents' WTP for ecological quality is relatively high. Class 2 respondents' WTP for ecological water quality far exceeds that of the other two classes.

Figure 29: visual representation of WTP derived from the 3 class LC model



The differences in WTP by the three classes for the different water quality categories are explicitly shown in Figure 29.

Several of the postestimation results shown on Table 43 have been reported previously. However, the remaining postestimation results yield some interesting insights, which are now reported.

With regard to distance and trip information, the postestimation results predict that Class 3 respondents visit the Yare less frequently and do less activities when they do visit.

Table 43: post estimation results for the 3 class LC model

	Class 1	Class 2	Class 3	Whole Sample
Class probability	.31	.585	.105	1.0
Number of respondents	62	117	21	200
Mean age	53.2 (19.5)	50.0 (17.5)	48.9 (20.7)	50.9 (20.7)
Gender(% male)	43.5	43.6	47.6	44
Employment, income, education and environmental affiliation				
Employed full time (%)	35.5	48.7	47.6	44.5
Employed in environmental occupations (%)	11.3	7.7	4.8	8.5
mean income (£1000s)	28.3 (21.3)	29.5 (21.8)	24.9 (18.4)	28.6 (21.4)
% respondents with degree level education or higher	45.2	49.6	28.6	46
mean number of environmental memberships per respondent	0.61 (0.83)	0.52 (0.69)	0.14 (0.35)	0.51 (0.72)
Predicted number of each respondent type in class				
Anglers	1	14	1	16
Experts	2	4	1	7
Rowers	7	3	0	10
Swimmers	4	1	0	5
General public (door to door)	37	87	15	139
General public (Whitlingham)	12	22	5	39
Distance and trip information				
Mean distance respondents live from the Yare (km)	7.8 (10.3)	7.8 (10.9)	9.7 (8.5)	8.0 (10.5)
Mean number of all river trips taken by respondents in the last year	48.7 (70.9)	35.2 (66.2)	31.2 (51.4)	39.0 (66.7)
Mean number of activities at the most visited site in the last year	1.7 (1.3)	2.1 (1.5)	1.6 (1.2)	1.9 (1.5)
% of respondents visiting the Yare in the last year	61.3	59.0	33.3	57.0
Mean number of Yare trips taken by respondents in the last year	40.0 (67.7)	22.7 (54.9)	3.9 (11.1)	26.1 (57.6)
Mean number of activities at the Yare in the last year	1.4 (1.3)	1.5 (1.5)	0.7 (1.2)	1.4 (1.4)
Mean number of Yare trips in the next year if water quality improvements are made	44.1 (68.1)	40.4 (79.5)	5.5 (11.1)	37.9 (72.6)
Importance of issues when making choice decisions (% respondents)				
The size of the bill	59.7	56.4	100.0	62.0
The distance to any proposed improvement	41.9	43.6	52.4	44.0
The ecological quality of the river	98.4	99.1	66.7	95.5
The recreational quality of the river	75.8	54.7	19.0	57.5
Protesting?	12.9	12.8	19.0	13.5

Standard errors in parenthesis

However, when we examine the data relating to trips to *all* river locations, a different picture begins to emerge. Despite Class 3 respondents visiting the survey river stretch on the Yare infrequently, they visit other rivers on average 31.2 (s.e. 51.4) times a year. We have previously seen that Class 3 respondents have consistent preferences for water quality – they do care about water quality – but they tend to live further from the survey stretch and appear to prefer to visit other river sites instead. These results are in line with the survey interviewer's experiences: several respondents within the Beccles area told the interviewer that they would rather visit the River Waveney and they had no desire to fund improvements at a river site that they do not visit.

Class 1 respondents make the greatest average number of trips each year, 48.7 (s.e. 70.9), but take the majority of their trips, 40.0 (s.e. 67.7), at the survey river stretch on the Yare. Class 1, as we have seen, contains the majority of the recreational users, who use the recreational facilities on the Yare frequently.

Class 2 respondents have a more balanced mix of destinations. They visit the survey river stretch 22.7 (s.e. 54.9) times and take the remainder of their 35.2 (s.e. 66.2) annual trips elsewhere.

During the interviews respondents were asked how often they would visit the Yare in the coming year if the ecological and recreational quality of the water was guaranteed to be high. The *a priori* expectation was that if high water quality was guaranteed, respondents who currently visit the Yare would visit more often and respondents who do not visit may start visiting.

Class 3 respondents' apathy towards recreation at the survey river stretch can be seen in their stated number of future visits, which rises only slightly, from 3.9 (s.e. 11.1) current trips to 5.5 (s.e. 11.1) future trips. This negligible change suggests that their demand for recreation at the Yare is not at all related to the quality of river water. It appears that they do not wish to visit the survey river whatever its quality.

Class 1 respondents' proposed future trip frequency is also inelastic, showing a small rise, if high water quality is guaranteed. This may in part be explained by the high frequency of visits made by the rowers and swimmers who would, as previously discussed, find it difficult to visit more often.

In contrast to respondents in the other classes, Class 2 respondents future recreational demand is far more elastic. Their proposed future trip frequency rises from an average of 22.7 (s.e. 54.9) trips over the last year to an average of 40.4 (s.e. 79.5) trips in the coming year. This suggests a class of respondents who would be willing to pay more for ecological improvements and would be keen to visit the river more often to enjoy those improvements.

Respondents were asked to quantify the importance of different issues when making their choice decisions. These issues were bill size, the ecological quality, the recreational quality and the distance from where they lived to where the improvement would happen. Their answers to these questions, on a class by class basis, tend to confirm and reinforce the results of the LC analysis.

Within this LC model, Class 3 respondents are the most sensitive to choices with higher prices. Postestimation results predict that all of the class 3 respondents thought that the size of the increases to their water were important when making their choice decisions. In contrast, 56.4% of Class 2 respondents felt bill size was important when making their choice decisions. As discussed previously, Class 2 respondents have the highest percentage of respondents in full time employment and the highest mean income which may support their relative indifference towards price increases.

Given that proportionally less Class 3 respondents visit the Yare and they visit less frequently, it is unsurprising that these respondents are less concerned about the ecological quality of the water at that site.

The importance of the recreational water quality when making choice decisions also corresponds with respondents' WTP for recreational quality. Recreational quality was important to 75.8% of Class 1 respondents (the class which contains the majority of recreational users), 54.7% of Class 2 respondents (who tend to prefer the ecological quality of the water), whereas only 19% of Class 3 respondents (who tend not to use the river) felt recreational quality was important.

During the survey, respondents were asked how likely it was that improvements to the river water quality would actually happen in the future (see question 20 of the survey in Appendix III). Respondents who thought that improvements were highly unlikely were defined as protestors. These respondents had little faith that management agencies would improve water quality. Consequently, it was thought that these respondents may not have seriously considered choice options which offered High water quality, as they had little faith such improvements could, or would, be achieved. The protest rate across the sample was 13.5%. Class 3 contained the highest proportion of protestors, 19%, which may go some way to explain their unwillingness to select choice options which contained higher prices. Despite these observations, protestors were included in the analysis as their omission did not significantly affect results.

From a policy perspective, the three class model described above, may not best represent the general population, as the data on which it is based over-samples recreational users, river management and water quality experts. A criticism is that Model 7 may produce misleading WTP estimates. The solution to this criticism was to perform a LC analysis which partitioned recreational users and experts away from the respondents drawn from the general public – effectively sampling the general public without interference from the other respondent types. This reanalysis is now discussed.

The starting point to this reanalysis was to generate information criteria values, shown on Table 44, for LC models of differing numbers of classes, to find the optimum number of classes to best represent the latent preferences of respondents drawn from the general public.

Table 44: information criteria values for a partitioned LC model

Number of classes	BIC	AIC	CAIC
2	1706.58	1665.22	1719.58
3	1576.28	1512.65	1596.28
4	1580.98	1495.08	1607.98
5	1600.31	1492.12	1634.31

For the reasons discussed previously, it was found that a 3 class model was the optimal solution.

Partitioned LC models using combinations of the full range of socio-economic variables were explored. One of the best specified models, LC Model 8, described in Table 45, uses the same class membership covariates and choice attribute coefficients as used in CL Model 4 (Appendix V). The results from Model 8, its corresponding WTP estimates, using Low water quality as the baseline (reported in Table 46), and its postestimation results, reported in Table 49, are now discussed.

Model 8, shown in Table 45, contains four separate classes, the first 3 of which contain only respondents drawn from the general public. The fourth, partitioned class, contains only recreational user and expert groups. The most striking initial observation of the 4 class partitioned model is the similarity in composition of the first three classes to the 3 classes in the 3 class model, reported in Table 41. *A priori* expectations, were that by removing experts and users, there would be a significant impact on the marginal WTP estimates within the first three classes. This hasn’t been the case. Instead, the first 3 classes have remained relatively stable. The 4 class partitioned LC model has revealed 3 distinct types of latent preferences among respondents drawn from only the general public. These 3 classes are now reported, with attention paid to the differences to the previous LC model.

Table 45: Model 8, a 4 class partitioned LC model with distance and environmental membership as class membership covariates

	Class 1	Class 2	Class 3	Class 4#
Class membership covariate coefficients				
Intercept	0.575** (0.303)	1.025*** (0.259)	0.499 (0.305)	-2.099*** (0.503)
Environmental memberships (EnvMemberCont)	-0.039 (0.259)	0.104 (0.226)	-1.335** (0.535)	1.270*** (0.281)
Distance (DistanceBin)	-0.011 (0.319)	-0.205 (0.276)	1.197*** (0.414)	-0.980** (0.506)
Choice attribute coefficients				
Price	-0.015*** (0.005)	-0.023*** (0.004)	-0.301*** (0.076)	-0.016*** (0.006)
Medium ecological quality	1.208*** (0.192)	3.588*** (0.299)	0.853** (0.465)	1.466*** (0.231)
High ecological quality	1.713*** (0.291)	5.833*** (0.392)	1.683** (0.753)	2.487*** (0.382)
Medium recreational quality	1.814*** (0.214)	0.464** (0.189)	1.047 (0.825)	1.449*** (0.264)
High recreational quality	2.678*** (0.257)	0.959*** (0.185)	2.628* (1.517)	2.215*** (0.297)
Class probability	0.26	0.53	0.1	0.11
Predicted number of respondents	52	106	20	22
Total respondents = 200				
Total observations = 4800				

*, ** and *** = significance at 10%, 5% and 1% levels. Standard errors in parenthesis
#Users and experts are partitioned into class 4

Class 1 has a class probability of 0.26. 52 of the 200 respondents are estimated to be in this class. Despite the removal of recreational users, Class 1 respondents continue to have the highest likelihood of choosing a choice alternative with improved levels of recreational quality. It was erroneously assumed that because Class 1 in the three class LC model, shown in Table 41, contained the largest proportion of primary and secondary contact users, it was these recreators who were driving the positive recreational water quality coefficients. This is not the case. Class 1 consists of members of the general public who value the recreational quality of the water highly and chose High recreational quality when that option was available.

Class 2 continues to have the largest class probability, 0.53, and is estimated to contain 106 respondents. The number of environmental memberships held, and class members' distance from the river continue to be insignificant class membership covariates.

Despite the removal of users or experts, these respondents continue to have the greatest preference for ecological quality. This class continues to contain the majority of the anglers (12 respondents) who, as we have seen in the CL modelling, care passionately about the ecological quality of the river's water.

Class 3 is least affected by the removal of users and experts, as it contained only one expert, and no users, originally. The class continues to have the smallest class probability, 0.1, and is estimated to contain 20 respondents. The number of environmental memberships held, and the distance Class 3 respondents live from the survey river, continue to be significant class membership covariates. The strength of these variables has increased. As the number of environmental memberships held by respondents increases, respondents continue to be less likely, -1.335** (s.e. 0.281), to be assigned to Class 3. Postestimation results support this coefficient, as it is predicted that Class 3 respondents hold, on average, only 0.1 (s.e. 0.3) memberships, the lowest of the four classes. As distance to the river increases, respondents are, again, significantly more likely, 1.197*** (s.e. 0.414), to be assigned to Class 3. Postestimation results estimate a mean distance from the river to the average Class 3 respondent's home of 10.0km (s.e. 8.6km), 2km further than the distance of the average respondent.

Class 3 respondents continue to have complete and transitive preferences for water quality. These respondents continue to be the most averse to choice options containing increased price. It was previously hypothesised that the average Class 3 respondent had the lowest mean income and preferred their disposable income to be spent on goods and services other than riverine improvements. It then became apparent, from the 3 class model's postestimation results, Table 43, that Class 3 respondents did visit rivers, but tended not to visit the survey river stretch. Within this partitioned model the motivations of Class 3 respondents can be seen more clearly. Although they have the lowest preference for price increases, this doesn't appear to wholly attributable to low income, as, following the partition, the average respondents from classes 1 and 3 now have similar mean incomes, £24,600 and £24,800 respectively. Relatively low income has not prevented the average Class 1 respondent from having clearly defined preferences for recreational water quality. Nevertheless, Class 3 respondents sensitivity to choices with increased price continues to impact on their willingness to pay for water quality at the survey river stretch. It would seem, with reference to the postestimation results on Table 49, that Class 3 respondents are averse to paying higher prices at the Yare because they do not want to pay for improvements at a location they do not use. Class 3 respondents live the highest mean distance from the Yare and they visit infrequently, if at all. However, this class of respondents has the highest frequency of visits to all river sites, visiting 30.2 (s.e. 52.5) times each year compared with 24.4 (s.e. 48.3) for Class 1 and 28.3 (s.e. 55.4) for Class 2.

Class 4, the partitioned class, contains the 22 respondents who identified as recreational users or experts. Referring to the results of CL model 4 (see Appendix V) and the 3 class LC model, it can be seen that the respondents in Class 4 produce confounded results. Some Class 4 respondents (swimmers, the majority of the rowers) prefer recreational water quality, while others (the remainder of the rowers, the majority of the experts) prefer ecological water quality.

The number of environmental memberships held by Class 4 respondents and the distance they live from the survey river are both significant determinants of class membership. On average, these respondents live 3.8km (s.e. 3.4km) from the Yare. This is much closer than the average respondent. They also hold the most environmental memberships, 1.3 (s.e. 0.8), per person.

All choice attribute coefficients are highly significant within Class 4. These respondents are only slightly less likely to choose an option with increased price. This may in part be because these respondents have the highest predicted mean income. Due to their higher disposable income, coupled with their desire for high water quality, the maximum price level of £100 within the choice experiment may not be sufficiently high to adversely impact upon these respondents' choice decisions.

3.5.12 Marginal WTP estimates derived from the 4 class partitioned LC model

WTP derived from the 4 class partitioned LC model are now discussed. This first analysis assesses changes in utility using Low water quality to define V^0, the baseline. This is followed by an analysis using respondents' perceptions of water quality to define V^0.

A priori expectations were that the partitioning of recreational users and experts may cause WTP for the general public to be significantly reduced. This hasn't been the case. WTP for the three classes composed only of the general public in the partitioned model are similar to the WTP estimates generated from the coefficients of the 3 class model. With closer examination there are some interesting observations to be made.

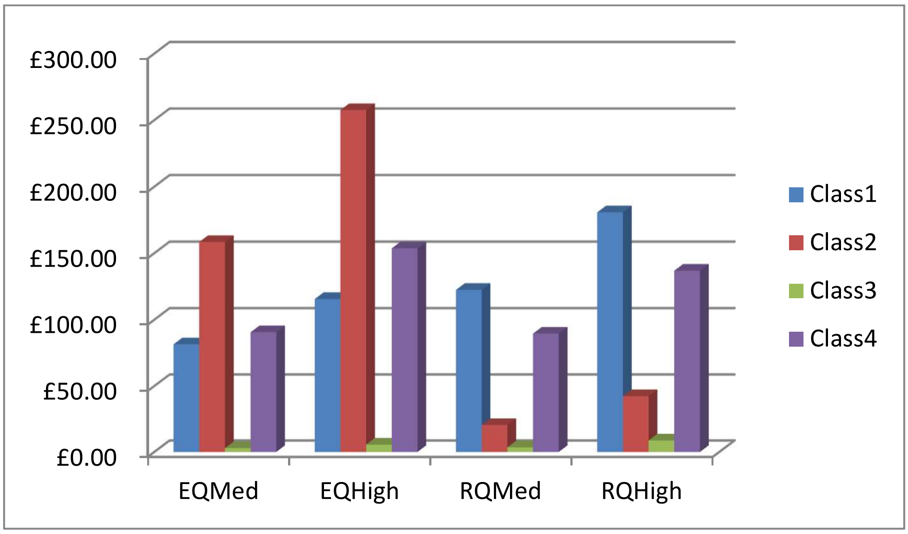
Table 46: marginal WTP estimates derived from the 4 class partitioned LC model

Water quality type	Class 1		Class 2		Class 3		Class 4	
	WTP (£)	s.e. (£)	WTP (£)	s.e. (£)	WTP (£)	s.e. (£)	WTP (£)	s.e. (£)
Medium ecological	£81.54	£29.13	£158.45	£26.63	£2.84	£1.66	£90.59	£32.16
High ecological	£115.58	£40.87	£257.58	£40.10	£5.60	£2.76	£153.69	£52.15
Medium recreational	£122.44	£43.27	£20.48	£9.66	£3.48	£2.12	£89.58	£31.56
High recreational	£180.72	£60.06	£42.33	£9.88	£8.74	£3.32	£136.87	£43.56

(WTP £, per household, per year for the 20km survey river stretch)

Class 1 respondents continue to have the highest WTP for recreational quality. The removal of recreational users and experts from this class has caused WTP for Medium recreational quality to fall by 14.4% and WTP for High recreational quality to fall by 13.8% compared with the values predicted in the 3 class model. Class 1 respondents' WTP for ecological quality continues to remain high, relative to the other classes. Class 2 respondents' WTP for ecological water quality has risen slightly, despite the removal of recreational users and experts. Class 3 respondents continue to have the lowest WTP for water quality of the three classes. On Table 46, we see that their WTP has barely changed. Class 4 respondents, due to their confounded preferences for water quality, have relatively high WTP for all types of water quality, as shown in Figure 30, below.

Figure 30: visual representation of WTP derived from the 4 class partitioned LC model



The heterogeneous differences in preferences across classes are clearly shown in Figure 30. Class 1 respondents prefer, and are willing to pay for recreational quality. Class 2 respondents are willing to pay for ecological quality and Class 3 respondents have consistently low WTP across the two water quality attributes. The WTP estimates for the 4 class partitioned model, reported in Table 46, avoid the criticism that the estimates are biased by the overrepresentation of user and expert groups in the sample. However, by calculating the above WTP estimates as marginal changes from Low water quality, it may be that respondents' WTP is overestimated. Table 47 shows the results of a reanalysis of WTP with the baseline water quality set to the level of respondents' perceived water quality.

Table 47: WTP estimates derived from the 4 class partitioned LC model, based on respondents' perceptions of water quality

	Class 1		Class 2		Class 3		Class 4	
	WTP (£, per household, per year for the 20km survey river stretch)							
Water quality type	**V^0=Low water quality**	**V^0 = respondents' perception**	**V^0 = Low water quality**	**V^0 = respondents' perception**	**V^0 = Low water quality**	**V^0 = respondents' perception**	**V^0 = Low water quality**	**V^0 = respondents' perception**
Medium ecological	£81.54	£1.57	£158.45	£4.48	£2.84	£0.00	£90.59	£0.00
High ecological	£115.58	£29.79	£257.58	£92.53	£5.60	£2.31	£153.69	£54.44
Medium recreational	£122.44	£4.71	£20.48	£0.39	£3.48	£0.00	£89.58	£0.00
High recreational	£180.72	£49.54	£42.33	£19.14	£8.74	£5.00	£136.87	£30.09

Estimates of WTP where V^0 = low water quality transcribed from Table 46 for ease of comparison.

As with the estimates of perceptions-based WTP shown on Table 38, we find that respondents' WTP for Medium water quality levels are greatly reduced. Respondents within the smaller classes, classes 3 and 4, have zero WTP for Medium ecological and Medium recreational water quality. These zero values are due to those respondents' perceptions of current water quality. None of the respondents partitioned into Class 4 (e.g. the rowers, swimmers and experts, see Table 32) thought the water quality at the survey river was lower than Medium, therefore they would receive no marginal benefit from the future water quality remaining at Medium. No Class 4 visitors thought the water quality was lower than Medium, and Class 4 non-visitors' (who were unable to provide perceptions data) had their perceptions set to Medium, the modal level of perceived water quality across the sample. WTP for the higher levels of ecological and recreational water quality are much reduced: those respondents who perceive the current water quality as High receive no benefit from the future water quality remaining, or changing to High. Those respondents who currently perceive the water quality to be Medium receive the marginal benefit from a change from Medium to High. The patterns within the results of the perceptions-based analysis of LC Model 8 remain as previously discussed: Class 1 respondents prefer recreational improvements, Class 2 respondents hold a preference for improved ecological water quality and Class 3 respondents have relatively low WTP values for either form of water quality improvement.

The perceptions-based WTP values reported on Table 47 can be aggregated to provide averaged WTP for water quality attributes. These are shown on Table 48. Average WTP values for the whole sample differ marginally from the averaged WTP of classes 1-3 (which excludes rowers, swimmers and experts). The slight differences can be accounted for by rowers and swimmers holding a higher preference for the recreational, rather than ecological, quality of rivers. Experts tend to have higher perceptions of current water quality (Table 32), which serves to reduce the utility of Medium water quality levels.

Table 48: averaged perceptions-based WTP derived from LC Model 8

	WTP (£, per household, per year for the 20km survey river stretch)	
Water quality type	**Classes 1-3 only**	**Whole sample**
Medium ecological quality	£3.13	£2.78
High ecological quality	£64.06	£63.01
Medium recreational quality	£1.61	£1.43
High recreational quality	£26.43	£26.83

The remaining postestimation results derived from the 4 class partitioned model are now reported.

Table 49: post estimation results for the 4 class partitioned LC model

	Class 1	Class 2	Class 3	Class 4	Whole Sample
Class probability	.26	.53	.1	.11	1.0
Number of respondents	52	106	20	22	200
Mean age	57.4 (18.9)	50.0 (17.6)	49.8 (20.7)	40.8 (13.9)	50.9 (18.6)
Gender (% male)	40.4	44.3	45.0	50.0	44.0
Employment, income, education and environmental affiliation					
Employed full time (%)	26.9	48.1	45.0	68.2	44.5
Employed in environmental occupations (%)	1.9	2.8	0.0	59.1	8.5
mean income (£1000s)	24.6 (16.1)	28.2 (21.1)	24.8 (18.8)	43.5 (28.3)	28.6 (21.4)
% respondents with degree level education or higher	36.5	50.0	25.0	68.2	46.0
mean number of environmental memberships per respondent	0.4 (0.7)	0.5 (0.6)	0.1 (0.3)	1.3 (0.8)	0.5 (0.7)
Predicted number of each respondent type in class					
Anglers	0	12	0	4	16
Rowers, swimmers and experts	0	0	0	22	22
General public (door to door)	38	86	15	0	139
General public (Whitlingham)	14	20	5	0	39
Distance and trip information					
Mean distance respondent lives from the Yare (km)	9.4 (11.1)	7.8 (11.2)	10.0 (8.6)	3.8 (3.4)	8.0 (10.5)
Mean number of all river trips taken by respondent in the last year	24.4 (48.3)	28.3 (55.4)	30.2 (52.5)	133.0 (87.4)	39.0 (66.7)
Mean number of activities at the most visited site in the last year	1.4 (1.3)	2.0 (1.6)	1.5 (1.2)	2.8 (0.9)	1.9 (1.5)
% of respondents visiting the Yare in the last year	51.9	55.7	30.0	100.0	57.0
Mean number of Yare trips taken by respondents in the last year	15.6 (37.6)	16.9 (49.3)	3.8 (11.4)	115.2 (77.1)	26.1 (57.6)
Mean number of activities at the Yare in the last year	1.8 (1.3)	1.4 (1.5)	0.7 (1.2)	2.5 (0.9)	1.4 (1.4)
Mean number of Yare trips in the next year if water quality improvements are made	17.0 (32.2)	33.8 (72.7)	5.1 (11.3)	137.0 (89.3)	37.9 (72.6)
Importance of issues when making choice decisions (% respondents)					
The size of the bill	67.3	55.7	100.0	45.5	62.0
The distance to any proposed improvement	42.3	42.5	50.0	50.0	44.0
The ecological quality of the river	98.1	99.1	65.0	100.0	95.5
The recreational quality of the river	67.3	52.8	20.0	90.9	57.5
Protesting?	15.4	13.2	15.0	9.1	13.5

Standard errors in parenthesis

Class 1 has seen a substantial reduction in the number of respondents visiting, from 61.3% to 51.9%. Class 2 has seen a slight reduction from 59.0% to 55.7% of respondents visiting. These reductions in visit frequency are due to the partitioning of the high frequency visitors into class 4, all of whom visited in the last year. The percentage of Class 3 respondents who visit the river has fallen slightly to 30%. The mean number of trips to the Yare taken by Class 3 respondents is almost unchanged at 3.8 (s.e. 11.4). By removing high frequency recreational visitors, the mean number of trips taken by Class 1 respondents has fallen substantially from 40.0 (s.e. 67.7) to 15.6 (s.e. 37.6) trips. Class 2 respondents visited 16.9 (s.e. 49.3) times. Class 4 respondents visited just under three times a week, at an average of 115.2 (s.e. 77.1) times each year.

Respondents from classes 1 and 3 do not appear to be motivated to visit more frequently by the promise of high quality water. Interestingly, despite the removal from the class of users and experts, the remaining respondents in Class 2 stated that they would visit far more frequently if water quality was high. Their proposed number of trips rises from 16.9 (s.e. 49.3) to 33.8 (s.e. 72.7) trips in the future. This may be due to a latent demand for improved opportunities for ecologically focussed activities, e.g. nature watching or photography. Class 4 respondents also stated that they would visit more often, rising from 115.2 (s.e. 77.1) to 137.0 (s.e. 89.3) trips in the future. As discussed, this rise may not be due to rowers visiting more frequently, as rowing club members typically visit 3 or more times at present. Part of this increase may be attributable to improved recreational opportunities for swimmers. In interview several of the experts professed a desire to take more trips for recreation and social activities if water quality was guaranteed to be high.

Data on the importance of the different issues when making choice decisions continues to reinforce the results of the LC analysis. The mean income for Class 1 has dropped substantially. It now has the lowest mean income, £24,600 (s.e. £16,100), of the four classes. Consequently, Class 1 respondents have become more sensitive to the issue of bill size, with 67.3% feeling it was very important. Interestingly, Class 3 respondents, with a slightly higher mean income (£24,800), continue to be the most sensitive to the issue of increased price. All Class 3 respondents thought bill size was very important. 55.7% of Class 2 and 45.5% of

Class 4 respondents felt bill size was important when making their choice decisions.

There is not a great deal of variation between the four classes on the importance of distance from the survey river stretch to their home. Equal proportions of Class 3 and Class 4 respondents stated that the issue was important, despite a large difference in the mean distance these two classes live from the river stretch. It may be that the two classes felt the issue to be important for opposing reasons. Class 3 respondents, living an average of 10.0km (s.e. 8.6) from the river, may have felt the issue important because the river was relatively distant. Class 4 respondents, living an average of 3.8km (s.e. 3.4) away from the river, may have felt the issue was important because the river was relatively close. The usefulness of this variable is reduced due to this ambiguity.

The three classes (1, 2 and 4) of respondents who visit the river relatively frequently are in agreement on the importance of the ecological water quality. Over 98% of these respondents thought the issue was important when making their choice decisions. In contrast, only 65% of Class 3 respondents felt the issue was important. This may not be because they feel that the ecological quality of rivers is unimportant per se, as they visit other rivers frequently and have positive coefficients for ecological water quality within the LC modelling. It appears more likely that they place less importance on the ecological quality of a river which they rarely visit.

Recreational quality was important to 67.3% of Class 1 respondents (the class which has the highest WTP for recreational quality), 54.7% of Class 2 respondents (who tend to prefer the ecological quality of the water) and 90.9% of Class 4, which has a large proportion of respondents who frequently use the river for recreation. Only 20% of Class 3 respondents thought that the recreational water quality was important. Again, this may be because they rarely visit the site and distance was an extremely important factor to them when making their choice decisions.

Class 4 has the lowest proportion of respondents, 9.1%, defined as protestors, e.g. respondents who thought that improvements were highly unlikely. There is little difference in the protest rate across the other three classes, in which the protest rate ranges from 13.2% for Class 3 to 15.4% for Class 1. As discussed previously, protestors were included in the analysis as their omission did not significantly affect results.

3.6 Discussion

Two groups of complimentary models examine different aspects of the same data to find solutions to the research questions. Comparing and contrasting the results of the different types of models reveals significant answers to the questions of who cares about river water quality and by how much.

The analysis uses choice experiment methods to disaggregate the value of recreational and ecological characteristics of river water quality. This is both feasible and necessary since, contrary to previous econometric valuation practices (Bateman et al., 2011; Ferrini et al., 2014), these facets of water quality can be completely uncorrelated. The CE featured an efficient experimental design and the sample included both non-visitors and a wide variety of recreational users. Face-to-face surveys presented respondents with choices across a range of future water quality scenarios, differentiated in terms of the survey river's ecological and recreational quality attributes and hypothetical remediation costs. CL and LC analyses identified a number of preference predictors including respondents' spatial relationship to rivers and their socio-economic characteristics. The willingness to pay measures derived from CL and LC models revealed clear differences in preferences between respondent groups.

Of the CL models, Model 5 is most suited to providing an overview of the water quality preferences of less specialised respondents (i.e. non-visitors and casual visitors from the general public) and the more specialised respondent groups (i.e. rowers and swimmers, and anglers). The less specialised respondents held higher values for improved ecological quality, rather than recreational enhancements. Similar preference orderings, but at higher levels of WTP, were revealed by anglers. However, other users, such as swimmers and rowers, prioritised recreational over ecological improvements. Three other preference predictors, environmental memberships held, distance and income, were identified. There were positive correlations between respondents' WTP and their income and also positive correlations between respondents' WTP and the number of environmental memberships they held. A significant distance decay in values away from the sites of any proposed investment was also observed. It may be argued that CL Model 5 presents respondents as one dimensional beings, e.g. anglers preferring EQ and rowers and swimmers preferring RQ. This outcome arises from parameterising the model in the most efficient way, to best represent the preferences of those user groups. It would be strange if those relationships did not conform to *a priori* expectations in the way that they do: one would then have to consider if the results from the survey sample were random. Another criticism is that Model 5 is heavily parameterised to arrive at its results. Provencher and Bishop (2004) caution us that heavily parameterised models tend not to be robust. To avoid this criticism, CL Model 6 provides a simple, robust and parsimonious solution which suggests that the average respondent holds greater preferences for ecological water quality improvements. In addition to these weaknesses, the CL models fail to meet critical assumptions underpinning RUT. Several of the parameters of model 6 were found to exhibit heterogeneous variance, shown on table 39. The CL model structure also failed to control for intra-respondent panel data, but instead treated all unobserved factors across observations as independent and unique. These shortcomings are acknowledged. The use of CL modelling, as a starting point of the data analyses, is defended because CL is a simple method by which trends within data can be examined.

The rudimentary findings of the CL models are highlighted by the results from the LC models. These unpick the simple characterisation of respondents to reveal three statistically distinct types of respondent defined by their latent preferences. While the partitioned LC model confirms that the majority of the general public have a preference for ecological quality, it also reveals a sizeable minority which hold a preference for recreational water quality, and a third group, which holds relatively low values for either form of river water quality improvements at the survey site.

Of the two LC models, it can be argued that the preferred model is the four class partitioned model. From a policy perspective, the first three classes are more representative of the preferences of the general public, as recreational users and experts are omitted from these classes. Their omission is acceptable given that the results of CL Model 5 confirms their water quality preferences and because the postestimation results (Table 43) from the 3 class LC model suggests the classes to which these respondents may have been ascribed, had they not been excluded.

It appears that Class 1 respondents within the partitioned LC model, despite having a relatively high WTP for improved recreational water quality, do not actually want to use potentially improved recreational water. Although they have a high WTP for improved recreational water quality, the number of their proposed future trips barely increases if recreational quality is guaranteed to be high. This begs the question: why do they have a clear preferences for improved recreational water quality? It is unlikely that these respondents are 'yea-saying' as the CE format presents repeated options with assorted combinations of choice attributes to the respondent. It may be that Class 1 respondents are anthropic in their outlook. Although they have a high regard for both aspects of water quality, they appear to hold preferences which improve the environmental safety for other humans, even if they do not directly benefit themselves. Anthropic behaviour does occur. Hanley et al. (2003) found that despite the majority of respondents preferring improved water quality at Scottish beaches, those respondents would not start swimming as a result of improved water quality.

Within the 4 class partitioned LC model we find that the majority of respondents prefer, and have a higher willingness to pay for, ecological water quality. We also find that there appears to be a strong latent demand by that majority for an increased enjoyment of ecologically based leisure activities, if ecological improvements are made.

Within the LC modelling a small class of respondents is revealed, the majority of which tend not to visit or participate in river based recreational activities at the Yare. These respondents are apathetic in their WTP for ecological or recreational water quality improvements at the Yare. However, the postestimation results suggest that these respondents typically take their river based recreation at substitute locations. Given that these respondents frequently enjoy trips to other rivers, it is likely that these respondents may prefer river water quality improvements at sites closer, and more convenient, to their homes.

3.7 Conclusions

The legislative imperative of the WFD requires improvements to river water quality. Poor river water quality imposes costs onto society, as do the remedial measures to reduce that river pollution. To minimise the incidence of derogations, relaxing the requirement to achieve ‘good ecological status’ on grounds of excessive cost, it is necessary that the benefits of reducing pollution are comprehensively assessed. To aid cost/benefit assessments, the present research has sought to disentangle and examine the relationships between the different sources of non-market values, thereby allowing decision makers to understand the consequences of adopting alternative investment strategies. These strategies may favour either ecological or recreational improvements, or a mix of the same to improve benefits. To improve our understanding of the consequences of alternative strategies, this research has used attribute based valuation methods and a novel survey design to analyse the way in which individuals value the recreational and environmental functions of rivers.

With regard to the question of who cares about river water quality, results were found to be stable over the alternative choice models estimated. These models identified significant heterogeneity in water quality preferences across the different respondent types. Clearly the answer to “who cares?” depends on who is being asked, and for what reason. Previous research revealed that recreational water users are willing to pay relatively more to secure higher recreational water quality. This expectation is confirmed within these results. What was more unexpected was the heterogeneity across preferences found within the general population, composed primarily of non- and infrequent visitors. Three distinct respondent types were revealed within the general public. The majority hold a preference for enhanced ecological quality, a minority are motivated by recreational quality improvements and a yet smaller proportion typically prefer to visit substitute river venues and are ambivalent about the water quality at the Yare.

Topography, human population density, land use type and land use intensity causes spatial differences in pollution types, pollution vectors and pollution concentrations across UK rivers (Hampson et al., 2010; Haygarth et al., 2005). Previous research has shown that it is technically infeasible and prohibitively expensive for all UK rivers to be brought to 'good ecological status' within the near future (Defra, 2008c; Wither et al., 2005). Derogations on grounds of unacceptable financial costs and technical infeasibility will be necessary.

This research adds to the literature by further demonstrating that positive non-market benefits are likely to accrue from remediation schemes. It also shows that the non-market benefits which may accrue from different types of water quality improvements are nuanced in terms of their environmental impacts, their potential beneficiaries and, by inference, their overall value and policy implications. It is important that research outcomes demonstrate to policymakers that different remedial measures (aimed at either ecological or microbial water quality improvements) may trigger entirely different benefits and differing levels of market and non-market values. Decision makers need to be able to understand these differences and be able to access simple, quantifiable data in order to maximise the effectiveness of the limited resources available for improvements.

So, what are the implications and how should the policy and management community react? As the costs of pollution and the benefits of remediating that pollution are unequally distributed, it is simply not cost effective to direct scarce financial resources at pollution remediation equally across all rivers. It appears that the policy and management communities must be pragmatic, accept that not all riverine pollution can be solved in the short term, and adopt a focussed and targeted approach to pollution remediation schemes. Close examination of the net benefits of different patterns of investment, at different locations, will ensure that the allocation of scarce resources yield the maximum net benefits across locations. Areas of high net benefit will vary spatially, not only due to the characteristics of riverine pollutants, but also because beneficiaries are unevenly distributed. Using this approach, the policy and management communities should well be advised to focus on the highest value areas for immediate attention. There are solid reasons why this should be and the foundations for this approach are already in place. Successive UK governments have adopted and maintained evidence based decision-making processes as their leitmotif. Decisions require the consistent and effective monitoring of pollution in watercourses and the accurate calculation of the costs of remediation schemes. Governmental regulatory organisations, such as the Environment Agency and Defra, are well placed to calculate and assess damage to the environment (and costs to society) arising from pollution, the benefits of alleviating that pollution, and, in collaboration with the agricultural sector and the privately owned utility companies, the costs of schemes to remediate identified pollution sources. Economists will find that the net costs arising from river water quality improvements are bound to change over time. The location of the pollution source yielding the highest net benefit, once remediated, may not be the location of the source yielding the next best net benefit, *ad infinitum*.

The benefits of pollution remediation are relatively uncertain given the elusive nature of non-market values. Precise calculations of the benefits arising from pollution remediation appear to be location specific, requiring detailed and costly research and analysis to reveal. What is being increasing confirmed and revealed by recent work, e.g. Metcalfe et al., (2012), the findings of which are reinforced by the research presented here, are the spatial conditions and patterns pertaining to the non-market benefits which may be available following pollution remediation. This research finds that the location of the pollution remediation, the type of remediation, the intended beneficiaries and the distance to those beneficiaries, all significantly affect benefit values.

The results of this research suggest that, with regard to microbiological river water quality, the optimal remediation solutions may be to supply a relatively few numbers of high quality recreational sites close to population centres across the UK. CL models 5 and 6 found significant distance decay in benefit values. CL Model 4 found distance to have a step function, with a preference parameter at 8km from the river. LC models 7 and 8 found distance to be a significant determinant of class membership. These results corroborate the findings of Metcalfe et al. (2012), that the highest benefits will be obtained from improvements undertaken relatively close to densely populated settlements. LC Model 8 suggests that Class 3 respondents hold relatively low preferences for improvements at the survey river as they may be motivated by preferences for improvements at substitute sites which they visit more frequently (see Table 49). Given that the focus of this study was to develop a method to disentangle preferences for ecological and recreational water attributes, it did not include analyses of respondents' preferences for substitute sites, benefits transferability, framing or respondents' insensitivity to scope. With this in mind, it is not prudent to extrapolate the results of the models reported here to estimate preferences with accuracy spatially elsewhere. These issues are discussed further below. More generally, within this case study area the location which would appear to yield the highest net recreational benefits would be the area surrounding the centre of Norwich. This would take full advantage of the close proximity of the largest proportion of potential recipients.. On a national scale, the locations which may yield the highest net benefits from recreational improvements could include sites such as the River Trent at the location of the National Water Sports Centre, which, as discussed previously, suffers financial losses and disruption due to poor water quality. Other river locations with substantial numbers of recreational users, such as lengthy stretches along the River Thames upstream of Westminster, may also yield high use values.

A targeted approach to remediation schemes may help to minimise the tangible financial losses to boating, swimming and other river recreation clubs and the losses due to ill-health. Improved recreational water quality may help to promote recreational use of rivers leading to, among other benefits, increased revenues once rivers have gained reliable reputations as safe venues due to consistently high water quality.

The major motivation for the majority of the respondents within this research concern the ecological quality of river water. With this in mind we should expect proportionately more resources to be targeted towards improving the ecological quality of our rivers. Bateman et al. (2006a) speculate that we might expect to find less distance decay in pure non-use ecological values. Within CL Model 4 (Table 51, Appendix V) in the present study there is a distance threshold in ecological values, e.g. baseline respondents living further than 8km from the river are WTP 38.8% less for high ecological water quality. It is important to note that this reduced benefit value does not necessarily indicate a distance decay in pure non-use values. Within this analysis non-users (who were not intentionally sampled) are conflated with non-visitors (i.e. those who did not visit the survey river in the year prior to the study). This issue is discussed in further detail below.

Given a targeted approach to ecological pollution remediation by management authorities, within the geographical area of the present study, we should perhaps expect the majority of ecological quality improvements to be made close to, but upstream of, the major urban area, Norwich. The reasons for this assumption are twofold. This study finds that ecological improvements close to urban areas adjacent to the river would receive the largest numbers of beneficiaries and, accordingly, higher levels of ecological use benefit values. Secondly, areas downstream of Norwich may require more costly pollution remediation measures due to pollution derived from human effluent emissions into the Yare from the wastewater treatment works at Postwick, downstream of the survey river stretch. Previous research (Hampson et al., 2010) found that the water quality in upland headwaters tends to be relatively high. Although remediation costs in those areas may be low, the number of beneficiaries in those relatively unpopulated and inaccessible areas is also low, which generates low use benefit values.

What these last few paragraphs have highlighted is the emphasis which must be made on spatial location in planning decisions. The benefits of remediation vary substantially across space, such that there is little point in spending scarce resources on remediation measures in locations where there are little or no net benefits. We find that respondents have preferences not only intricately related to the nature of the good, but also related to its distance from them and related to the availability of substitute goods. For example, within this research, respondents at Wroxham or Bungay have little or no incentive to visit the Yare as they have high quality recreational sites available locally. A close examination of the effects of distance decay on benefit values, and the availability of substitute venues, will better enable planners to examine the benefits arising in different locations and help transform scenario analyses into optimal analyses for different types of investment and remediation schemes.

In recent years the field of environmental economics has increasingly sought to develop spatially transferable models. As we have seen in the preceding chapters, transferable results are advantageous in terms of reducing not only the financial and temporal analytical costs, but also because such models provide insights into the optimal locations in which to undertake more comprehensive assessments of the effects of environmental change. The transferability of predictive models is a contentious issue within water valuation research given the frequently disparate nature of highly localised factors. Different areas are subject to different pressures, in terms of pollution sources, concentrations and vectors. These same areas differ in terms of the spatial distributions and densities of the potential beneficiaries arising from remedial measures. This complex interplay between costs and benefits and the nature of potential remedial measures, has not been fully addressed by this research. The conclusion to this thesis explores the limitations of the research and examines ways in which this work may be improved and extended in order to produce more robust, transferable outcomes.

3.7.1 Limitations and potential improvements to the research within Chapter 3

As mentioned above, it may not be prudent to extrapolate the WTP estimates obtained by this research to calculate total welfare estimates for ecological and recreational improvements held by the wider population. There are several reasons for this.

At worst case, WTP values are potentially too high to be meaningful. Although the perceptions-based estimates of annual household WTP (i.e. High ecological quality = £55.46/household/year and High recreational quality = £21.03/household/year from CL Model 6 (Table 38), and High ecological quality = £63.01-64.06/household/year and High recreational quality = £26-43-26.83/household/year from LC Model 8 (Table 48)) produce values not too dissimilar from Metcalfe et al. (2012) (i.e. £66.40-76.20/household/year/up to 2015 for a bundle of attributes (Table 24), Hanley et al. (2006) (i.e. £28.57-42.99/household/year for individual attributes at the Clyde (Table 24)), or Doherty et al. (2012) (i.e. €129/household/year for a bundle of waterbody attributes at 'good' quality), the scale of the improvements proposed within this study differs markedly. Within this CE respondents were responding to potential changes to a 20km stretch of the River Yare. Respondents in the Hanley et al. study were providing welfare estimates for improvements to the whole of the River Clyde, whereas in Metcalfe et al., respondents provided welfare estimates for either improvements to rivers in their local area or, as in Doherty et al., rivers nationally. Clearly the scale of proposed waterbody improvements is a limiting factor to the reliability of WTP values reported here: it isn't reasonable to scale WTP values for a 20km stretch up to provide estimates of WTP for the thousands of kilometres of rivers nationally. Within the present study a defined stretch of a single survey river was chosen to examine potential effects of distance decay on respondents' utility for ecological or recreational or both water quality attributes. A stretch-based approach was also used to aid compatibility with the stretch-based approach used in the ChREAM project. It is entirely possible that if the wording of the survey used here (see Appendix III) was framed differently (e.g. prompted respondents to consider changes to all rivers, rather than to a 20km stretch of a single river), more realistic, transferable, WTP estimates may have been obtained. Metcalfe et al. accounted for scale. They observed WTP of £66.40/household/year for 95% of local rivers to be improved to 'good' ecological status and £76.20/household/year for 95% of national rivers to be improved to the same standard.

There are other reasons why the WTP values reported here may be unrepresentative. The most obvious reason is that visitors to the survey site and recreational users (i.e. swimmers and rowers) were over-sampled. This was by design: to reveal and examine a suite of preferences for different aspects of river water quality and recreational functions of the survey river. Issues of sample weighting and sample representativeness are discussed further below.

Combined WTP values for both water quality attributes to be brought to 'good' ecological status in the present analysis range from £76.49/household/year (CL Model 6, Table 38) to £90.49/household/year (LC Model 8, whole sample, Table 48). These estimates exceed the value of the bundle of water quality attributes at the River Wear (£36.93/household/year) but are less than the aggregated welfare of improvements at the river Clyde (£110.26/household/year) (Hanley et al. (2006), Table 24). Hanley et al. found that, despite the same survey design being employed in both places and despite both rivers being superficially similar, benefits transfer tests were rejected as preferences differed significantly across the two studies, potentially due to differences in unobserved psychological characteristics and/or cultural values. Socio-economic and quantitative demographic characteristics aside, it is not possible to say whether the latent psychological and cultural preferences for river quality held by the present sample of respondents is representative of those held by people in other areas within the

UK. The increased use of quantified qualitative research methods in environmental economics, such as Q Methodology (see for example Brown (1980), or Watts and Stenner (2012)), may aid the characterisation of respondents' psychological preferences. A mixed methods[27] approach towards incorporating such qualitative data into economic valuation (e.g. Aldrich et al. (2007) or Hunter et al. (2012)), may help improve our understanding of how populations differ culturally and also help us to identify the characteristics of a psychologically representative sample.

Although the number of attributes used in this study was deliberately kept low, the DCE multiple-attribute design format used here may introduce error in welfare estimates. Previous studies such as Foster and Mourato (2003) and Hanley et al. (1998) have found that the DCE multiple-attribute format can produce higher values for a package of improvements rather than a CV focused on a single aggregated policy change. Metcalfe et al. (2012) suggest that this may be a function of respondents placing less weight on the cost attribute when it is varied simultaneously with other attribute.

During the experimental design the range of price levels was intentionally kept relatively low. There are several reasons for this: it would be unrealistic for respondents' annual water bills to rise dramatically (e.g. there would be a public outcry if bills were to double or triple for the purpose of river pollution remediation) and such extreme increases, within the experimental setting, may have led to attribute non-attendance (i.e. respondents ignoring the ecological and recreational attributes to focus on selecting the lowest available price alternative) or hypothetical bias (i.e. respondents rejecting the premise of the choice experiment as they believed it to unrealistic, and, consequently, providing ill-considered answers). Naturally there will be respondents for whom increased price is a considerable burden and, for those respondents, a low range of prices will be sufficient to capture their sensitivity to price. For other respondents, for a variety of reasons (e.g. higher income, greater wealth, etc.), increased price may not be so constraining: an increase of £100 per year (£1.92 per week) may be

[27] Where mixed methods is defined as the "collection or analysis of both quantitative and qualitative data in a single study in which the data are collected concurrently, both are given priority, and involve the integration of the data at one or more stages in the process of research." (Cresswell et al., 2003).

insignificant. The price vector used in this research may have been too low to fully capture their sensitivity to price. A higher maximum value, e.g., £200, as used by Metcalfe et al. (2012), may be more appropriate.

A further criticism of the selection of price levels is that a price level of zero is unrealistic (e.g. all improvements are costly). There are a number of experimental reasons why we would want to have a zero cost in the choice alternative attribute bundle. Firstly, a zero price enables a sense check to see if respondents behave in an illogical manner (e.g. it would be worrying if more people rejected, rather than selected, a zero price), or check if respondents are protesting. We may also want to see the possible shape of the trade-off between money and the different improvements, to allow for a non-linear relationship: if we only had prices of, for example, £3, £5 or £10, the 'real' price may be outside of that range (e.g. £1, or £12) and we may then need to extrapolate outside of the range of the data to calculate the relationship between benefits and costs, e.g. calculate what respondents' WTP may be for different levels of the water quality attributes.

Another problem of a zero cost may be that it may promote strategic behaviour, where the respondent, realising that on one choice occasion was able to get a bundle of goods for zero, begins to act strategically and rejects non-zero, or the higher price of the alternatives, in other choice sets. This suggests a failure of the design to be incentive compatible, e.g. respondents may provide misleading answers. However, this argument can occur at any price level, e.g. the next lowest price level. As soon as respondents realise that that could get an improvement for, say, £3, they may then begin to behave strategically and reject price options that exceed £3.

We may consider if having a zero cost undermines the credibility of the experiment. For example, if a zero cost introduces hypothetical bias, where respondents disbelieve the premise that improvements, or a bundle of attributes, can be provided at zero cost. However, from a purely experimental perspective, having a zero cost is perfectly reasonable: an improvement in the recreational attribute may be accompanied by both a deterioration in the ecological attribute and by a zero price increase - the implication is that an increased level of one service is offset by a lower service elsewhere with no net increase in price.

During post-survey interviews, a small proportion of respondents reported lexicographic preferences or attribute non-attendance, depending on their personal preferences. Lexicographic behaviour is non-compensatory: respondents do not consider all attributes but instead adopt an attribute processing strategy to ease their decision-making, such as always choosing the cheapest alternative (Campbell and Lorimer, 2009).

Overall, all levels of all attributes within both the CL and LC results were significant, complete and transitive, and respondents' preferences for attribute levels appear consistent (Tables 29-31), which suggests a reasonably low level of attribute non-attendance. At a group level, we find that certain respondents, e.g. swimmers, are not prepared to trade off reductions in recreational water quality, suggestive of a lexicographic preference for water quality within this group: several of the swimmers reported selecting only the choice options which maximised recreational water quality. Other respondents reported that they were motivated by the smallest increases in the cost attribute.

The above examples may be considered rational circumstances in which value estimates demonstrate attribute non-attendance, lexicographic preferences or responsiveness to scope (Heberlein et al., 2005; Rollins and Lyke, 1998), that is, such preferences reasonably reflect the manifestation of underlying preferences that are truly lexicographic (Atkinson et al., 2000). Consequently, such 'irrational' responses were retained (Lancsar and Louviere, 2006). The issues of attribute non-attendance and lexicographic preferences could be further analysed but, within this analysis, as in Campbell et al. (2011), a latent class framework that defines classes based on rules that recognise the non-attendance to one or more attributes, was preferred.

Within the survey design, to minimise the complexity of choice options, alternative river sites (or substitute activities) were not offered to respondents within the choice options. Consequently, there was not sufficient information collected to determine variations in insensitivity to scope. While scope is an important issue with a contentious literature (e.g. Diamond and Hausman 1994; Hausman, 2012; Heberlein et al. 2005; Powe and Bateman 2004; Veisten et al., 2004), the focus within the present research was on disentangling and examining preferences for ecological and recreational water quality for a single site, not on measuring respondents' preferences for the survey river against substitute river sites or against alternative recreational activities. The availability of, and respondents' preferences for, alternative recreational sites undoubtedly influenced respondents' choice decisions, particularly among respondents who lived at greater distances from the Yare: during the post-survey interviews it was found that respondents living closer to alternative high quality sites, e.g. the Wroxham Broads or the River Waveney at Beccles, preferred to visit those sites, rather than travel to the Yare.

A relatively small sample has been used in this research. There are a number of reason for this which are outlined from p.189. A small sample is probably less representative of the overall population. There is an argument that the sample should be weighted to try to make it more representative of society, e.g. add weight to population groups underrepresented within the sample, so that such groups have more weight within the analysis, and, therefore, the analysis is more representative of the overall population. With a sample of 200, some of the weights could be quite strong because in some respects the profile of the sample are quite different from the wider population (e.g. different ethnicities are underrepresented within the sample). Weighting could have been done if the purpose of the analysis was to produce data representative of the UK population. However, the purpose of this work was not to produce values suitable for CBA for national decision making. What has been done is the development of a method that allows us to distinguish between ecological and recreational preferences, in a format that can be adapted for use within the ChREAM project, and to assess those differences within the present sample. The results from this research, with further analysis and with weighting applied, can be applied to inform the ChREAM study (in which there is currently no way to disentangle preferences for ecological and recreational attributes). The results obtained using the current methodology are informative for several reasons. The significant impact of distance (in CL Model 4) on respondents' preferences for high recreational quality suggests that it is likely that recreational values are relatively spatially confined (and ecological values relatively less spatially confined): if there are rivers close to areas where demand for recreational use is high, preferences for recreational quality are likely to be higher whereas for more distant rivers preferences for recreational use are likely to fall away. What could not be distinguished from the ChREAM results, is whether recreational values fall with distance: if the preferences motivating ChREAM respondents are predominantly recreational it would suggest that there is little value in improving remote rivers. However, in common with Doherty et al. (2014), this work shows that ecological values are typically higher than recreational values. Metcalfe et al. (2012) suggest that the relative values of the non-cost attributes derived from a DCE can be considered reliable, but that total values, which depend on cost, may be biased upward. Further analysis of the ChREAM data should be taking account of the results presented here. It is accepted that the unweighted values reported here should not be used in their raw form within CBA.

Previous research has shown that, with an efficient experimental design, effective results are possible with small samples (e.g. Hanley et al., 1998, 2005, 2006; Rolfe et al., 2000; Wattage et al., 2005). However, a larger sample for any future extension of this research is desirable for several reasons. Firstly, a larger sample may reduce noise in the error term and produce narrower confidence intervals. Although the results reported here are broadly similar to those obtained by ChREAM research, there are differences. We see that CL Model 2 (Table 51) has overlapping confidence intervals for Yellow and Green ecological quality levels (cf. the confidence intervals of the Yellow and Green water quality levels of the larger ChREAM survey, on Table 25, which do not overlap). Secondly, the small sample size (and the low mean distance respondents live from the survey river) causes difficulties in estimating the impact distance has on respondents' WTP for river improvements. A larger sample size, and more variability in respondents' distances from the survey river, may enable the experiment to better reveal respondents' nuanced distance-sensitive preferences for water quality attributes.

This research fails to capture the non-use values which may exist within a proportion of the sample. The survey summary statistics, on Table 32, show that 43% of respondents didn't visit the survey river stretch of the Yare over the year prior to the study. Furthermore there appears to be a relationship between the frequency of visitation and WTP (see, for example, the rate of visitation of Class 3 respondents in the LC models). Nevertheless, there is no evidence that any of the respondents are pure non-users. It is of course feasible that non-users might hold differing values for improving the River Yare. Arguably a less methodologically inclined study which focusses solely upon estimating the total value for improvements should include pure non-users. However, this was not the purpose of this study which was more concerned with developing an approach to disaggregate the effects of ecological and recreational improvements upon WTP values.

As mentioned during the conclusions to this chapter, the transferability of research findings is a contentious issue within water valuation research (e.g. Bateman et al., 2011; Hanley et al., 2006b). Although this research primarily sought to disentangle and examine the relationships between the different sources of non-market values, it was also designed to supplement and integrate with the analysis of the wider effects of land use change conducted under the ChREAM research programme (Bateman et al., 2006a). The survey instruments used in this research were designed to capture data that may be used to produce alternative valuation assessments, e.g. travel cost assessments of benefit values, in such a way that integrates with the ChREAM dataset. It may also be possible to use this research to devise methods by which the larger ChREAM dataset may be reanalysed to disentangle and measure the differences in preferences between ecological and recreational benefit values. The results presented here may also help guide interpretation of any future LC analysis of the ChREAM dataset. Given the similarities in their research designs, the larger ChREAM dataset may be useful in helping to test the spatial transferability of the results obtained during this analysis. Given the limitations of the present research, outlined above, it would not be prudent to use the estimates of WTP in their present form for benefits transfer.

Norwich may be atypical among UK cities in providing high quality riverine recreation facilities, as the city caters for a range of tastes by offering a variety of river recreation clubs, venues and activities. This research found that the majority of recreational users live close to the sources of their recreation. The situation invokes Gidden's theory of structuration (Giddens, 1984), as the nature of the relationship between the agents (rowers, swimmers) and the structures in place to support them (clubs, recreation infrastructure) is unclear from the findings of this research. For example, have the recreationalists (agents) been drawn to the recreational structures, or have the recreational clubs (structures) developed to meet agents' demands for river recreation? Within Norwich the relationships between structures and agents may be difficult to assess due to the stability and longevity of those recursive relationships. This research found that those respondents whose primary motivation is high recreational water quality, despite having high WTP values, may not be able, or are unwilling, to visit more frequently. However, in other cities where the infrastructure supporting river recreation is less developed, there may be a high latent demand among potential recreationalists for increased recreational opportunities. Assessing whether the necessary structures are in place for those agents to enjoy, or whether the demand from agents is sufficient so that structures ought to be put in place, may be desirable as these considerations may affect non-market use values considerably.

Within this research, potential reasons why several respondent types are unwilling to increase their number of visits are discussed. For example, rowers typically visit in excess of three times per week, and, as such, could be assumed to receive no marginal benefit from improved water quality. The elasticities of future visits of the different respondent classes within the LC models were also discussed. During that discussion it was found that while classes 1 and 3 were somewhat ambivalent about future trips, the postestimation estimates for Class 2 respondents (who prefer the ecological quality of the river) shown on Table 49, predicted that they would double their annual number of trips if the water quality was guaranteed to be high. Do these results simply suggest that ecological quality improvements should be the preferred remediation strategy, given the potentially large increase in utility by those respondents who predominantly favour ecological water quality? This type of research question is worthy of future investigation, particularly if increased uptake of river recreation is desirable as a policy objective.

Postscript

The implications for water policy and the utility of this research following the decision to leave the EU

The decision to leave the EU may potentially result in legislative changes in the mid- to long-term. As mentioned in the introduction, the research presented in this thesis can be decoupled from the legislative imperatives of the WFD. And, depending on the shape of future legislation, this research may have enhanced relevance as UK policymakers become tasked with modifying legislation to suit an altered socio-economic climate. The potential shape of the UK's post-Brexit water policy, and the implications for the value of this research, are now discussed.

At the time of writing (November 2016) it is uncertain what the future holds for UK water policy or what impacts the decision to leave the EU will have. Environmental protection does not exist in a vacuum but must be viewed alongside a range of interrelated and competing factors, including the perceived need to control immigration or the economic imperative to retain access to the EU common market. Lloyd Martin, chief executive of British Water, said "following the result of the referendum on EU membership, industry finds itself in uncertain yet stimulating times" (Freyberg, 2016). The same uncertainty holds for the agricultural sector. Although the government were quick to reassure the sector by guaranteeing the current level of direct agricultural subsidies up to 2020, in line with the current CAP funding period (H. M. Treasury, 2016), many commentators agree that post-Brexit, post-CAP fiscal constraints will reduce the level of support available for UK agriculture (ADAS, 2016; House of Commons Environmental Audit Committee, 2016). This uncertainty has prompted Defra to delay publication of its 25-year Food and Farming Plan until there is a greater understanding of the outcome of the EU-exit negotiations (Farming Online, 2016).

The direction new legislation may take is dependent on two major factors: the future relationship the UK has with the EU and the objectives the UK government chooses to adopt regarding environmental issues. These factors are not mutually exclusive. In the short term, EU directives that have already been adopted into domestic UK legislation will continue to apply until they are repealed or amended by the government (Institute for European Environmental Policy, 2016). The future relationship with the EU is the key determinant of future UK water pollution legislation, as that relationship may be conditional upon the UK retaining aspects of EU policy.

In the near future, the UK government aims to embark on unprecedented Article 50 exit negotiations and is reticent to reveal its negotiating position. The shape of the UK's future relationship with the EU has been the subject of much conjecture and commentators foresee several possible scenarios, the two extremes of which will now be outlined in turn.

Stanley Johnson (Boris Johnson's father), the co-chair of Environmentalists for Europe and one of the original authors of EU environmental legislation, believes that future environmental regulations are currently a low priority for the UK government (Neslen, 2016). Of greater priority is the need to try to retain unfettered access to the EU's single market which, in 2015, accounted for 44% of UK exports and 53% of UK imports (Institute for European Environmental Policy, 2016; Miller, 2016). In addition, the government is under pressure to retain harmonised regulations and standards from innumerable sources, across all sectors of the UK economy, including manufacturing (Ministry of Foreign Affairs of Japan, 2016) and the financial services sector (Elgot, 2016). One post-Brexit relationship which would achieve these objectives would be to become a member of the European Economic Area (EEA), which allows members to freely move goods, services and capital within the EU single market. This solution, widely touted by the press as a "soft Brexit," or "the Norway model," would result in the UK being bound by many existing pieces of EU environmental legislation, including the WFD. However, EEA countries do not have to comply with all EU legislation: most crucially for the present research, the UK would no longer be bound by the rBWD or the SWD (Institute for European Environmental Policy, 2016), which have been the key drivers of improvements to UK water and wastewater quality (Freyberg, 2016).

Membership of the EEA would preserve the free movement of EU citizens into the UK. Given the Brexit Leave Campaign's vociferous focus on controlling immigration, the government may consider EEA membership to be politically untenable (Davies-Boren, 2016). An alternative scenario, “hard Brexit,” would position the UK outside of the EU entirely. This would enable the UK to have greater control over immigration policy, at the expense of preferential access to the internal market (Institute for European Environmental Policy, 2016). Under this scenario the UK would cease to be bound by EU environmental legislation.

Naturally, there are a number of variations to the relationship the UK could adopt with the EU, post-Brexit; but the two scenarios outlined above describe the extremes of the continuum with regard to the UK’s continued adherence to EU water quality legislation. Whatever the future holds, it is clear that in the mid- to long-term the UK’s environmental policy has the potential for change, the extent of which is subject to competing ideologies.

Currently, EU water policy is transposed into domestic legislation. Given the transboundary nature of environmental pollution, it may be unlikely for the government to abandon co-operative arrangements to tackle environmental degradation. Defra minister, Rory Stewart, has stated "the basic structure of European environmental law in relation to our Department, I think, is very close to what we think is sensible. It is what we would intend to do in the United Kingdom" (BBC, 2016). This sentiment is echoed by the Chartered Institution of Water and Environmental Management, which believes “the logic of an EU-led initiative on the environment is sound” (Freyberg, 2016).

Others are far less optimistic. In the worst case, Dr Charlotte Burns, of Friends of the Earth, warns of an erosion of UK environmental policy whereby the UK regains a reputation for being the ‘Dirty Man of Europe’ (2016). To a lesser degree, this pessimism is shared by the House of Commons Environmental Audit Committee, which recently reported “the overwhelming evidence is that EU membership has improved the UK’s approach to the environment and ensured that the UK’s environment has been better protected… Many witnesses implied that if the UK were free to set its own environmental standards, it would set them at a less stringent level than has been imposed by the EU" (House of Commons Environmental Audit Committee, 2016). Kerry McCarthy MP, the Shadow Secretary of State for the Environment, Food and Rural Affairs, shares these concerns that Brexit jeopardises the future quality of the UK's environmental quality (Landscape Institute, 2016).

While it is unlikely that the UK's environmental policy would change immediately following Brexit, there is an emerging consensus that the future for water policy lies in subsequent amendments to existing legislation (Country Land and Business Association Limited, 2016; Shepherd and Wedderburn LLP, 2016). This view is shared by Antoine Simon, legal expert at Friends of the Earth Europe, who said "what would change is that future governments would be able to review the environmental legislation in place and apply the standards they deem useful, or reasonable, or necessary" (Davies-Boren, 2016).

Post Brexit, there are opportunities for a simpler agricultural policy which could focus on the UK's priorities for a competitive agricultural sector, with an increased emphasis on efficiency, streamlining and deregulation (H.M. Treasury, 2016; Cowell and Owens, 2016). The Eurosceptic, George Eustice, promoted by Prime Minister May to Minister of State at Defra (Tasker, 2016), believes the UK could develop a more flexible approach to environmental protection following Brexit (Neslen, 2016). This perspective has raised doubts concerning the implementation and enforcement of environmental legislation, as the UK becomes subject to a simplified judicial process (Miller, 2016). There are also concerns that policymakers may return to legislating on environment issues in a fragmented and ad-hoc manner (Lowe and Carter, 1994), or that policymakers may expand ineffective voluntary regulatory frameworks (Mitchell, 2016). Environmental organisations warn of public outcry if environmental legislation were to be seriously compromised. The head of WWF UK, David Nussbaum, warns "there will be one mighty battle if the government uses Brexit to try to reduce standards on the environment. Why should we in the UK have a worse environment than our neighbours?" (Harrabin, 2016; Mitchell, 2016).

Whatever the exit conditions, UK agricultural policy is likely to change as, even with a "soft Brexit," EEA members do not participate in the CAP (Burns, 2016). The government has indicated that without the CAP it would be unlikely to maintain current levels of agricultural subsidy (H.M. Treasury, 2016). If direct subsidies to the agricultural sector are reduced, the House of Commons

Environmental Audit Committee (2016) predicts that structural changes will occur within the farming sector, with a transition to fewer, larger producers, which are better able to remain competitive due to improved economies of scale. Structural change carries risks of detrimental consequences for the environment: the LUAM 'healthy diet' scenario analysis (Figure 18 in Chapter 2) predicted a rise in localised FIO pollution corresponding with a structural change towards larger, more intensive dairy enterprises.

Environmental groups have expressed concerns about the overall level of funding for agri-environmental protection schemes outside of the CAP (H.M. Treasury, 2016). The government recognises environmental degradation as a market failure in the agricultural sector and, in return for continued agricultural subsidies, may require environmentally sensitive farming practices governed by an ecosystem services approach (ADAS, 2016; H.M. Treasury, 2016) or a move away from the blanket principles of EU legislation, towards a greater utilisation of CBA to assess specific changes in the agricultural sector (H. M. Treasury, 2003).

As with the agricultural sector, the water industry may be subject to legislative changes. Recent analysis by Defra found no significant changes in the overall number of water bodies classified at 'excellent' or 'good' surface water status between 2008 and 2012 (Institute for European Environmental Policy, 2016). In the face of frequent criticism for non-compliance with EU rBWD standards, and in the case of a soft Brexit, the UK may decide to relax bathing water legislation (DWF LLP, 2016). Under Brexit outside of the EEA there could be a retreat from the tough objectives of the WFD (Institute for European Environmental Policy, 2016). British Water, the industry trade association, has said that leaving the EU is "certain to have a significant impact on a sector where considerable investment is driven by EU directives on water, wastewater and the environment" and has begun lobbying for a relaxation of water quality legislation (British Water, 2016; Freyberg, 2016).

This section has provided a broad overview of the likely challenges faced by policymakers, the agricultural sector and the water industry as a result of Brexit. The UK may obtain greater flexibility to shape the future direction of environmental legislation and, if so, it is likely that the UK will make greater use of CBA within decision making. Defra will become tasked with shaping legislation to suit an altered political and economic climate and will require guidance in the coming years. Post-Brexit, the research presented within this thesis - modelling FIO pollution in response to land use change and assessing the non-market benefits of alternative water quality investment strategies - becomes more relevant.

Appendices

Appendix I: Agreement between Professor Dave Kay (CREH) and Professor Ian Bateman (UEA) concerning collaborative work under the ChREAM project

Rationale for the agreement

This agreement sets out plans and conditions regarding the above collaboration. It is necessitated because of the commercial sensitivity regarding the data and models held and to be generated by CREH concerning the determinants of faecal indicator organism (FIO) levels in UK rivers. The agreement is designed to facilitate full collaboration between CREH and UEA so as to fulfil the requirements of the ChREAM project. It is also hoped that this may provide a basis for more long-term collaboration.

Introduction to the research

The ChREAM project covers a wide variety of work. However, of particular relevance to this agreement is the modelling of the impact of land use change (typically initiated by policy implementation or shifts in market forces) upon river water quality. While other work considers impacts in terms of nutrient levels, the CREH/UEA collaboration examines fluctuations in FIO levels as a result primarily (but not exclusively) of land use change. Carlo Fezzi and Ian Bateman of the ChREAM project at UEA are developing a land use model intended to predict changes in land use pattern as a result of the above policy and market drivers. This modelling exercise combines temporal and cross sectional panel data on agricultural land use and farm finances to predict land use under a variety of scenarios.

The key research objective of the planned collaboration is to link a unified model of FIO fluctuations to the land use model. This will allow inspection of how changes in land use affect FIO response.

There are a number of stages in achieving this. In overview these are as follows:

1. UEA is in the process of providing CREH with land use, environmental characteristic and population data covering all of the catchments for which CREH hold FIO records. That data is also held for the whole of the UK; a point which

will be vital to the transfer exercise described subsequently. For example, it is expected that shifts in population and climate change may also be amenable to incorporation within the FIO modelling exercise.

2. CREH have already undertaken research to combine all of their FIO catchment data together. They will use the UEA data to provide a consistent set of predictors of FIO fluctuations. CREH will undertake a new modelling exercise using this consistent set of predictors. Once the initial analysis is complete CREH will invite Danyel Hampson (PhD student, UEA – Dave Kay being an external supervisor and Ian Bateman being the lead internal supervisor) to be involved in further aspects of this process (those aspects to be agreed), the intention being that his work will form a valid element in his thesis. Carlo Fezzi and Ian Bateman will liaise with CREH to ensure that the FIO model is compatible with the land use model.

3. CREH will provide the full details of FIO model (including estimated parameters, standard errors, etc.) to Ian Bateman who will take responsibility for its confidentiality at UEA (see notes on publication etc. subsequently).

4. The FIO model will be linked to the land use model in such a manner that each scenario run of the latter generates estimates of FIOs. The objective is to allow the joint model to run to optimise in a manner that maximises farm profit subject to constraints imposed by river FIO and nutrient levels. Further scenarios will also be considered for alternative optimisations (e.g. in respect of minimising FIOs, etc.)

5. Because the joint land use and FIO model runs on common variables, and the latter are held for all points across the UK, the intention is to extrapolate to this wider area.

6. A range of further collaborative extensions are foreseen. For example, as part of his thesis Danyel Hampson will undertake a study into the various economic values (including informal recreation) that may be generated by reductions in FIO levels. Danyel will also be assisted by Dave regarding modelling of the health consequences of such changes.

Conditions of agreement

Professors Dave Kay and Ian Bateman jointly agree to the above work plan subject to the following conditions:

1. At the end of the ChREAM project any data supplied and all details of the FIO model will be deleted or returned to Professor Dave Kay unless he asks otherwise.

2. At no point will any data, model parameters or any information be published or otherwise disseminated which could allow a third party to operationalise the FIO model.

3. Instead it is foreseen that publications will focus upon summary statistics and mapped outputs from the model and relationships described in terms of directional responses rather than quantified parameters.

4. To guarantee the above, Professor Dave Kay will hold a veto over any publication or other output concerning the FIO aspect of ChREAM.

Ian Bateman

26th February 2008

i.bateman@uea.ac.uk

Appendix II: list of publications and presentations

Publications

Crowther, J., Hampson, D., Bateman, I. J., Kay, D., Posen, P., Stapleton, C. and Wyer, M. (2010). *Generic modelling of faecal indicator organism concentrations in the UK. Working paper ECM 10-01*, Norwich: The Centre for Social and Economic Research on the Global Environment (CSERGE), University of East Anglia.

Crowther, J., Hampson, D., Bateman, I. J., Kay, D., Posen, P., Stapleton, C. and Wyer, M. (2011). Generic modelling of faecal indicator organism concentrations in the UK. *Water*, 3(2), pp. 682-701.

Hampson, D., Crowther, J., Bateman, I. J., Kay, D., Posen, P., Stapleton, C., Wyer, M., Fezzi, C., Jones, P. and Tzanopoulos, J. (2010). Predicting microbial pollution concentrations in UK rivers in response to land use change. *Water Research*, 44(16), pp. 4748-4759.

Hampson, D. I., Ferrini, S., Rigby, D., and Bateman, I. J. (2017). River Water Quality: Who Cares, How Much and Why? *Water*, 9(8), p.621.

Presentations

Audsley, E., Barns, S., Bateman, I. J., Binner, A., Brouwer, R., Crowther, J., Coombes, E., Davies, H., Day, B., Deflandre, A., Ferrini, S., Fezzi, C., Hadley, D., Hampson, D., Hime, S., Hutchins, M., Jones, A., Kay, D., Leeks, G., Lovett, A., Neal, C., Pearn, K., Posen, P., Rigby, D., Sandars, D., Turnbull, D., Turner, K. and Willoughby, B. (2009). “The future for water quality”, poster and group presentation at The Future of Rural Land Use, Rural Economy and Land Use (RELU) conference, 4 June 2009, London.

Bateman, I. J., Binner, A., Coombes, E., Day, B. H., Ferrini, S. Fezzi, C., Hampson, D. and Posen, P. (2010). “The ChREAM and SEER Projects”,

poster and group presentation to HRH The Prince of Wales, 26 January 2010, University of East Anglia, Norwich.

Bateman, I. J., Coombes, E. and Hampson, D. (2010). “The ChREAM Project”, poster presentation at the Royal Norfolk Show, 29 June 2010, Norfolk Showground, Norwich.

Hampson, D., Bateman, I. J. and Lovett, A. (2008). “Predicting faecal organism pollution in riverine environments in response to land use changes”, paper presented at the University of East Anglia School of Environmental Sciences Seminar Series, 5 June 2008, School of Environmental Sciences, University of East Anglia, Norwich.

Hampson, D., Bateman, I. J. and Lovett, A. (2009). “Benefits of reducing microbial river pollution”, poster presentation at the Environment Agency ‘Better Environments, Better Lives’ Conference, 27 February 2009, Birmingham, England.

Hampson, D., Crowther, J., Bateman, I. J., Kay, D., Posen, P., Stapleton, C., Wyer, M., Fezzi, C., Jones, P. and Tzanopoulos, J. (2010). “Predicting Microbial Pollution Concentrations in UK Rivers in Response to Land Use Change”, paper presented to the University of East Anglia Water Security Research Centre, 5 March 2010, University of East Anglia, Norwich, England.

Hampson, D., Crowther, J., Bateman, I. J., Kay, D., Posen, P., Stapleton, C., Wyer, M., Fezzi, C., Jones, P. and Tzanopoulos, J. (2010). “Predicting Microbial Pollution Concentrations in UK Rivers in Response to Land Use Change”, paper presented at the International Water Association Water Research Conference, 11-14 April 2010, Lisbon, Portugal.

Hampson, D., Ferrini, S., Rigby, D. and Bateman, I. J. (2016). “River water quality: who cares, how much and why?”, paper presented to the School of Agriculture, Food and Rural Development, Newcastle University, 3 February 2016, Newcastle. (paper presented by S. Ferrini).

Hampson, D., Ferrini, S., Rigby, D. and Bateman, I. J. (2016). “River water quality: who cares, how much and why?” Paper presented at the European Association of Environmental and Resource Economists 22nd Annual Conference, 22-25 June 2016, Zurich, Switzerland. Available online at: http://www.webmeets.com/files/papers/EAERE/2016/471/2016_EAERE_final.pdf, last accessed 12 August 2016. (paper presented by I. J. Bateman).

Press releases

Natural Environment Research Council (2011), Fences *Help Clean Up Livestock Pollution of Rivers,* 8 March 2011.

Rural Economy and Land Use (2010), *Fences Could Help Clean Up Watercourses*, 20 September 2010.

Finally, post-doctoral research, which incorporates elements of the research conducted within thesis, was presented at the 2013 annual meeting of UK Q users:

Hampson, D., Bateman, I. J., Simmons, P. and Rigby, D. (2013). “Using Q methodology to reveal non-expert preferences on UK river water quality issues: a pilot study in East Anglia”, paper presented at the T and Q annual meeting of UK Q users, 17th May 2013, University of East London.

Appendix III: choice experiment questionnaire

Suitability questions

[POTENTIAL RESPONDENT MUST BE 18 YEARS OF AGE OR OVER]

Q1. Hello, is there anyone over the age of eighteen available to speak to?

IF NO – It's no problem, I just wanted to ask them some questions about a survey I'm doing. Thanks, goodbye.

Q2. Hello, my name is Danyel Hampson. I'm a university student and I'm conducting a survey as part of a University project looking at local environmental issues. I'm keen to get the views of local people. Would you help by completing a questionnaire? It lasts about 30 minutes at most and is completely anonymous and confidential. I have a letter from my supervisor and several other forms of ID. [OFFER PROOFS]

IF NO – Thanks for your time, goodbye. [COLLATE REFUSALS DATA]

Q3. Are any of the following statements true? [READ STATEMENTS FROM SHOWCARD 1]

IF YES – I'm sorry, but due to its format, you'll be unable to take part in this survey, but thank you for your time.

Introduction

This survey is mainly about water quality in rivers. This does not have any effect on the quality of your tap water; it only affects river plants, animals and fish and the types and quality of recreation that visitors can enjoy. I want to get a balanced picture, and am just as interested in talking to people who don't visit rivers as those who do.

Q4. How long you have lived at this address or an address in the surrounding mile or so? [RECORD RESPONSE]

Q5. Please could you tell me your home postcode? [RECORD RESPONSE]

Questions about actual river use

To start with I'd like to ask you some questions about recreational trips you take which involve a visit to a river. By a 'recreational trip' I mean when you leave your home with the deliberate purpose of visiting such a place rather than, for example, just passing the river on your way somewhere else.

So this would include trips where you go out for a walk or bike ride alongside the river, as well as trips to go fishing or canoeing in the river.

Q6. Looking at the categories in [SHOWCARD 2], which best describes how often, over the past 12 months, you have been on trips to rivers or riverside sites?

[RECORD RESPONSE]

Q7. Please take a look at the following map of the area [SHOWCARD 3]. Can you point to where your home is? [MARK MAP ON ANSWER SHEET WITH 'x' SYMBOL]. [ALL RESPONDENTS MUST IDENTIFY THEIR HOME EVEN IF THEY DO NOT VISIT RIVERS]

IF RIVER TRIPS [IN Q6] = 0, GO TO 'POLLUTION INFORMATION', OTHERWISE CONTINUE

Q8. You said that you took [READ TOTAL RIVER TRIPS FROM ANSWER SHEET, Q6] trips to rivers each year. How many trips were to sites on this map? [SHOWCARD 3]. [RECORD RESPONSE]

IF RIVER TRIPS IN AREA [IN Q8] = 0, GO TO 'POLLUTION INFORMATION', OTHERWISE CONTINUE

Q9. Can you point to the place you have visited most often on a river in this area? [SHOW SHOWCARD 3, RECORD RESPONSE AS 'o' ON MAP]

Q10. From these categories [SHOWCARD 4] can you tell me the main purpose of your visit(s) to that location? [RECORD RESPONSE]

Q11a. What other things have you done on visits to any of the rivers in the area? [SHOWCARD 4] [RECORD RESPONSES]

Q11b. [ASK IF ANSWERS TO Q10 AND Q11A DO NOT INCLUDE SWIMMING, BOATING OR CANOEING, OTHERWISE CONTINUE TO 'POLLUTION

INFORMATION'] Do you ever do any riverside activity in the area which involves contact with the river water? This would include things like swimming, paddling, boating or canoeing [RECORD RESPONSE]

Pollution Information

I now want to show you some information about pollution problems in rivers. As this card shows [SHOW SHOWCARD 5], there are two main type of pollution that affect UK rivers.

Ecological pollution [INDICATE] from washing detergents and farm fertilisers cause algae to grow in rivers. This reduces the oxygen available for fish and water plants, etc. However, this does not pose major risks for human health.

Biological pollution, such as sewage, [INDICATE] comes mainly from households and farm animals and makes rivers unsuitable for recreation such as paddling, swimming or boating. However, most water plants and most fish can tolerate a fair amount of biological pollution. Biological pollution can be harmful to human health, as I will shortly explain.

[SHOW ECOLOGICAL QUALITY LADDER - SHOWCARD 6a]

This picture marked with a blue circle shows a river of the highest ecological quality [INDICATE PICTURE]. This symbol [POINT TO GAME FISH SYMBOL] shows that rivers like this are suitable for pollution sensitive game fish such as salmon and trout. This symbol [POINT TO COARSE FISH SYMBOL] shows it's suitable for coarse fish, such as carp and chub while this one [POINT TO BIRD SYMBOL] shows it's suitable for all bird species. As you can see, there is a wide variety of plants in and around the river which has very clear water.

This green circle [INDICATE] indicates the presence of some ecological pollution with far fewer game fish [POINT TO GAME FISH SYMBOL]. But there is no reduction in coarse fish or birds [POINT TO SYMBOLS]. The variety of plants in and around the river is slightly lower but the water is still fairly clear.

The yellow circle [INDICATE] shows still higher levels of ecological pollution, with virtually no game fish [POINT TO GAME FISH SYMBOL] and significantly less coarse fish [POINT TO COURSE FISH SYMBOL]. The variety of plants is lower and algae has substantially reduced the water clarity although there will still be a number of birds.

The red circle [INDICATE] shows the highest level of ecological pollution with virtually no fish, few birds or water plants and very cloudy water.

Is there anything you want to ask about ecological pollution in rivers or these pictures? [ANSWER ANY QUESTIONS]

[PUT ECOLOGICAL LADDER ASIDE]

[SHOW RECREATIONAL QUALITY LADDER - SHOWCARD 6b].

I am going to use these images to show the presence of biological pollution in rivers and the effect it has on recreation.

Biological pollution can cause a variety of illnesses ranging from nausea and diarrhoea to, very occasionally, more serious illnesses which can, very rarely result in death.

The more contact a person has with biologically polluted water, the more likely it is that they will get ill. Someone in the water swimming has a higher risk of illness than a person in a canoe who only gets splashed with the water. A person on the river bank, who has no contact with the water, has no increased risk of getting ill.

As the amount of biological pollution increases, the risk of illness to recreational river users increases. This affects the types and quality of recreation that users can enjoy.

We'll use these images [INDICATE RECREATIONAL QUALITY LADDER – SHOWCARD 6b] to show different recreational qualities. As you can see the pictures are arranged from higher to lower quality [INDICATE]

These blue images, [INDICATE 1st ROW], show a river of the highest recreational quality. The risk of illness is low. These symbols [INDICATE SYMBOLS] show that a river of this quality is suitable for swimming and boating.

The next type of river [INDICATE 2nd ROW] has a higher risk of illness. This type of river is suitable for boating but is no longer suitable for swimming. [POINT TO SYMBOLS].

These red images show a river of the lowest recreational quality [INDICATE 3rd ROW]. This river has the highest risk of illness and isn't suitable for swimming or boating [POINT TO SYMBOLS]

Is there anything you want to ask about biological pollution in rivers or these pictures? [ANSWER ANY QUESTIONS]

[PLACE SHOWCARD 6a NEXT TO SHOWCARD 6b – ENSURE RESPONDENT CAN SEE BOTH]

I am going to use pictures from these cards to illustrate the different combinations of ecological pollution and biological pollution sewage in the survey river.

An important fact is that ecological pollution and sewage in rivers can happen independently from each other. For example, a river of high ecological quality [INDICATE BLUE ECOLOGICAL PICTURE] might look inviting but may have levels of biological pollution which makes it unsafe for either swimming or boating [INDICATE RED RECREATION ICONS]. On the other hand the water in an ecologically polluted river [INDICATE RED ECOLOGICAL PICTURE] may have no biological pollution and be perfectly safe for people to swim in [INDICATE BLUE RECREATION ICONS]

We'll be using these pictures throughout the interview so please take your time to get used to them. [PAUSE]

IF RIVER TRIPS IN AREA [IN Q8] = 0, GO TO 'INTRODUCING THE SURVEY STRETCH', OTHERWISE CONTINUE

Now, please think again about the river site [INDICATE THE SITE GIVEN IN THE ANSWER TO Q9, On SHOWCARD 3] that you visited most often.

Q12, Looking at the ecological quality pictures [SHOWCARD 6a] which colour best describes the ecological quality of the river at your most visited site? [RECORD RESPONSE]

Q13, Looking at the recreational quality pictures [SHOWCARD 6b] which colour best describes the recreational quality of the river at your most visited site? [RECORD RESPONSE]

Introducing the survey stretch

I'd now like you to think about the river along this stretch, highlighted in purple [INDICATE PURPLE STRETCH ON SHOWCARD 7]. As you can see, this stretch of the river flows from here [INDICATE], through the city centre, to here [INDICATE] just to the south-east of Norwich.

IF RIVER TRIPS IN AREA [IN Q8] = 0, GO TO 'INTRODUCTION TO CHOICE EXPERIMENT', OTHERWISE CONTINUE

Q14. How many of your river trips in the area were to sites along the purple stretch? [SHOWCARD 7] [RECORD RESPONSE]

IF RIVER TRIPS ALONG PURPLE STRIP = 0, GO TO 'INTRODUCTION TO CHOICE EXPERIMENT', OTHERWISE CONTINUE

Q15. From these categories [SHOWCARD 4] what was the main purpose of that/those visit(s)? [RECORD RESPONSE]

Q16. What other things did you do on that/those visit(s)? [SHOWCARD 4] [RECORD RESPONSES]

Q17a. Looking at the ecological quality ladder [SHOWCARD 6a] which colour do you think best describes the actual ecological quality of the river along this stretch? [RECORD RESPONSE]

Q17b. Looking at the recreational quality ladder [SHOWCARD 6b] which colour do you think best describes the recreational quality of the river along this stretch? [RECORD RESPONSE]

Choice experiment section

Introduction

The next few questions are among the most important of this interview.

In 2004 the UK government agreed new laws to improve the quality of certain rivers. One of these rivers is the river along this purple stretch. [INDICATE PURPLE STRETCH ON SHOWCARD 7]. Other rivers in the area are not highlighted [INDICATE] as we will not be considering the quality of these rivers.

[IF ASKED, SAY THAT THE ACTUAL WATER QUALITY OF THE RIVERS IS UNKNOWN, ONLY THAT WATER QUALITY CHANGES ARE BEING CONSIDERED]

Choice experiment.

The following questions ask you to choose between two future options for the water quality of this stretch of river [INDICATE PURPLE STRETCH, SHOWCARD 9]. The options are labelled A and B [INDICATE OPTIONS ON SHOWCARD 8.1]. Each option shows the ecological [INDICATE] and recreational [INDICATE] qualities of the river stretch, and the level of your annual water bill [INDICATE]. In all cases this bill will either be unchanged or increase. This is because improving river water quality requires investments which would have to be paid for by higher water bills. All water users would have to pay, including industry and farmers, but also households, because they also contribute to water pollution. Any increase in bills would start in early 2014 and the improvements to water quality would be finished by 2015.

For each question simply choose the situation you would prefer for the purple river stretch. In comparing A and B please consider the location of the river, [INDICATE PURPLE STRETCH, SHOWCARD 7] how close it is to your home [INDICATE RESPONDENTS HOME (ANSWER TO Q7)], and whether you would benefit from them. Please remember that any increases in water bills would mean you have less money to spend on other things.

Q18a. Let's look at the first question [SHOWCARD 8.1]. Here you can see that under option A the ecological quality is [STATE COLOUR], the recreational quality is [STATE COLOUR] and the increase to your annual water bill is [STATE AMOUNT]. Under option B the ecological quality is [STATE COLOUR], the recreational quality is [STATE COLOUR] and the annual water bill increase is

[STATE AMOUNT]. Take your time to consider these two options and then let me know which one you would prefer for the purple stretch [RECORD RESPONSE]

Q18b-l. The next few questions have the same format. Here is the next question, [SHOWCARD 8.2] where option A is like this [INDICATE] and option B is like this [INDICATE]. Again, which one would you prefer for the purple stretch? [RECORD RESPONSE]

[WORK THROUGH EACH QUESTION WITH THE RESPONDENT, RECORDING RESPONSES]

[AFTER SHOWCARD 8.9, STATE 'THERE ARE 3 MORE TO GO', AFTER SHOWCARD 8.10, STATE 'THERE ARE 2 MORE TO GO', AFTER SHOWCARD 8.11 STATE 'THIS IS THE FINAL QUESTION ON THIS']

Questions on respondents' future river use

I'm interested in your how your use of this river might change in the future.

Q19a. If you haven't already, would you visit or use this river stretch [INDICATE PURPLE STRETCH ON SHOWCARD 7] if improvements were made so that the water quality was guaranteed to be like this? [SHOWCARD 9] [RECORD RESPONSE]. [IF YES GO TO 19b, IF NO GO TO Q20]

Q19b. If yes, how many days do you think you would visit the river over the next year? [RECORD RESPONSE]

Q19c. From these categories [SHOWCARD 4] what might be the main purpose of that/those visit(s)? [RECORD RESPONSE]

Q19d. What other things might you do on that/those visit(s)? [SHOWCARD 4] [RECORD RESPONSES]

Q20. How likely do you feel it is that the river quality proposed in the last question [SHOWCARD 9] would be provided as described? [SHOWCARD 10] [RECORD RESPONSE]

Choice experiment control questions

Q21. Overall, how easy or difficult did you find it to answer the questions involving changes in water quality and water bills? [SHOWCARD 11] [RECORD RESPONSE]

I'd like to know, from these categories [SHOWCARD 12], how important each of the following issues were in determining your answers to the choice questions.

Q22a. The distance from where you live to where the improvement would happen? [RECORD RESPONSE]

Q22b. The size of the water bill increases? [RECORD RESPONSE]

Q22c. The size of the ecological quality improvements? [RECORD RESPONSE]

Q22d. The size of the recreational quality improvements? [RECORD RESPONSE]

Q22e. Any other? [RECORD RESPONSES]

Q22f. How important was this other issue? [RECORD RESPONSES]

Please tell me, from these categories [SHOWCARD 13], who you think should pay for water quality improvements.

Q23a. The government or council? [RECORD RESPONSE]

Q23b. Water companies? [RECORD RESPONSE]

Q23c. Domestic water and sewerage customers? [RECORD RESPONSE]

Q23d. The agricultural sector? [RECORD RESPONSE]

Q23e. The polluter? [RECORD RESPONSE]

Q23f. The recreational user? [RECORD RESPONSE]

Q23g. Any other? [RECORD RESPONSE]

Thank you for your help with that.

Survey control questions

To finish off, I just have a few more questions about you and your household. These will only be used for statistical purposes to see if we have interviewed a fair range of people and please remember that all of these answers are completely confidential.

Q26a. What is your age? [RECORD RESPONSE]

Q26b. What is your ethnic background? [SHOWCARD 15] [RECORD RESPONSE]

Q26c. What is your religion? [SHOWCARD 15a] [RECORD RESPONSE]

Q26d. From this list what is your highest educational qualification? [SHOWCARD 16] [RECORD RESPONSE]

Q27. How many people including yourself are in your household, by which I mean you your partner and any members of your family that you currently live with? [RECORD RESPONSE]

Q28. How many of them are younger than 18? [RECORD RESPONSE]

Q29. Looking at these categories [SHOWCARD 17] could you tell me which best approximates your total household income before tax? [SELECT ONE ONLY]

[IF NECESSARY, REASSURE RESPONDENT THAT ALL INFORMATION IS COMPLETELY CONFIDENTIAL AND THIS IS THE BEST INDICATOR OF WHETHER I HAVE INTERVIEWED A REPRESENTATIVE RANGE OF PEOPLE]

Q30a. Looking at this list of organisations [SHOWCARD 18] please tell me which, if any, you are a member of. You can select more than one. [RECORD RESPONSES] [IF SPORTS CLUB OR OTHER, RECORD TYPE]

Q30b. Which organisations are any others in your household a member of? [RECORD RESPONSES] [IF SPORTS CLUB OR OTHER, RECORD TYPE]

Q31. Which of these statements [SHOWCARD 19] best describes your current employment status? [SELECT ONE ONLY] [RECORD RESPONSE]

Q32a. Do you go fishing? [RECORD RESPONSE] [If NO, GO TO Q34]

Q32b. Do you hold a fishing licence? [RECORD RESPONSE] [If NO, GO TO Q33a]

Q32c. How much does the licence cost per year? [RECORD RESPONSE]

Q33a. Do you belong to a fishing club? [RECORD RESPONSE][If NO, GO TO Q34]

Q33b, How much does membership cost per year? [RECORD RESPONSE]

Q34. Do you own any of the following craft? [SHOWCARD 20] [RECORD RESPONSE]

Q35a. Do you belong to a rowing, canoeing, or any other river based recreation club? [RECORD RESPONSE] [IF NO, GO TO ‘LEAD OUT’]

Q35b. What Type? [RECORD TYPE] [IF NO, GO TO ‘LEAD OUT’]

Q35c. What is the membership cost per year? [RECORD RESPONSE]

Lead out

That was the last of my questions. This survey will continue for several weeks. At the end of that time there is a possibility that my supervisor might have some follow up questions - this would be for quality control purposes only and not to ask any further questions about rivers. Could you please give me a telephone number where you can be contacted and your first name? This data will be kept strictly confidential and held for 3 months following this survey after which it will be destroyed. [RECORD RESPONSE]

That's the end of the interview. Thank you very much for your time and help, it is very much appreciated.

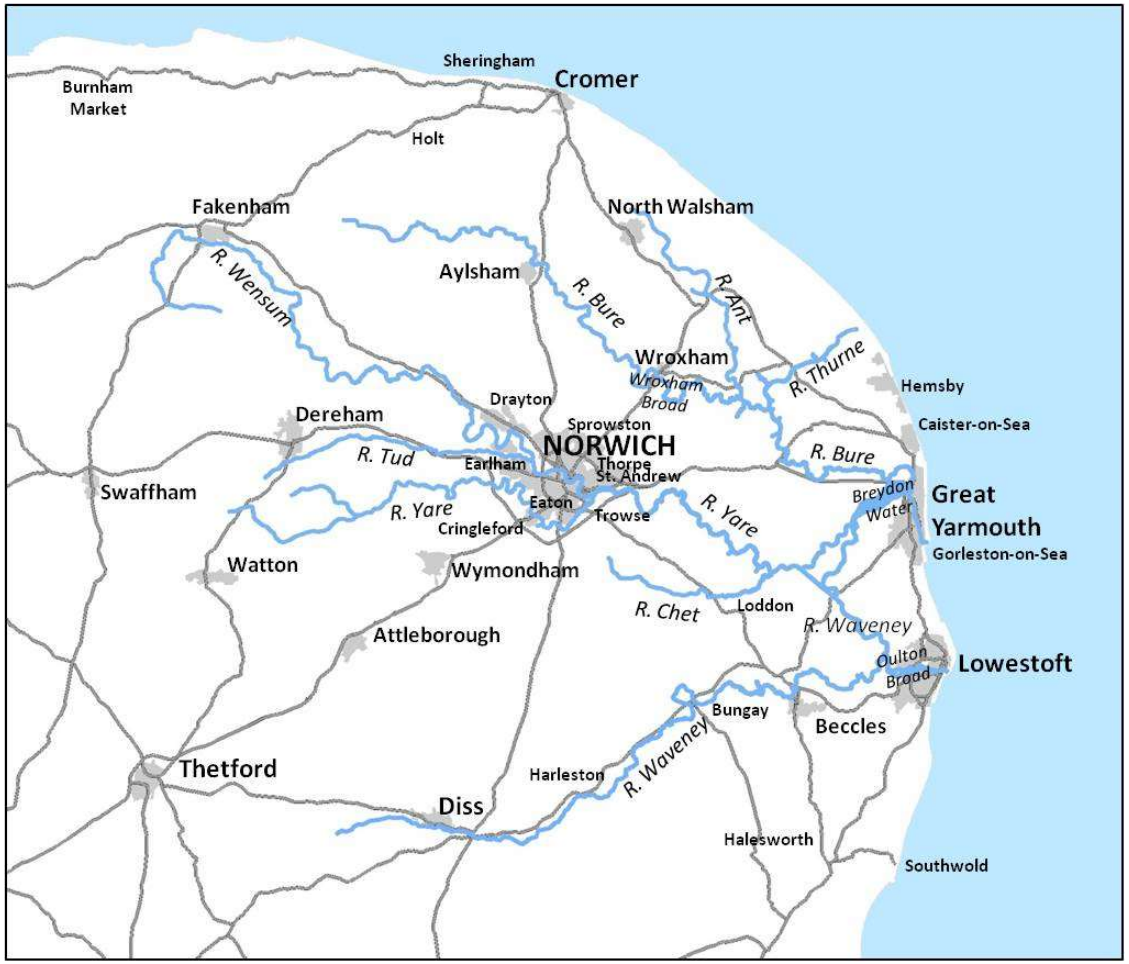
SHOWCARD 1 – Participation Statements

Statement	True	False
I am colour blind		
Myself or my partner are NOT responsible for the water bill		
I have lived at this address, or within the surrounding mile for less than a year		

SHOWCARD 2 – Number of trips

Code	Trip frequency	Trips per year
A	Did not go	No trips
B	Once in the last year	1 trip
C	Twice in the last year	2 trips
D	Once every three months	4 trips
E	Once every month	12 trips
F	Once a fortnight	26 trips
G	Once a week	52 trips
H	Twice a week	104 trips
I	More than twice a week	208 trips
J	Every day	365 trips

SHOWCARD 3 – Survey area



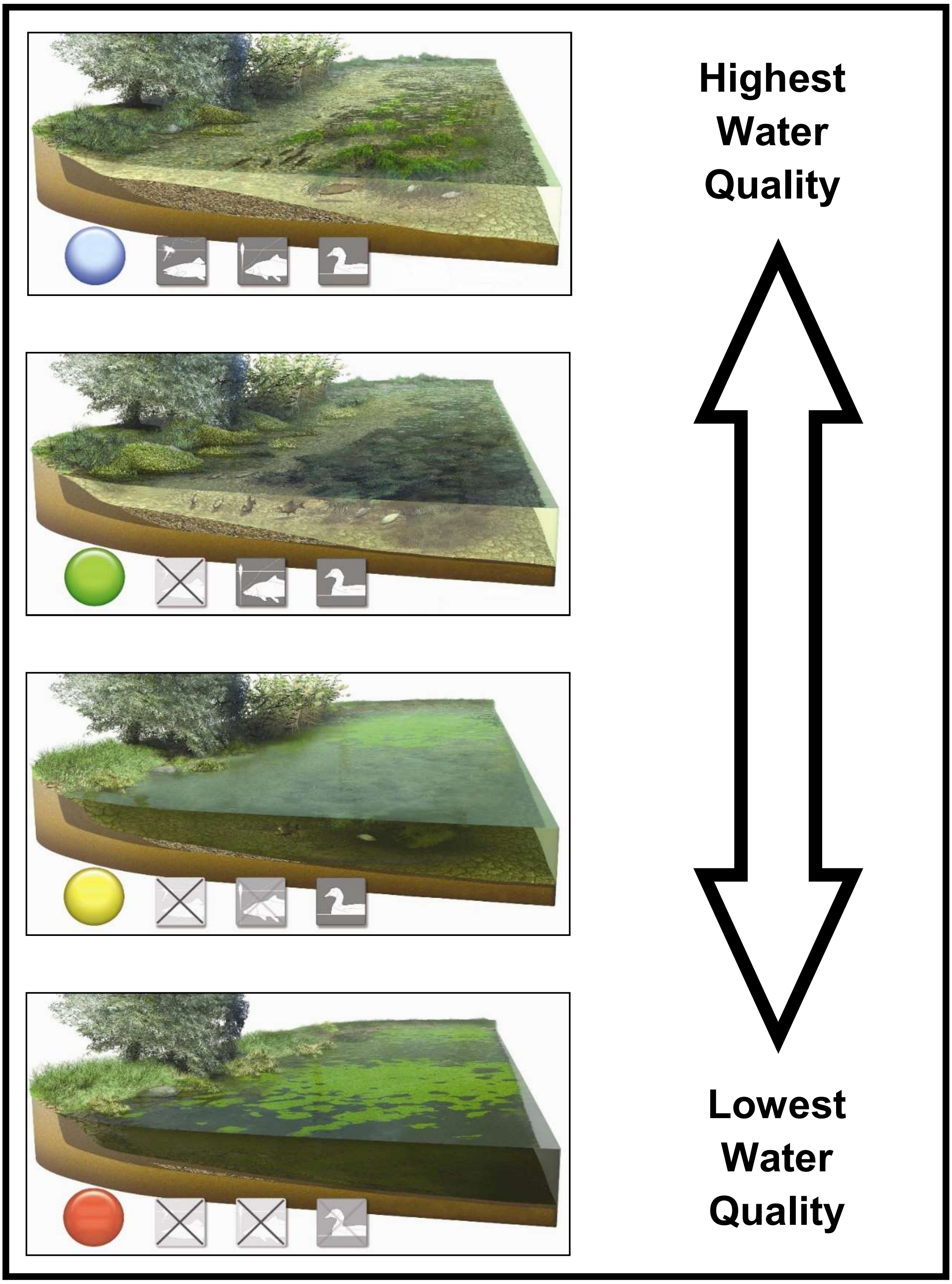
SHOWCARD 4 – Purpose of visit

Code	Activity
A	Walking / Rambling
B	Dog walking
C	Running
D	Picnic
E	Feeding birds
F	Wildlife watching
G	Cycling
H	Motorised Boating
I	Canoeing / Rowing (or other non-motorised water recreation)
J	Swimming / paddling
K	Fishing / Angling
L	Other (please state)

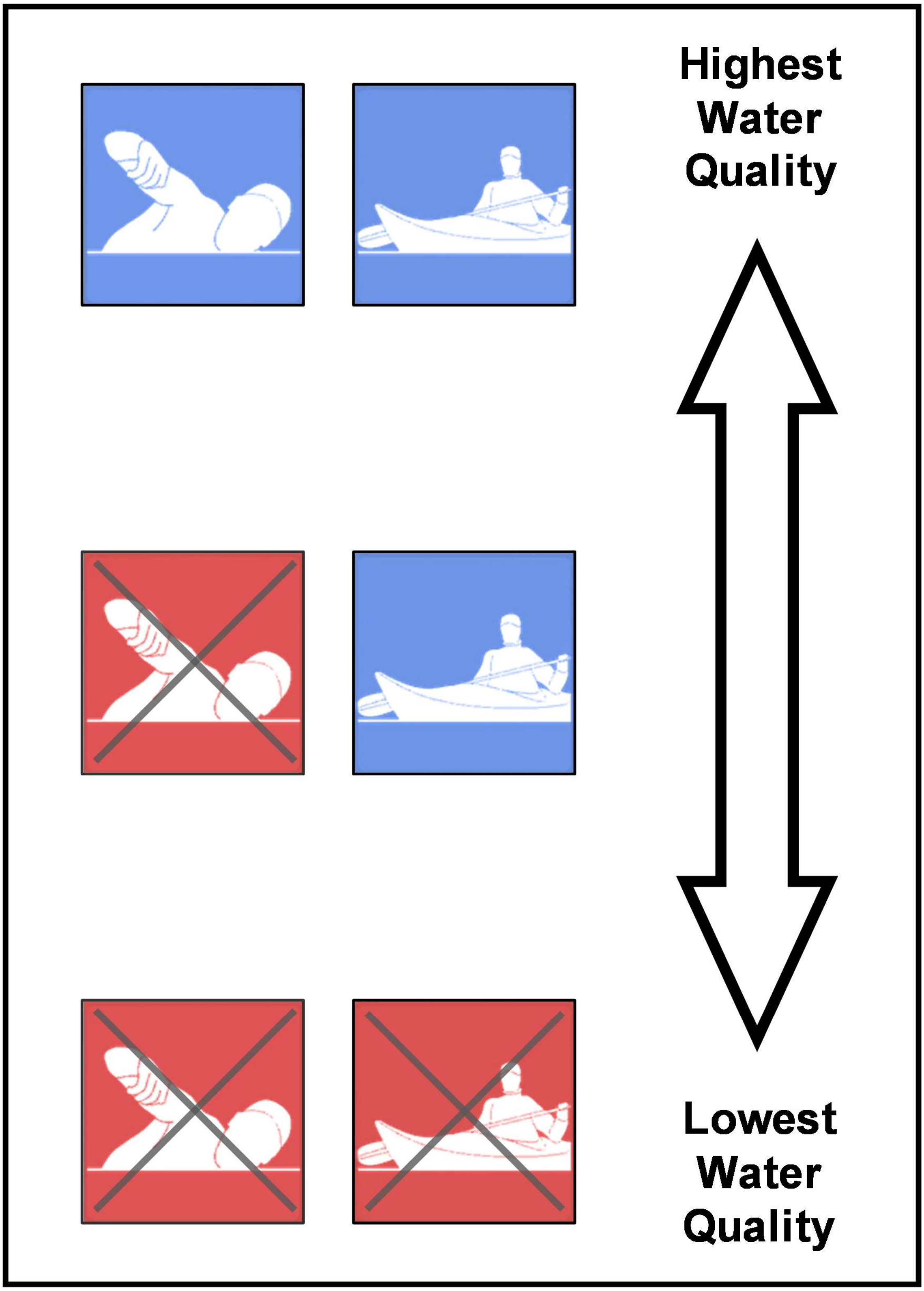
SHOWCARD 5 – UK river pollution problems

Pollution type	Main sources	Main effects
Ecological	• Washing Machines, detergents, etc. • Farm fertilizers, etc.	• Causes algae to grow • Reduces oxygen in rivers • Harms fish, river plants, etc.
Biological	• Households, e.g. sewage • Farm animals	• Makes rivers unsuitable for recreation

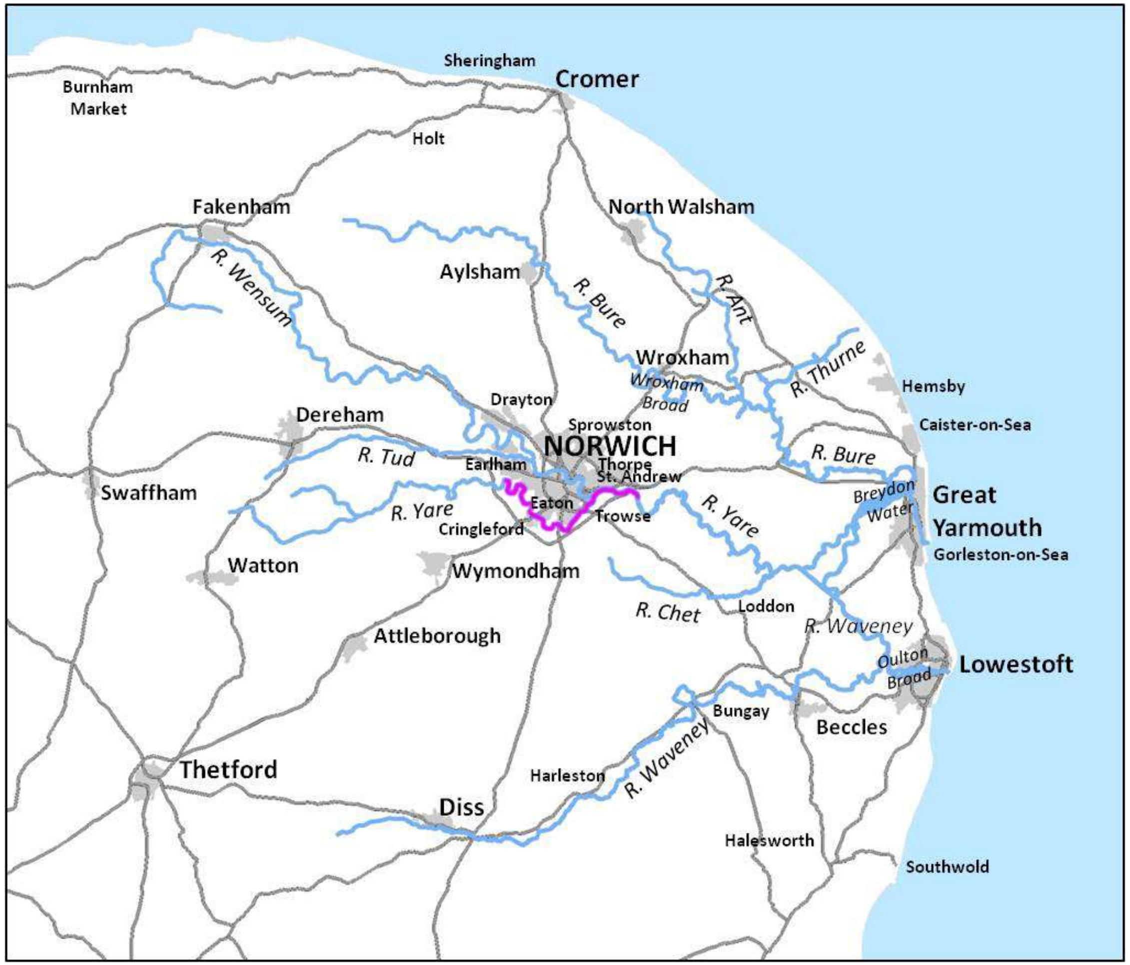
SHOWCARD 6a – Ecological quality ladder



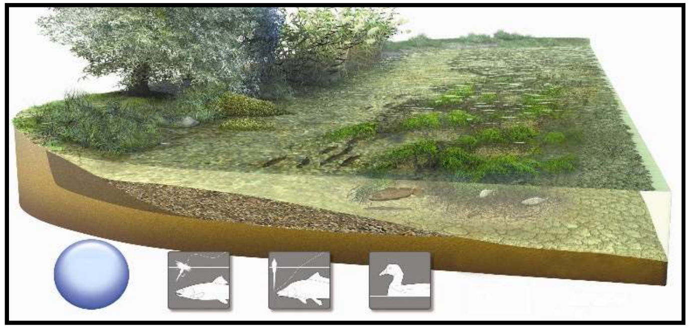
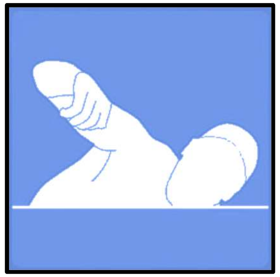
SHOWCARD 6b – Recreational quality ladder



SHOWCARD 7 – Survey river stretch



SHOWCARD 9 – Future visits

ECOLOGICAL QUALITY	RECREATIONAL QUALITY
	 

SHOWCARD 10 – Future river improvements

Code	Likelihood of improvements being made
A	Very likely
B	Somewhat likely
C	Neither likely or unlikely
D	Somewhat unlikely
E	Very unlikely

SHOWCARD 11 –Question difficulty

Code	Difficulty level
A	Very easy
B	Fairly easy
C	Neither easy or difficult
D	Fairly difficult
E	Very difficult

SHOWCARD 12 –Importance of issues

I thought this issue was...

Very Important	Important	Neither	Unimportant	Completely Unimportant
A	B	C	D	E

SHOWCARD 13 – Who should pay for water quality improvements

Improvements to water quality should be paid for by...

Strongly Agree	Tend to Agree	Neither	Tend to disagree	Strongly Disagree
A	B	C	D	E

SHOWCARD 14 – Question difficulty

Code	Difficulty level
A	Very easy
B	Fairly easy
C	Neither easy or difficult
D	Fairly difficult
E	Very difficult

SHOWCARD 15 – Ethnic background

Code	Ethnicity
A	White
B	Mixed/Multiple ethnic groups
C	Asian/Asian British
D	Black/African/Caribbean/Black British
E	Other

SHOWCARD 15a – Religion

Code	Religion
A	no religion
B	Christian
C	Buddhist
D	Hindu
E	Jewish
F	Muslim
G	Sikh
I	Other

SHOWCARD 16 – Highest educational qualification

Code	Qualification type
A	No qualifications
B	NVQ Level 1, Foundation GNVQ
C	1 - 4 O levels/CSEs/GCSEs (any grades), Foundation Diploma
D	NVQ Level 2, Intermediate GNVQ, City and Guilds Craft, BTEC First/ General Diploma, RSA Diploma
E	5+ O levels/CSEs (grade 1)/GCSEs (grades A*- C), Higher Diploma
F	NVQ Level 3, Advanced GNVQ, City and Guilds Advanced Craft, ONC, OND, BTEC National, RSA Advanced Diploma
G	2+ A levels/VCEs, 4+ AS levels, Higher School Certificate, Advanced Diploma
H	NVQ Level 4 - 5, HNC, HND, RSA Higher Diploma, BTEC Higher Level
I	Degree (for example BA, BSc)
J	Higher degree (for example MA, PhD, PGCE)
K	Professional qualification(e.g. teaching, nursing, accountancy)
L	Other (please state)

SHOWCARD 17 – Total household income before tax

Code	Annual income £s	Monthly income £s
A	less than 6,000	less than 500
B	6,001 – 12,000	501 – 1,000
C	12,001 – 18,000	1,001 – 1,500
D	18,001 – 24,000	1,501 – 2,000
E	24,001 – 30,000	2,001 – 2,500
F	30,000 – 36,000	2,501 – 3,000
G	36,001 – 42,000	3,001 – 3,500
H	42,001 – 48,000	3,501 – 4,000
I	48,001 – 54,000	4,001 – 4,500
J	54,001 – 60,000	4,501 – 5,000
K	60,001 – 66,000	5,001 – 5,500
L	66,001 – 72,000	5,501 – 6,000
M	Over 72,001	over 6,001

SHOWCARD 18 –Organisation memberships

Code	Membership type
A	Religious, or faith group
B	School fundraising group / PTA / School Governors
C	Scouts, Guides, cadets, etc.
D	Lions club / Rotary club / other community volunteering group
E	Walking club / Ramblers Association
F	Fishing / Angling club
G	Rowing / Canoeing club
H	National Trust / RSPB / English Nature
I	Greenpeace, Friends of the Earth, WWF, other environmental group
J	Climbing club
K	Women's Institute
L	Not a member of any similar organisations
M	Other (please state)

SHOWCARD 19 – Employment Status

Code	Employment type
A	Self Employed
B	Employed full-time (more than 30 hours per week)
C	Employed part-time (less than 30 hours per week)
D	Student
E	Unemployed – seeking employment
F	Unemployed – other
G	Looking after the home /children
H	Retired
I	Unable to work due to sickness or disability
J	Other

SHOWCARD 20 – Recreational craft owned

Code	Craft type
A	Canoe / Rowing boat
B	Narrowboat / Widebeam / Cruiser
C	Yacht
D	Other (please state)

Appendix IV: CL model on the pilot data

conditional logit model on price ecological quality recreational quality

Conditional (fixed-effects) logistic regression Number of obs. =480

LR chi2(3) = 85.03

Prob. > chi2 = 0.00

Log likelihood = -123.84 Pseudo R2 = 0.2556

Choice	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
price	-.008	.005	-1.64	0.102	-.017	.002
ecological quality	.725	.110	6.59	0.000	.509	.941
recreational quality	.646	.120	5.39	0.000	.411	.881

Appendix V: chronology of CL modelling development

This section reports a chronology of CL modelling development. As with the pilot data, the CL modelling of the main survey data was undertaken using Stata 13.1 (StataCorp L. P., 2013). Table 51, on the following page, shows the results of four CL models, each of which evolve towards an increasingly complex modelling solution. Descriptions and definitions of these variables are shown below.

Table 50: variables used within preliminary econometric modelling

Dependent variables	
Price	Respondents' response to cost of water quality, expressed as a continuous variable.
EQ	Ecological water quality, expressed as a continuous variable.
Yellow ecological quality	Yellow ecological water quality, expressed as a categorical variable. 1=yellow ecological water level, 0=other ecological levels.
Green ecological quality	Green ecological water quality, expressed as a categorical variable. 1=green ecological water level, 0=other ecological levels.
Medium ecological quality	Yellow and green ecological quality categories combined, expressed as a categorical variable. 1=yellow and green ecological water levels, 0=other ecological levels.
Blue/High ecological quality	Blue/High ecological water quality, expressed as a categorical variable. 1=blue/high ecological water level, 0=other ecological levels
RQ	Recreational water quality, expressed as a continuous variable.
Medium recreational quality	Medium recreational water quality, expressed as a categorical variable. 1=medium recreational water level, 0=other recreational levels.
High recreational quality	High recreational water quality, expressed as a categorical variable. 1=high recreational water level, 0=other recreational levels.
RQ*EQ	Variable describing the interaction between recreational and ecological water quality, expressed as a continuous variable.
Independent variables	
Swimmers	Respondent who is an active river swimmer, recruited via Tri-Anglia Triathlon Club. Binary variable: 0=respondent is not a swimmer, 1=respondent is a swimmer.
Rowers	Respondents who are rowers, recruited via rowing clubs. Binary variable: 0=respondent is not a rower, 1=respondent is a rower.
Anglers	Respondents who are anglers. Binary variable expressed as 0=respondent is not an angler, 1=respondent is an angler.
EnvMemberCont	The total number of environmental organisation memberships held by the respondent, expressed as a continuous variable.
DistanceBin	The distance the respondent lives from the closest point of the Yare. Binary variable: 0=respondent lives <8km, 1=the respondent lives >8km (Mean distance=7.99km).

Model 1 has the same specification as the CL model performed on the pilot survey data, to check whether anomalies existed within the main survey data. Model 2 offers a refinement in that, rather than treat ecological and recreational water quality as continuous variables, it uses separate coefficients to represent the different levels of ecological and recreational water quality. This enabled a better understanding of respondents' preferences, e.g. how they distinguished between the different attribute levels. Model 3 goes further as it examines an interaction within respondents' preferences when improved levels of ecological and recreational water quality are simultaneously available. The final preliminary model, Model 4, incorporates socio-economic variables, identifies and isolates different user types and reports their preferences as separate variables. Each model is now discussed.

Table 51: preliminary CL models

	Model 1		Model 2		Model 3		Model 4	
Variable	**Coef.**	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**	**Coef.**	**s.e.**
Baseline Coefficients								
Price	-.018***	.002	-.012***	.002	-.020***	.002	-0.020***	.003
EQ	.929***	.039						
RQ	.601***	.040						
Yellow ecological quality			1.781***	.101				
Green ecological quality			1.872***	.103				
Medium ecological quality					1.601***	.112	1.395***	.148
Blue/High ecological quality			3.072***	.124	2.717***	.169	2.546***	.223
Medium recreational quality			.775***	.080	.578***	.102	0.590***	.138
High recreational quality			1.264***	.083	.911***	.145	0.910**	.178
RQ*EQ					.121***	.042	0.141***	.044
Socio-Economic Coefficients								
Swimmers								
SwimmerxPrice							-0.019	.020
SwimmerxMedium ecological quality							-1.612***	.616
SwimmerxHigh ecological quality							-3.152***	.971
SwimmerxMedium recreational quality							0.964	.747
SwimmerxHigh recreational quality							2.405**	1.191
Anglers								
AnglersxPrice							-0.008	.008
AnglersxMedium ecological quality							0.840**	.425
AnglersxHigh ecological quality							1.678***	.633
AnglersxMedium recreational quality							-0.721*	.382
AnglersxHigh recreational quality							-0.870**	.386
Rowers								
RowerxPrice							-0.006	.012
RowerxMedium ecological quality							-0.771*	.465
RowerxHigh ecological quality							-0.862	.801
RowerxMedium recreational quality							1.218**	.504
RowerxHigh recreational quality							1.290**	.598
Environmental Memberships (continuous variable)								
EnvMemberContxPrice							0.005	.003
EnvMemberContxMedium ecological quality							0.701***	.171
EnvMemberContxHigh ecological quality							0.908***	.241
EnvMemberContxMedium recreational quality							0.095	.139
EnvMemberContxHigh recreational quality							0.314**	.157
Distance from river (binary 0=closer than 8km, 1=further than 8km)								
DistanceBinxPrice							-0.008*	.004
DistanceBinxMedium ecological quality							0.006	.193
DistanceBinxHigh ecological quality							-0.374	.275
DistanceBinxMedium recreational quality							-0.113	.175
DistanceBinxHigh recreational quality							-0.444**	.186
Pseudo R^2	0.322		0.354		0.356		0.396	
Log Likelihood	-1127.501		-1074.67		-1070.782		-1004.083	

Model 1: Simple CL model of Price, EQ and RQ. Model 2: CL model of Price, with ecological and recreational water quality attributes split into categorical levels. Model 3: CL model of Price with categorical levels of ecological and recreational water quality for main effects, and a continuous interaction term (RQ*EQ) between ecological and recreational water quality. Model 4: CL model of price, Categorical water quality levels, EQ*RQ interaction term, with anglers, swimmers, rowers, environmental membership and distance as categorical covariates. Note *, ** and *** = significance at 10%, 5% and 1% levels

The results of the simple CL analysis of price, ecological and recreational water quality reported in Model 1, correspond with the pattern of results seen in the pilot study data. As expected, because of the increased sample size, the confidence intervals for the three variables have narrowed and the estimated coefficients are more narrowly defined. Importantly, along with the coefficients for ecological and recreational water quality, the coefficient for price is now highly significant. The coefficients of Model 1 confirm that respondents are less likely to choose an alternative with increased price, *ceteris paribus,* and respondents have higher preferences for ecological, rather than recreational, water quality. This simple model assumes that the water quality variables, defined within the survey, are continuous. In reality water quality characteristics span a continuum but, for simplicity, within the experiment water quality was defined by categorical levels. The water quality data should, more accurately, be analysed as categorical levels and this refinement is explored in Model 2.

Model 2 verifies whether respondents' preferences for ecological and recreational water quality present non-linear effects and identifies respondents' preferences for individual water quality levels. The log likelihood of Model 2 (-1074.67) shows an improvement over that of Model 1 (-1127.501). Respondents clearly prefer improved water quality as both ecological and recreational water quality levels have an increased likelihood of being chosen as their quality improves. Respondents continue to prefer to choose choice options which contain lower price. All variables in Model 2 are highly significant.

This categorical model has overlapping confidence intervals for Yellow and Green ecological water quality levels. A Wald test on their coefficients (Prob. > chi2 = 0.9398), confirmed that they are insignificantly different from one another: respondents did not differentiate between Yellow and Green ecological categories when they made their choice decisions. This is in contrast to the confidence intervals of the Yellow and Green coefficients of the CL model on ChREAM data, which do not overlap (see Table 25). The standard errors of the water quality coefficients in the ChREAM model are much smaller (Yellow s.e. 0.012, Green s.e. 0.011, Blue s.e. 0.012) and the ChREAM model's confidence intervals are more narrowly defined. This is almost certainly due to the larger sample size (1100 respondents) used for the ChREAM model. It is possible that, if the sample used for this present research was of a comparable size, that larger sample may reduce noise in the error term and produce narrower confidence intervals, resulting in significant differences between coefficients for Yellow and Green ecological water quality levels.

It is important to remember that the ChREAM water quality levels conflate ecological and recreational attributes into a single attribute of preference, whereas, in this present research, ecological and recreational attributes are disaggregated. An alternative explanation for the insignificant difference in respondents' preferences for Yellow and Green ecological water quality levels may be that, by removing the recreational attribute inherent within the ChREAM water quality levels[28], the actual differences between the Yellow and Green ecological quality levels depicted here may have been diminished to the point where respondents feel that the differences that remain are insignificant.

Further cross-referencing of Model 2 against the ChREAM model suggests that the specification of Model 2 (and the efficiency of the CE design underpinning the choice data) is essentially sound: respondents' preferences for both ecological and recreational water quality attributes are complete and transitive, as we would expect them to be. Despite this, it may not be worthwhile excessively comparing the strength of the coefficients of Model 2's ecological water quality levels against the ChREAM coefficients for water quality levels simply because, being disaggregated, the attributes represented by the ecological water quality levels are now qualitatively different: some of the variation in respondents' preferences is now undoubtedly contained within the coefficients for recreational water quality levels.

A parsimonious solution to the insignificant difference between Yellow and Green ecological water quality levels was to collapse them into one intermediate variable, Medium, which is used within models 3 and 4. For clarity and conformity Blue ecological quality is renamed High. In terms of the log likelihood, Model 3, LL -1070.782, represents a small improvement over Model 2. As with Model 2,

[28] Apart from the differences in ecological quality, ChREAM's Yellow water quality is suitable for boating, but not swimming, and the Green level used in ChREAM is suitable for boating and swimming. These differences in recreational quality add to the overall differences between Yellow and Green water quality levels and to the distinctions respondents can make between them.

we see that respondents continue to have consistent preferences for water quality, preferring to choose higher levels of ecological and recreational quality, where those options are available. Respondents continue to avoid choice options which have increased price, *ceteris paribus*.

Model 3 includes another significant variable, RQ*EQ, which describes a significant positive interaction when respondents choose options containing the higher categorical levels of ecological and recreational quality. Respondents are significantly more likely to choose options with higher levels of recreational quality (or ecological quality) if higher levels of ecological quality (or recreational quality) are also available. This interaction variable treats recreational quality and ecological quality as continuous rather than categorical. Different specifications of RQ*EQ interaction variables were explored, e.g. using categorical levels of recreational quality and ecological quality, but their usefulness was slight and their interpretation problematic. The effect of the RQ*EQ variable on respondents' WTP is discussed below.

Given the objectives of disentangling and quantifying different types of respondents' preferences for ecological or recreational water quality, one of the optimal CL models, Model 4, also uses Price, ecological and recreational water quality levels and RQ*EQ coefficients. However, within Model 4, these six coefficients are baseline variables that represent respondents who have no environmental memberships, live within 8km of the survey stretch on the Yare and have not used the river for primary or secondary recreational activities (e.g. swimming, boating, fishing) over the last year, although they may have visited for other purposes. Model 4 also contains three groups of coefficients used to distinguish the preferences of the three primary recreational users; anglers, swimmers and rowers. There are also two other groups of socio economic variables used in Model 4. These model the impact that the number of environmental memberships held by respondents and the distance respondents live from the Yare, have on respondents' preferences. The additional groups of socio-economic interaction variables modify the baseline variables, depending on the socio-economic and use characteristics of the respondents. All of the coefficients are now discussed in greater detail.

The sign, strength and significance of the baseline coefficients in Model 4 reasonably describe the preferences for water quality held by a baseline respondent.

Several of the groups of socio-economic variables produce coefficients with wide confidence intervals, primarily due to small group size. Consequently some of the socio-economic coefficients in Model 4 are insignificant and cannot be taken at face value. However, these insignificant variables are useful as a guide to identify trends in the data. Despite this caveat, many of the socio-economic variables *are* significant, particularly those which have the highest utility for each of the different types of recreational user. It is important to remember that these socio-economic variables are interaction terms which modify the baseline coefficients, and, with this in mind, they become very useful.

In common with baseline respondents, swimmers' utility decreases as Price increases. Swimmers are more likely to choose an option with High recreational quality, than the other three water quality states. Both interaction terms for ecological water quality are negative, which appear reasonable, given that swimmers have primary contact with the water and would be expected to prefer swimmable water quality.

The net result of modifying the baseline coefficient for High ecological quality, to account for the preferences of the average swimmer, produces a negative coefficient for High ecological quality, e.g.

2.546 + -3.152 = -0.606

It is important to consider that the High ecological quality coefficient for a swimmer becomes further modified by the environmental membership and distance variables. All of the swimmers who were interviewed were members of at least one environmental organisation and, because of this, their coefficient for High ecological quality becomes further modified. Interactions between multiple socio-economic variables, and the effect these interactions have on the WTP estimates of respondents', is discussed further below.

Rowers exhibit similar choice behaviour to that of the swimmers. Rowers have significant, positive preferences for recreational water quality. There is an insignificant difference in preferences between Medium and High recreational quality coefficients (Prob. > chi2 = 0.8630). It seems that rowers, in common with swimmers, choose the recreational quality level most appropriate to their needs as Medium recreational quality is sufficient for rowers to enjoy their activities safely.

Anglers are significantly more likely to choose higher levels of ecological quality, when the option to do so is available. Interestingly, anglers have negative preferences for both recreational quality levels. A Wald test (Prob. > chi^2 = 0.7286) confirms that anglers' preferences for recreational water quality are insignificantly different from one another. All of these coefficient modifiers make sense when viewed from an angler's perspective. Anglers require rivers with Medium ecological quality for coarse fishing and High ecological quality for game fishing. In interview, several anglers stated preferences for quiet locations where fish are undisturbed by swimmers, rowers or other human interference. It is quite reasonable that anglers choose water at lower recreational quality, to inhibit swimming and boating.

During the interview process respondents were asked if they held personal memberships for any environmental organisations. Examples include the National Trust, Royal Society for the Protection of Birds (RSPB), English Nature, Norfolk Wildlife Trust and environmental recreation clubs. The environmental membership variable is coded as a continuous variable. Having at least one environmental membership has a highly significant positive effect on the respondent's probability of choosing a choice option with higher levels of ecological quality. A Wald test (Prob. > chi^2 = 0.1331) confirms that these respondents do not distinguish between the two ecological water quality states, but it is clear that respondents holding environmental memberships have significant preferences for higher ecological water quality. These respondents are also significantly more likely to choose options with High recreational water quality.

The distance variable is defined by the distance the respondent lives from their home to the closest point of the survey river stretch. Within Model 4, the distance variable captures a distance threshold, at 8km, where respondents who live further from the environmental improvement are less likely to choose an option with higher prices to pay for that improvements. Distance is a binary variable where respondents who lived within the mean distance of 7.99km from the Yare were coded as 0 and respondents who lived further than the mean were coded as 1[29]. Respondents who live further away from the river are slightly less likely, to choose an option with increased price (-0.008* (s.e. 0.004)). This coefficient is close to the 5% significance level at p=0.053. Respondents who live further away from the river are also significantly less likely to choose options with High recreational water quality. The distance variable has an insignificant effect on the remaining three water quality variables but the sign of the coefficients on these variables suggest that respondents are less likely to choose higher levels of water quality, as distance increases. Distance was also analysed as continuous and log transformed variables, but was insignificant when expressed in either of those forms.

A selection of other socio-economic variables collected during the survey, e.g. age, gender, education, were assessed for their suitability as explanatory variables but these were insignificant determinants of choice behaviour.

Experts were treated as a discrete group of interest during data collection but during the analysis treating experts as an independent group produced no significant results. However, experts were selected from a range of disciplines and may have confounding motivations for water quality improvements, which may explain why they proved insignificant as a defining group.

Marginal willingness to pay estimates derived from preliminary CL models

This section reports the marginal WTP estimates for changes in attribute levels in CL models 3 and 4. The marginal WTP estimates are the negative of the ratio between the mean coefficients for each attribute and the mean coefficient of the payment attribute (please see Equation 16, p.197). The marginal WTP estimates, shown on Table 52, are derived from Model 3. This model is parsimonious, as,

[29] A binary variable using median distance was explored but, due to the skewed distance respondents lived from the survey river (Figure 25), this variable was insignificant. Distance is further refined and expressed as the inverse multiplicative ($1/x$) in CL models 5 and 6.

using only six variables, it obtains highly significant WTP estimates which are useful to examine the overall trends in the data.

Table 52: marginal WTP estimates derived from Model 3

	Medium ecological quality	High ecological quality	Medium recreational quality	High recreational quality	RQ*EQ
WTP for all respondents (200 respondents)					
WTP (£)	£78.72***	£133.61***	£28.44***	£44.80***	£5.92***
95% confidence intervals					
Lower limit	£62.37	£107.96	£17.51	£29.43	£1.90
Upper limit	£95.07	£159.26	£39.37	£60.17	£9.96

WTP=£, per household, per year for the 20km survey river stretch. Low water quality defines V^0, the baseline. Note *, ** and *** = significance at 10%, 5% and 1% levels.

Table 52 provides marginal WTP estimates for improvements in water quality from Low ecological or recreational quality (the baseline, V^0) to either Medium or High levels of those water quality attributes. Respondents have significantly higher WTP for ecological quality, rather than recreational quality. As discussed previously, the variable RQ*EQ reflects the respondent's preferences to choose options with higher categorical levels of water quality simultaneously. Within Model 3, we see that RQ*EQ is highly significant and the effect this coefficient has on respondents' WTP is shown in Table 53.

Table 53: additional WTP derived from the interaction effect of RQ*EQ in Model 3

Water quality (level)	Low ecological quality (0)	Medium ecological quality (1)	High ecological quality (2)
Low recreational quality (0)	0	0	0
Medium recreational quality (1)	0	£5.92	£11.84
High recreational quality (2)	0	£11.84	£23.68

WTP=£, per household, per year for the 20km survey river stretch. Low water quality defines V^0, the baseline.

We see that where a respondent is able to choose Medium levels of ecological and recreational quality together they are willing to pay an additional £5.92 on their water bill. Where they are able to choose High ecological quality with Medium recreational quality (or *vice versa*) they are willing to pay an additional £11.84 for improved water quality. Where they can choose High levels of both ecological and recreational water quality, they are willing to pay an additional £23.68 towards water quality improvements at the survey stretch.

Although Model 3 is parsimonious and highly significant, it may be argued that the WTP estimates derived from Model 3 do not adequately represent the general population as the sample contains disproportionately high numbers of rowers, swimmers and experts, which may skew the mean coefficient estimates. Hence the development of Model 4, which identifies and distinguishes recreational users from the general population. Unfortunately, due to the non-probabilistic sampling scheme, it cannot be stated with certainty that the coefficients of Model 4 are actually more representative of the wider population. Despite these criticisms, these models do achieve one of the primary objectives of this research: to disaggregate the values respondents hold for different attributes of water quality.

The marginal WTP estimates for different water quality levels, user groups and socio-economic variables, derived from Model 4, are reported in Table 54.

Table 54: marginal WTP estimates derived from Model 4

	Medium ecological quality	High ecological quality	Medium recreational quality	High recreational quality	RQ*EQ
	WTP (£, per household, per year for the 20km survey river stretch)				
Baseline, no primary or secondary river contact (178 respondents)					
WTP < 8km	70.32*** (11.317)	128.30*** (18.651)	29.71*** (7.864)	45.86*** (10.147)	7.10*** (2.392)
WTP > 8km	50.61*** (7.791)	78.44*** (11.097)	17.21*** (5.806)	16.84** (7.068)	5.09*** (1.666)
Anglers (16 respondents)					
WTP < 8km	80.78*** (23.371)	152.62*** (37.279)	-4.74 (13.598)	1.43 (14.337)	
WTP > 8km	63.10*** (16.956)	108.37*** (24.720)	-6.87 (11.419)	-11.38 (12.889)	
Swimmers (5 respondents)					
WTP < 8km	-5.59 (16.966)	-15.65 (29.044)	40.11** (18.780)	85.56*** (28.876)	
WTP > 8km	-4.52 (14.488)	-21.03 (26.255)	30.93** (14.627)	61.62*** (19.174)	
Rowers (10 respondents)					
WTP < 8km	23.86 (20.683)	64.34 (41.851)	69.06* (36.993)	84.05** (42.453)	
WTP > 8km	18.54 (15.575)	38.51 (27.371)	49.80** (23.380)	51.61** (24.178)	
Baseline respondent holding 1 environmental membership					
WTP < 8km	141.70*** (31.795)	233.45*** (50.285)	46.29*** (15.196)	82.74*** (21.880)	
WTP > 8km	92.86*** (16.615)	136.03*** (22.890)	25.26*** (8.958)	34.47*** (10.675)	

Standard errors in parenthesis. *, ** and *** = significance at 10%, 5% and 1% levels. Low water quality defines V^0, the baseline.

The mean WTP values for the baseline respondents who live within 8km of the Yare, who do not hold environmental memberships or use the river for recreation, are broadly similar to the mean WTP values derived from Model 3. Among the baseline respondents living closer to the river we see a clear willingness to pay more for the ecological quality of the survey river and consistent willingness to pay more to obtain the highest water quality of both water quality types.

As discussed previously, distance has a significant effect on respondents' probability of choosing an option which, in turn, leads to a highly significant decrease in respondents' WTP for all water quality types if they live further than 8km from the river. These respondents continue to have consistent preferences in their WTP for higher ecological water quality as ecological quality improves, but appear to be unwilling to pay more for High recreational water quality. As discussed previously, baseline respondents tend to be more ambivalent about recreational quality and, it would appear, they become more ambivalent about High recreational quality as they live further away from the river.

The additional explanatory variables in Model 4 tend to reduce the central WTP results on the common, baseline, variables in Table 54. We find that the variation has moved into the user and socio-economic variables. Several of the user groups produce very wide estimates due to small group size. Consequently, some of the WTP estimates reported in Table 54 are insignificantly different from zero and cannot be taken at face value - but they are useful as a guide to identifying trends in the data. The variables of interest relating to the different user groups are significant, e.g. anglers prefer ecological quality and their WTP for ecological quality is significant, rowers and swimmers prefer RQ and their WTP for RQ is significant.

Of the three user groups, anglers living close to the river have the highest WTP for both Medium ecological quality and High ecological water quality. In contrast, for the reasons discussed previously, anglers have the lowest mean WTP for recreational water quality. The WTP values for recreational water quality for anglers are insignificantly different to zero.

Swimmers have the highest WTP for High recreational water quality. Swimmers' WTP for ecological quality is the lowest of all respondent types. Although the WTP values held by swimmers for ecological water quality are not significantly different from zero and have very wide confidence intervals, the values do suggest that swimmers would prefer water quality investments to be directed towards improving the recreational quality of water. It is important to note that the correct value of swimmers' WTP for ecological quality may not be negative as all of the swimmers who were interviewed held at least one environmental membership, which positively modifies their willingness to pay for ecological quality. For example, a swimmer, holding one environmental membership and living within 8km of the river has a positive willingness to pay, £8.97 (s.e. 26.309), for High ecological water quality.

Like swimmers, rowers also have higher willingness to pay for recreational water quality. Rowers have the highest WTP, of all respondent types, for Medium recreational water quality. This is, as discussed previously, the water quality type from which they gain the greatest utility. Rowers' WTP for High recreational water quality is slightly lower than that of the swimmers. Like swimmers, rowers WTP for ecological water quality is less than the values held by anglers or baseline respondents.

Having one (or more) membership(s) of an environmental organisation has a large impact on the baseline respondents' mean willingness to pay for both water quality types. In Model 4, the number of environmental memberships held by respondents is specified as a continuous variable. However, we would not expect environmental memberships to have a linear relationship with WTP. To prevent this issue, Model 5 contains a respecified, binary, variable to describe the effect of Environmental Membership where 0=no memberships and 1=one or more memberships.

Within Model 4, RQ*EQ varies depending on whether the respondent lives within, or further than, 8km of the survey river. For respondents living closer to the river the effect is £7.10*** (s.e. 2.392), and £5.09*** (s.e. 1.666) for respondents living farther away. The effects of the RQ*EQ interactions on different combinations of RQ and ecological quality are further described in Table 55.

Table 55: additional WTP derived from the interaction effect of RQ*EQ in Model 4

Water quality (level)	Low ecological quality (0)	Medium ecological quality (1)	High ecological quality (2)
	WTP (£, per household, per year for the 20km survey river stretch)		
Respondents living < 8km			
Low recreational quality (0)	0	0	0
Medium recreational quality (1)	0	£7.10***	£14.20***
High recreational quality (2)	0	£14.20***	£28.40***
Respondents living > 8km			
Low recreational quality (0)	0	0	0
Medium recreational quality (1)	0	£5.09***	£10.18***
High recreational quality (2)	0	£10.18***	£20.36***

Low water quality defines V^0, the baseline.

References

Adamowicz, W. L. (2004). What's it worth? An examination of historical trends and future directions in environmental valuation. *Australian Journal of Agricultural and Resource Economics*, 48(3), pp. 419–443.

ADAS. (2016). What does Brexit mean for the food and farming sector? [online] Available at: http://www.adas.uk/News/what-does-brexit-mean-for-the-food-and-farming-sector-21 [Accessed 23 July 2016].

Agbotui, P.Y., Kportufe, G.S., Braimah, S. and Brookman-Amissah, M. (2014). The Chemical Composition of the River Aire and Anthropogenic Impacts on the River from its Source to Leeds. *Journal of Environment and Earth Science*, 4(10), pp.100-107.

Aitken, M. N. (2003). Impact of agricultural practices and river catchment characteristics on river and bathing water quality. *Water Science and Technology*, 48(10), pp. 217-224.

Aldrich, G. A., Grimsrud, K. M., Thacher, J. A. and Kotchen, M. J. (2007). Relating environmental attitudes and contingent values: how robust are methods for identifying preference heterogeneity? *Environmental and Resource Economics*, 37(4), pp. 757–775.

Allenby, G. M. (1990). Hypothesis testing with scanner data: the advantage of Bayesian methods. *Journal of Marketing Research*, 27(4), pp. 379–389.

Andreopoulos, D., Damigos, D., Comiti, F. and Fischer, C. (2014). Public Preferences for Climate Change Adaptation Policies in Greece: A Choice Experiment Application on River Uses. In: C. Zopounidis, N. Kalogeras, K. Mattas, G. van Dijk and G. Baourakis, eds., *Agricultural Cooperative Management and Policy: New Robust, Reliable and Coherent Modelling Tools,* 1st ed. Switzerland: Springer International Publishing, pp. 163–178.

Andrews, K. and Zabel, T. (2003). Achievements of Co-operative Agreements. In: F. Brouwer, I. Heinz and T. Zabel, eds., *Governance of Water Related Conflicts in Agriculture,* 1st ed. Dordrecht: Kluwer Academic Publishers, pp. 45-66.

Appleton, T. (2008). More about the Humber River Basin District. [online] Environment Agency. Available at: http://www.environment-agency.gov.uk/subjects/waterquality/955573/1458449/1458624/1462818/?version=1&lang=_e [Accessed 20 May 2008].

Atkinson, G., Machado, F. and Mourato, S. (2000). Balancing competing principles of environmental equity. *Environment and Planning A*, 32(10), pp. 1791-1806.

Bailar, B., Bailey, L. and Stevens, J. (1977). Measures of interviewer bias and variance. *Journal of Marketing Research*, 14(3), pp. 337–343.

Balana, B. B., Vinten, A. and Slee, B. (2011). A review on cost-effectiveness analysis of agri-environmental measures related to the EU WFD: Key issues, methods, and applications. *Ecological Economics*, 70(6), pp. 1021–1031.

Bann, C., Fisher, J. and Horton, B. (2003). The Benefits Assessment Guidance for PRO4: Review of Non-use Values for Water Quality and Water Resources and Values for Bathing Water Improvements. In: Environment Agency, *Report of an Environment Agency Expert Workshop in Peterborough, UK., May 2003*. Bristol: Environment Agency.

Barbier, E. B. (1993). Sustainable use of wetlands valuing tropical wetland benefits: economic methodologies and applications. *Geographical Journal*, 159(1), pp. 22–32.

Bateman, I. J. and Turner, R. K.. (1993). Valuation of the environment, methods and techniques: the contingent valuation method. In: R. K. Turner, ed., *Sustainable environmental economics and management: principles and practice.* London: Belhaven Press, pp.120-191.

Bateman, I. J., Turner, R. K. and Bateman, S. (1993). Extending cost benefit analysis of UK highway proposals: environmental evaluation and equity. *Project Appraisal*, 8(4), pp. 213–224.

Bateman, I. J., Willis, K. G. and Arrow, K. J. (2001). *Valuing environmental preferences: theory and practice of the contingent valuation method in the US, EU, and developing countries*. Oxford: Oxford University Press.

Bateman, I. J., Carson, R. T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Özdemiroglu, E. and Pearce, D. W., Sugden, R. and Swanson, J. (2002). *Economic valuation with stated preference techniques: a manual.* Cheltenham: Edward Elgar Publishing.

Bateman, I. J., Brouwer, R., Davies, H., Day, B. H., Deflandre, A., Falco, S. D., Georgiou, S., Hadley, D., Hutchins, M., Jones, A. P., Kay, D., Leeks, G., Lewis, M., Lovett, A. A., Neal, C., Posen, P., Rigby, D. and Turner, R. K. (2006a). Analysing the Agricultural Costs and Non-market Benefits of Implementing the Water Framework Directive. *Journal of Agricultural Economics*, 57(2), pp. 221-237.

Bateman, I. J., Day, B. H., Georgiou, S. and Lake, I. (2006b). The aggregation of environmental benefit values: welfare measures, distance decay and total WTP. *Ecological Economics*, 60(2), pp. 450–460.

Bateman, I. J., Ferrini, S. and Hime, S. (2008). *Aquamoney: UK case study report.* Norwich: The Centre for Social and Economic Research on the Environment, University of East Anglia.

Bateman, I. J., Brouwer, R., Hime, S. and Provins, A. (2009). *Development of Guidelines for Use of Benefits Transfer in Policy and Project Appraisal. Interim Report to the Department for Environment, Food and Rural Affairs, 17th February 2009.* London: Economics for the Environment Consultancy.

Bateman, I. J., Brouwer, R., Ferrini, S., Schaafsma, M., Barton, D. N., Dubgaard, A., Hasler, B., Hime, S., Liekens, I., Navrud, S. and De Nocker, L. (2011). Making benefit transfers work: deriving and testing principles for value transfers for similar and dissimilar sites using a case study of the non-market benefits of water quality improvements across Europe. *Environmental and Resource Economics*, 50(3), pp. 365–387.

BBC. (2008). Average water bill rising by 5.8%. [online] BBC News. Available at: http://news.bbc.co.uk/go/pr/fr/-/1/hi/business/7264594.stm [Accessed 1 March 2008].

BBC. (2016). Reality Check: Your questions about farming and fishing. [online] BBC News. Available at: http://www.bbc.co.uk/news/uk-politics-eu-referendum-36552642 [Accessed 19 June 2016].

Ben-Akiva, M. and Swait, J. (1986). The Akaike likelihood ratio index. *Transportation Science*, 20(2), pp. 133–136.

Bennett, J. and Blamey, R. (2001). *The choice modelling approach to environmental valuation*. Cheltenham: Edward Elgar Publishing.

Bentley, M. (2014). Whitlingham Country Park. [online] South Norfolk Council. Available at: http://www.south-norfolk.gov.uk/leisure/1226.asp [Accessed 28 March 2014].

Bergstrom, J. C., Stoll, J. R. and Randall, A. (1989). Information effects in contingent markets. *American Journal of Agricultural Economics*, 71(3), pp. 685–691.

Bernhard, A. E. and Field, K. G. (2000). A PCR assay to discriminate human and ruminant feces on the basis of host differences in Bacteroides-Prevotella genes encoding 16S rRNA. *Applied and environmental microbiology*, 66(10), pp.4571-4574.

Bicknell, B. R., Imhoff, J. C., Kittle, J. L., Jr., Donigian, A. S., Jr. and Johanson, R. C. (1997). *Hydrological Simulation Program Fortran, User's manual for version 11. EPA/600/R-97/080.* Athens, G.A.: Environmental Protection Agency National Exposure Research Laboratory.

Binkley, C. S. and Hanemann, W. M. (1978). *The recreation benefits of water quality improvement: analysis of day trips in an urban setting. EPA-600/5-78-010.* Washington, D.C.: Environmental Protection Agency, Office of Research and Development, Office of Health and Ecological Effects.

Birol, E., Karousakis, K. and Koundouri, P. (2006). Using economic valuation techniques to inform water resources management: A survey and critical appraisal of available techniques and an application. *Science of The Total Environment*, 365(1–3), pp. 105–122.

Bishop, O. S. (1966). *Statistics for biology*. 3rd ed. Harlow: Longman.

Blacklocke, S., Hooper, A., Rosenberg, M. and Earle, R. (2006). Cost effective programmes of measures: Theory versus reality. In: L. Gairns, C. Crighton and W. A. Jeffrey, eds., *Agriculture and the Environment VI: Managing Rural Diffuse Pollution. Proceedings of the SAC and SEPA Biennial Conference, Edinburgh, 5-6 April 2006.* Edinburgh: International Water Association, Scottish Agricultural College and Scottish Environmental Protection Agency, pp. 71-76.

Bliem, M., Getzner, M. and Rodiga-Laßnig, P. (2012). Temporal stability of individual preferences for river restoration in Austria using a choice experiment. *Journal of Environmental Management*, 103, pp. 65–73.

Boardman, Anthony E., Greenburg, David H., Vining, Aidan R. and Weimer, David. (2014). *Cost-Benefit Analysis: Concepts and Practice.* 4th ed. Upper Saddle River: Prentice Hall, Inc.

Boorman, D. B., Hollis, J. M. and Lilly, A. (1995*). Hydrology of soil types: a hydrologically based classification of the soils of the United Kingdom. Institute of Hydrology Report No. 126.* Wallingford: Institute of Hydrology.

Booth, W., Carubia, P. and Lutz, F. (1976). *A methodology for comparative evaluation of water quality indices, Final Report NTIS PB-251 572.* Washington, D. C.: The Council on Environmental Quality.

Bouwes, N. W. and Schneider, R. (1979). Procedures in estimating benefits of water quality change. *American Journal of Agricultural Economics*, 61(3), pp. 535–539.

Boxall, P. C. and Adamowicz, W. L. (2002). Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach. *Environmental and Resource Economics*, 23(4), pp. 421–446.

Boxall, P. C., Adamowicz, W. L., Swait, J., Williams, M. and Louviere, J. (1996). A comparison of stated preference methods for environmental valuation. *Ecological Economics*, 18(3), pp. 243–253.

Bradley, M. and Daly, A. (1994). Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. *Transportation*, 21(2), pp.167-184.

Bradley, G. and Hancock, C. (2003). Increased risk of non-seasonal and body immersion recreational marine bathers contacting indicator microorganisms of sewage pollution. *Marine Pollution Bulletin*, 46(6), pp. 791-794.

Braga, J. and Starmer, C. (2005). Preference Anomalies, Preference Elicitation and the Discovered Preference Hypothesis. *Environmental and Resource Economics*, 32(1), pp. 55-89.

Brainard, J. S., Jones, A. P., Bateman, I. J., Lovett, A. A. and Fallon, P. A. (2002). Modelling environmental equity: access to air quality in Birmingham, England. *Environment and Planning A*, 34(4), pp. 695–716.

British Rowing. (2017). Governance. [online] Available at: https://www.britishrowing.org/about-us/governance/ [Accessed 14 August 2017].

Breusch, T. S. and Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, 47(1), pp. 239–253.

Brey, R., Riera, P. and Mogas, J. (2007). Estimation of forest values using choice modeling: an application to Spanish forests. *Ecological Economics*, 64(2), pp. 305–312.

British Water. (2016). What are some potential positives for British Water following Brexit? [online] Available at: http://www.britishwater.co.uk/article/qandabrexit-185.aspx [Accessed 28 July 2016].

Brouwer, R., Martin-Ortega, J. and Berbel, J. (2010). Spatial Preference Heterogeneity: A Choice Experiment. *Land Economics*, 86(3), pp. 552–568.

Brouwer, R., Martin-Ortega, J., Dekker, T., Sardonini, L., Andreu, J., Kontogianni, A., Skourtos, M., Raggi, M., Viaggi, D., Pulido-Velazquez, M., Rolfe, J. and Windle, J. (2015). Improving value transfer through socio-economic adjustments in a multicountry choice experiment of water conservation alternatives. *Australian Journal of Agricultural and Resource Economics*, 59(3), pp. 458–478.

Brown, S. R. (1980). *Political subjectivity: Applications of Q methodology in political science*. New Haven: Yale University Press.

Brownstone, D. (2001). *Discrete choice modeling for transportation*. Berkeley: University of California Transportation Center.

Burge, P. and Rohr, C. (2004). DATIV: SP Design: Proposed approach for pilot survey. *Tetra-Plan in cooperation with RAND Europe and Gallup A/S*.

Burns, C. (2016). The EU Referendum and the Environment. [online] Friends of the Earth. Available at: https://www.foe.co.uk/sites/default/files/downloads/eu-referendum-environment-81600.pdf [Accessed 28 July 2016].

Burton, G. A., Gunnison, D. and Lanza, G. R. (1987). Survival of pathogenic bacteria in various freshwater sediments. *Applied Environmental Microbiology*, 53(4), pp. 633-638.

CaBA. (2016). Aire and Calder. [online] Catchment Based Approach. Available at: http://www.catchmentbasedapproach.org/humber/aire-calder [Accessed 25 July 2016].

Cabelli, V. J. (1983). *Health effects criteria for marine recreational waters. EPA-600/11-80-031.* Cincinnati: Environmental Protection Agency.

Camargo, J. A. and Alonso, Á. (2006). Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: a global assessment. *Environment International*, 32(6), pp. 831–849.

Campbell, D., Hensher, D. A. and Scarpa, R. (2011). Non-attendance to attributes in environmental choice analysis: a latent class specification. *Journal of Environmental Planning and Management*, 54(8), pp. 1061–1076.

Campbell, D. and Lorimer, V. S. (2009). “Accommodating attribute processing strategies in stated choice analysis: do respondents do what they say they do.” Paper presented at Comunicación presentada a la XVII Conferencia anual de la

Asociación Europea de Economía Ambiental y de los Recursos Naturales, June, 2009, Amsterdam. pp. 24-27.

Carlsson, F., Kataria, M., Krupnick, A., Lampi, E., Löfgren, Å., Qin, P. and Sterner, T. (2013). The truth, the whole truth, and nothing but the truth—A multiple country test of an oath script. *Journal of Economic Behavior and Organization*, 89, pp. 105-121.

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N. and Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications*, 8(3), pp. 559–568.

Carr, R. and Bartrum, J. (2004). The control envelope and risk management. In: J. A. Cotruvo, A. Dufour, G. Rees, J. Bartram, R. Carr, D. O. Cliver, G. F. Craun, R. Fayer and V. P. Gannon, eds., *Waterborne zoonoses*. London: IWA Publishing, on behalf of the World Health Organisation.

Carson, R. T. (1991). Constructed markets. In: C. D. Kolstad and J. B. Braden, eds., *Measuring the demand for environmental quality,* 1st ed. North-Holland: Elsevier Science Publishing Company, pp. 121–162.

Carson, R. T. (2000). Contingent valuation: a user's guide. *Environmental Science and Technology*, 34(8), pp. 1413–1418.

Carson, R. T., Mitchell, R. C., Hanemann, M., Kopp, R. J., Presser, S. and Ruud, P. A. (2003). Contingent valuation and lost passive use: damages from the Exxon Valdez oil spill. *Environmental and Resource Economics*, 25(3), pp. 257–286.

Carson, R.T., Groves, T. and List, J.A. (2014). Consequentiality: A theoretical and experimental exploration of a single binary choice. *Journal of the Association of Environmental and Resource Economists*, 1(1/2), pp. 171-207.

CDC. (2004). Surveillance for waterborne-disease outbreaks associated with recreational water - United States, 2001-2002. [online] Centers for Disease Control and Prevention. Available online at: http://www.medscape.com/medline/abstract/15499306. [Accessed 27 September 2017].

CEH. (2008a). Land Cover Map 2000 - Calibration and accuracy. [online] Centre for Ecology and Hydrology. Available at: http://192.171.153.213/sections/seo/accuracy1.html [Accessed 9 July 2008].

CEH. (2008b). The National River Flow Archive. [online] Centre for Ecology and Hydrology. Available at: http://www.ceh.ac.uk/data/nrfa/index.html [Accessed 3 June 2008].

Chadwick, D., Fish, R., Oliver, D. M., Heathwaite, L., Hodgson, C. and Winter, M. (2008). Management of livestock and their manure to reduce the risk of microbial transfers to water - the case for an interdisciplinary approach. *Trends in Food Science and Technology*, 19(5), pp. 240-247.

Champ, P. A., Boyle, K. J. and Brown, T. C. (2012). *A primer on nonmarket valuation, Volume3*. New York: Springer Science and Business Media.

ChoiceMetrics PTY Ltd. (2012). *NGene* (version 1.1) [computer software]. Sydney: ChoiceMetrics PTY Ltd.

ChoiceMetrics PTY Ltd. (2012). NGene Handbook for Version 1.1. [online] Available at: http://www.choice-metrics.com [Accessed 24 April 2017].Chorus, I., Falconer, I. R., Salas, H. J. and Bartram, J. (2000). Health risks caused by freshwater cyanobacteria in recreational waters. *Journal of Toxicology and Environmental Health Part B: Critical Reviews*, 3(4), pp. 323–347.

Christie, M., Fazey, I., Cooper, R., Hyde, T., Deri, A., Hughes, L., Bush, G., Brander, L., Nahman, A., de Lange, W. and Reyers, B. (2008). *An Evaluation of Economic and Non-economic Techniques for Assessing the Importance of Biodiversity to People in Developing Countries. CR 0391. Final Report to Defra.* London: Defra.

Ciriacy-Wantrup, S. V. (1947). Capital returns from soil-conservation practices. *Journal of Farm Economics*, 29(4, Part II), pp. 1181–1196.

Clawson, M. (1959). *Methods of measuring the demand for and value of outdoor recreation*. Washington, D.C.: Resources for the Future.

Correll, D. L. (1998). The role of phosphorus in the eutrophication of receiving waters: A review. *Journal of Environmental Quality*, 27(2), pp. 261–266.

Country Land and Business Association Limited. (2016). Regulatory Implications under BREXIT. [online] Available at: https://www.cla.org.uk/sites/default/files/Audit%20of%20EU%20Regulations.pdf [Accessed 3 August 2016].

Cowell, R. and Owens, S. (2016). The EU referendum, planning and the environment. [online] Town and Country Planning Association. Available at: http://www.tcpa.org.uk/data/files/May_2016_Sample.pdf [Accessed 3 August 2016].

Crabtree, B., Hickman, M. and Martin, D. (1999). Integrated water quality and environmental cost-benefit modelling for the management of the River Tame. *Water Science and Technology*, 39(4), pp. 213-220.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L. and Hanson, W. E. (2003). Advanced mixed methods research designs. In: A. Tashakkori and C. Teddlie, eds., *Handbook of mixed methods in social and behavioural research*. Thousand Oaks, CA: Sage, pp. 209–240.

Crowther, J., Kay, D. and Wyer, M. D. (2001). Relationships between microbial water quality and environmental conditions in coastal recreational waters: the Fylde coast, UK. *Water Research*, 35(17), pp. 4029-4038.

Crowther, J., Kay, D. and Wyer, M. D. (2002). Faecal-indicator concentrations in waters draining lowland pastoral catchments in the UK: relationships with land use and farming practices. *Water Research*, 36(7), pp. 1725-1734.

Crowther, J., Wyer, M. D., Bradford, M., Kay, D. and Francis, C. A. (2003). Modelling faecal indicator concentrations in large rural catchments using land use and topographic data. *Journal of Applied Microbiology*, 94(6), pp. 962–973.

Crowther, J. (2008a). Modelling: some thoughts/issues. [email 30 May 2008].

Crowther, J. (2008b). RE: FIO model variables: Discharge data. [email 22 May 2008].

Crowther, J. (2008c). Humber Catchment Data Request. [email 26 February 2008].

Crowther, J., Hampson, D. I., Bateman, I. J., Kay, D., Posen, P. E., Stapleton, C. M. and Wyer, M. D. (2011). Generic modelling of faecal indicator organism concentrations in the UK. *Water*, 3(2), pp. 682–701.

Cumming, R. G., Brookshire, D. S. and Schultze, W. D. (1986). *Valuing public goods: the contingent valuation method*. Totowa, New Jersey: Rowman and Allanheld Publishers.

Cummings, R. G., Ganderton, P. T. and McGuckin, T. (1994). Substitution effects in CVM values. *American Journal of Agricultural Economics*, 76(2), pp. 205–214.

Cuttle, S. P., Macleod, C. J. A., Chadwick, D. R., Scholefield, D., Haygarth, P. M., Newell-Price, P., Harris, D., Shepherd, M. A., Chambers, B. J. and Humphrey, R. (2007). An inventory of methods to control diffuse water pollution from agriculture, Defra Project ES02 03. [online] Defra. Available at: http://www.defra.gov.uk/farm/environment/water/csf/pdf/UserManual_Jan07.pdf [Accessed May 24 2008].

Davies-Boren, Z. (2016). How Brexit could make it easier to frack in the UK. [online] Greenpeace. Available at: http://energydesk.greenpeace.org/2016/06/21/how-brexit-could-make-it-easier-to-frack-in-the-uk/ [Accessed 26 June 2016].

Day, B., Bateman, I. J., Carson, R. T., Dupont, D., Louviere, J. J., Morimoto, S., Scarpa, R. and Wang, P. (2012). Ordering effects and choice set awareness in repeat-response stated preference studies. *Journal of Environmental Economics and Management*, 63(1), pp. 73–91.

Deeks, L. K., McHugh, M. and Owens, P. N. (2005). Faecal contamination of watercourses from farm waste disposal for three sites in the UK with contrasting soil types. *Soil Use and Management*, 21(2), pp. 212-220.

Defra. (2004). *Cost curve of nitrate mitigation options, Defra report number NT2511*. Devon: Institute of Grassland and Environmental Research.

Defra. (2008a). Water Quality - EC Bathing Water Directive. [online] Defra. Available at: http://www.defra.gov.uk/environment/water/quality/bathing/default.htm#guideline [Accessed 12 April 2008].

Defra. (2008b). Bathing Waters (England) Notice 2008 (Draft Version). [online] Defra. Available at: http://www.defra.gov.uk/environment/water/quality/bathing/pdf/bw-england2008-notice.pdf [Accessed 17 July 2008].

Defra. (2008c). *Overall Impact assessment for the Water Framework Directive (EC 2000/60/EC)*. London: Department for Environment, Food and Rural Affairs.

De Groot, R., Stuip, M., Finlayson, M. and Davidson, N. (2006). *Valuing wetlands: guidance for valuing the benefits derived from wetland ecosystem services*. Netherlands: International Water Management Institute.

Department for Transport. (2014). Supplementary Guidance: Mixed Logit Models. [online] Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427138/webtag-tag-supplementary-mixed-logit-models.pdf [Accessed 15 August 2017].

Desvousges, W. H., Smith, V. K., and Fisher, A. (1987). Option price estimates for water quality improvements: A contingent valuation study for the Monongahela river. *Journal of Environmental Economics and Management*, 14(3), pp. 248–267.

Diamond, P. A. and Hausman, J. A. (1994). Contingent valuation: Is some number better than no number? *The Journal of Economic Perspectives*, 8(4), pp. 45-64.

Dickson, J. W., Edwards, T., Jeffrey, W. A. and Kay, D. (2005). *Catchment scale appraisal of best management practices (BMPs) for the improvement of bathing water-Brighouse Bay*. Report prepared for Scottish Environment Protection Agency (SEPA), Reference 230/4187. Edinburgh: SAC Environmental, Auchincruive.

Dodds, W. K., Bouska, W. W., Eitzmann, J. L., Pilger, T. J., Pitts, K. L., Riley, A. J., Schloesser, J.T. and Thornbrugh, D. J. (2008). Eutrophication of US freshwaters: analysis of potential economic damages. *Environmental Science and Technology*, 43(1), pp. 12–19.

Doherty, E., Murphy, G., Hynes, S. and Buckley, C. (2014). Valuing ecosystem services across water bodies: results from a discrete choice experiment. *Ecosystem Services*, 7, pp. 89-97.

Donigian A. S. Jr., Bicknell, B. R., Imhoff, J. C. and Singh, V. P. (1995). Hydrological Simulation Program-Fortran (HSPF). In: V. P. Singh, ed., Computer models of watershed hydrology, 2nd ed. Littleton, C.O.: Water Resources Publications, pp. 395-442.

Dorevitch, S., Panthi, S., Huang, Y., Li, H., Michalek, A. M., Pratap, P., Wroblewski, M., Liu, L., Scheff, P. A. and Li, A. (2011). Water ingestion during water recreation. *Water Research*, 45(5), pp. 2020–2028.

Dufour, A. P. (1984). *Health effects criteria for fresh recreational waters. EPA 600/1-84-004.* Cincinnati: US Environmental Protection Agency.

Duncan, C., Jones, K. and Moon, G. (1993). Do places matter? A multi-level analysis of regional variations in health-related behaviour in Britain. *Social Science and Medicine*, 37(6), pp. 725-733.

Dwight, R. H., Baker, D. B., Semenza, J. C. and Olson, B. H. (2004). Health effects associated with recreational water use in urban vs. rural California. *American Journal of Public Health*, 94(4), pp. 565-567.

Dwight, R. H., Fernandez, L. M., Baker, D. B., Semenza, J. C. and Olson, B. H. (2005). Estimating the economic burden from illnesses associated with recreational coastal water pollution--a case study in Orange County, California. *Journal of Environmental Management*, 76(2), pp. 95-103.

DWF LLP. (2016). The implications for UK environmental legislation following Brexit. [online] Available at: https://www.dwf.law/news-events/legal-updates/2016/07/the-implications-for-uk-environmental-legislation-following-brexit/ [Accessed 30 July 2016].

Dwr Cymru. (2005). Welsh Water Monitoring Plan 2005-2010. [online] Available at: http://www.dwrcymru.co.uk/English/library/Reports/companyreports/investmentplans/Price%20Review/2005/MonitoringPlan2005.pdf [Accessed 10 April 2008].

Edberg, S. C. (1998). *The Germ Theory*. New Haven, CT: Medical Microbiology Lectures, Yale University.

Edberg, S. C. L., Rice, E. W., Karlin, R. J. and Allen, M. J. (2000). *Escherichia coli*: the best biological drinking water indicator for public health protection. *Journal of Applied Microbiology*, 88(S1).

EDINA. (2008a). *Agricultural census. Grid square agricultural census data for England, Scotland and Wales* [computer files]. Edinburgh: EDINA, University of Edinburgh.

EDINA. (2008b). Agcensus frequently asked questions. [online] Available at: http://edina.ac.uk/agcensus/support/faq.shtml#q15 [Accessed 18 May 2008].

EFTEC. (2002). *Valuation of Benefits to England and Wales of Revised Bathing Water Quality Directive and other Beach Characteristics Using the Choice Experiment Methodology. Report prepared for Defra.* London: Economics for the Environment Consultancy Ltd.

EFTEC. (2011). *Water Framework Directive: Disproportionate Costs. Report to UK Water Industry Research*. London: Economics for the Environment Consultancy Ltd.

Eijkman, C. (1904). Uber thermolabile stoffwechselprodukte als ursache der natuerlichen wachstumshemmung der mikroorganismen. *Zentralblatt fur Bakteriologie und Parasitenkunde,* Abt. I, 37, pp. 436-449.

Elgot, J. (2016). Banks could move assets out of UK by 2017 if 'EU passport' is lost. [online] The Guardian. Available at: https://www.theguardian.com/business/2016/oct/17/banks-could-move-assets-out-of-uk-by-2017-if-eu-passport-is-lost [Accessed 20 November 2016].

Environment Agency. (2000). *The Microbiology of Recreational and Environmental Waters 2000. Methods for the Enumeration of Waters and Associated Material; Standing Committee of Analysts*. Bristol: The Environment Agency.

Environment Agency. (2004). *A flood risk management strategy for the River Tame. Scoping Report*. Solihull: Environment Agency.

Environment Agency. (2008a). Regulatory Water Quality Decision Document. [online] Available at: http://www.environment-agency.gov.uk/commondata/acrobat/uv_decision_document_2034148.pdf [Accessed 28 May 2008].

Environment Agency. (2008b). Monitoring for the Water Framework Directive. [online] Available at: http://www.environment-agency.gov.uk/subjects/waterquality/955573/1001324/958407/?version=1&lang=_e [Accessed 2 February 2008].

Environment Agency. (2010a). Aire Catchment Flood Management Plan. Summary Report December 2010. [online] Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/289346/River_Aire_Catchment_Flood_Management_Plan.pdf [Accessed 8 November 2017].

Environment Agency. (2010b). *Public Attitudes to Angling 2010*. Bristol: Environment Agency.

Environment Agency. (2014). *The Broadland Rivers Management Catchment*. Bristol: Environment Agency.

Environment Agency. (2015a). Yare Operational Catchment. [online] Available at: http://environment.data.gov.uk/catchment-planning/OperationalCatchment/an-yare/Summary [Accessed 15 June 2015].

Environment Agency. (2015b). Water for life and livelihoods. Part 1: Anglian river basin district. River basin management plan. Bristol: Environment Agency.

Environment Directorate General. (2005). Introduction to the New EU Water Framework Directive. [online] Available at: http://ec.europa.eu/environment/water/water-framework/info/intro_en.htm [Accessed 27 November 2007].

ESRC. (2007). Research Ethics Framework. [online] Available at: http://www.esrcsocietytoday.ac.uk/ESRCInfoCentre/Images/ESRC_Re_Ethics_Frame_tcm6-11291.pdf [Accessed 12 February 2008].

ESRI Inc. (2005). *ArcMap* (version 9.1) [computer software]. Redlands, California: ESRI Inc.

ESRI Inc. (2006). ARCGIS Desktop Help Release 9.1: Georeferencing a raster dataset. [online] Available at: http://webhelp.esri.com/arcgisdesktop/9.1/index.cfm?id=410&pid=398&topicname=Georeferencing%20a%20raster%20dataset [Accessed 9 May 2008].

ESRI Inc. (2008). *ArcMap* (version 9.3) [computer software]. Redlands, California: ESRI Inc.

ESRI Inc. (2012). *ArcMap* (version 10.1) [computer software]. Redlands, California: ESRI Inc.

EU. (2000). Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy. *Official Journal of the European Union*, L327, pp. 1-73.

EU. (2006a). Directive 2006/7/EC of The European Parliament and of The Council of 15th February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC. *Official Journal of the European Union*, L64, pp. 37-51.

EU. (2006b). Directive 2006/113/EC of the European Parliament and of the council of 12 December 2006 on the quality required of shellfish waters. *Official Journal of the European Union*, L376, pp. 14-20.

European Commission. (2012). Report from the Commission to the European Parliament and the Council on the Implementation of the Water Framework

Directive 2000/60/EC River Basin Management Plans (COM(2012) 670 final of 14.11.2012). [online] Available at: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52012DC0670&from=EN [Accessed 8 November 2017].

European Commission CIRCABC (European Commission Communication and Information Resource Centre for Administrations, Businesses and Citizens). (2016a). The 2027 challenge. [online] Available at: https://circabc.europa.eu/webdav/CircaBC/env/wfd/Library/water_directors/xh%20-%20Documents%20Extraordinary%20WDs%20-%206%20October%20(Brussels)/20160822%20DISCUSSIONPAPER.pdf [Accessed 14 November 2017].

European Commission CIRCABC (European Commission Communication and Information Resource Centre for Administrations, Businesses and Citizens). (2016b). Statement as regards the continuation of the discussion on the 2027 challenge. [online] Available at: https://circabc.europa.eu/webdav/CircaBC/env/wfd/Library/water_directors/zi%20-%20Documents%20Slovakian%20WD%20(28-29%20November%2c%20Bratislava)/e%20-%20Presentations/WD2016-2_Item%201_TRIO%20Presidency%20Statement.pdf [Accessed 14 November 2017].

Farming Online. (2016). Gvt delays 25 year plans in light of Brexit vote. [online] Available at: http://www.farming.co.uk/news/article/12694 [Accessed 3 August 2016].

Farrell, J. and Gibbons, R. (1989). Cheap talk can matter in bargaining. *Journal of Economic Theory*, 48(1), pp. 221-237.

Ferley, J. P., Zmirou, D., Balducci, F., Baleux, B., Fera, P., Larbaigt, G., Jacq, E., Moissonnier, B., Blineau, A. and Boudot, J. (1989). Epidemiological Significance of Microbiological Pollution Criteria for River Recreational Waters. *International Journal of Epidemiology*, 18(1), pp. 198-205.

Ferrini, S. and Scarpa, R. (2007). Designs with a-priori information for nonmarket valuation with choice experiments: a Monte Carlo study. *Journal of Environmental Economics and Management*, 53(3), pp. 342-363.

Ferrini, S., Schaafsma, M. and Bateman, I. J. (2014). Revealed and stated preference valuation and transfer: A within-sample comparison of water quality improvement values. *Water Resources Research*, 50(6), pp. 4746–4759.

Fewtrell, L., Jones, F., Kay, D., Wyer, M. D., Godfree, A. F. and Salmon, B. L. (1992). Health effects of white-water canoeing. *The Lancet*, 339(8809), pp. 1587–1589.

Fewtrell, L., Kay, D., Salmon, R. L., Wyer, M. D., Newman, G. and Bowering, G. (1994). The health effects of low-contact water activities in fresh and estuarine waters. *Journal of the Chartered Institution of Water and Environmental Management*, 8(1), pp. 97-101.

Fezzi, C., Rigby, D., Bateman, I. J., Hadley, D. and Posen, P. (2008). Estimating the range of economic impacts on farms of nutrient leaching reduction policies. *Agricultural Economics*, 39(2), pp. 197–205.

Fezzi, C. and Bateman, I. J. (2009). "Structural Agricultural Land Use Modelling", paper presented at the International Association of Agricultural Economists 2009 Conference, August 16-22 2009, Beijing, China. [online] Available at: http://ageconsearch.umn.edu/handle/51423 [Accessed 9 September 2009].

Fezzi, C., Harwood, A. R., Lovett, A. A. and Bateman, I. J. (2015). The environmental impact of climate change adaptation on land use and water quality. *Nature Climate Change*, 5(3), pp. 255-260.

Field, B. C. (2008). *Natural resource economics: An introduction,* 2nd ed. Long Grove, Illinois: Waveland Press.

Filip, Z., Kaddu-Mulindwa, D. and Milde, G. (1988). Survival of some pathogenic and facultative pathogenic bacteria in groundwater. *Water Science and Technology*, 20(3), pp. 227-231.

Fleisher, J. M., Kay, D., Salmon, R. L., Jones, F., Wyer, M. D. and Godfree, A. F. (1996). Marine waters contaminated with domestic sewage: nonenteric illnesses associated with bather exposure in the United Kingdom. *American Journal of Public Health*, 86(9), pp. 1228–1234.

Fleming, C. M. and Cook, A. (2008). The recreational value of Lake McKenzie, Fraser Island: An application of the travel cost method. *Tourism Management*, 29(6), pp. 1197–1205.

Food Standards Agency. (2009). Healthy Diet. [online] Available at: http://www.eatwell.gov.uk/healthydiet/ [Accessed 10 January 2009].

Foster, V. and Mourato S. (2003). Elicitation format and sensitivity to scope: Do contingent valuation and choice experiments give the same results? Environmental And Resource Economics, 24(2), pp. 141-160.

Fraser, R. H., Barten, P. K. and Pinney, D. A. K. (1998). Predicting Stream Pathogen Loading from Livestock using a Geographical Information System-Based Delivery Model. *Journal of Environmental Quality*, 27(4), pp. 935-945.

Freyberg, T. (2016). Brexit: UK water sector reaction. [online] Water and Wastewater International. Available at: http://www.waterworld.com/articles/wwi/2016/06/brexit-uk-water-sector-reaction.html [Accessed 28 June 2016].

Fuller, R. M., Cox, R., Clarke, R. T., Rothery, P., Hill, R. A., Smith, G. M., Thomson, A. G., Brown, N. J., Howard, D. C. and Stott, A. P. (2005). The UK land cover map 2000: Planning, construction and calibration of a remotely sensed, user-oriented map of broad habitats. *International Journal of Applied Earth Observation and Geoinformation*, 7(3), pp. 202-216.

Gale, D. and Mas-Colell, A. (1978). On the role of complete, transitive preferences in equilibrium theory. In: G. Schwödiauer, ed., *Equilibrium and disequilibrium in economic theory: Proceedings of a Conference Organized by the Institute for Advanced Studies, Vienna, Austria July 3–5, 1974*. Netherlands: Springer, pp. 7–14.

Galioto, F., Marconi, V., Raggi, M. and Viaggi, D. (2013). An Assessment of Disproportionate Costs in WFD: The Experience of Emilia-Romagna. *Water*, 5(4), pp. 1967–1995.

GAMS Software GmbH. (2008). *GAMS Base Module*. [computer software]. Koln, Germany: GAMS Software GmbH.

Gassman, P. W., Reyes, M. R., Green, C. H. and Arnold, J. G. (2007). The soil and water assessment tool: historical development, applications, and future research directions. *Transactions of the ASABE*, 50(4), pp. 1211-1250.

Georgiou, S., Bateman, I. J., Cole, M. and Hadley, D. (2000). *Contingent ranking and valuation of river water quality improvements: Testing for scope sensitivity, ordering and distance decay effects. CSERGE Working Paper GEC 2000-18.* Norwich: Centre for Social and Economic Research on the Global Environment.

Georgiou, S., Langford, I. H., Bateman, I. J. and Turner, R. K. (1998). Determinants of individuals' willingness to pay for perceived reductions in environmental health risks: a case study of bathing water quality. *Environment and Planning A*, 30(4), pp. 577–594.

Gerba, C. P., Melnick, J. L. and Wallis, C. (1975). Fate of Wastewater Bacteria and Viruses in Soil. *Journal of the Irrigation and Drainage Division*, 101(3), pp. 157-174.

Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. Berkley and Los Angeles: University of California Press.

Glenk, K., Lago, M. and Moran, D. (2011). Public preferences for water quality improvements: implications for the implementation of the EC Water Framework Directive in Scotland. *Water Policy*, 13(5), pp. 645–662.

Goldstein, H. 1995. *Multilevel Statistical Models.* 2nd ed. London: Edward Arnold.

Görlach, B. and Pielen, B. (2007). "Disproportionate Costs in the EC Water Framework Directive – The Concept and its Practical Implementation", paper presented at the ENVECON 2007 Applied Environmental Economics Conference, 23 March 2007, London.

Grafton, Q., Adamowicz, W., Dupont, D., Nelson, H., Hill, R. J. and Renzetti, S. (2004). *The economics of the environment and natural resources*. Oxford: Blackwell Publishing.

Green, C. and Tunstall, S. (1999). A Psychological Perspective. In: I. J. Bateman and K. G. Willis, eds., *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries,* 1st ed. Oxford: Oxford University Press, pp. 207–257.

Greene, W. H. and Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37(8), pp.681-698.

Gujarati, D. N. (2003). *Basic Econometrics*. 4th ed. New York: McGraw-Hill.

Gustard, A., Bullock, A. and Dixon, J. M. (1992). *Low flow estimation in the United Kingdom. Report No. 108*. Wallingford: Institute of Hydrology.

Gysen, J., Bruyninckx, H. and Bachus, K. (2006). The Modus Narrandi: A Methodology for Evaluating Effects of Environmental Policy. *Evaluation*, 12(1), pp. 95–118.

Haas, C. N., Rose, J. B. and Gerba, C. P. (1999). *Quantitative Microbial Risk Assessment*. Chichester: John Wiley and Sons.

Hagedorn, C., Hansen, D. T. and Simonson, G. H. (1978). Survival and Movement of Fecal Indicator Bacteria in Soil under Conditions of Saturated Flow. *Journal of Environmental Quality*, 7(1), pp. 55-59.

Hampson, D., Crowther, J., Bateman, I. J., Kay, D., Posen, P., Stapleton, C., Wyer, M., Fezzi, C., Jones, P. and Tzanopoulos, J. (2010). Predicting microbial pollution concentrations in UK rivers in response to land use change. *Water Research*, 44(16), pp. 4748–4759.

Hanemann, W. M. (1984). Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics*, 66(3), pp. 332–341.

Hanley, N. Wright, R. E. and Adamowicz, V. (1998). Using choice experiments to value the environment. *Environmental and resource economics*, 11(3), pp.413-428.

Hanley, N., Mourato, S. and Wright, R. E. (2001). Choice modelling approaches: a superior alternative for environmental valuation? *Journal of Economic Surveys*, 15(3), pp. 435–462.

Hanley, N., Bell, D. and Alvarez-Farizo, B. (2003). Valuing the Benefits of Coastal Water Quality Improvements Using Contingent and Real Behaviour. *Environmental and Resource Economics*, 24(3), pp. 273–285.

Hanley, N., Adamowicz, W. and Wright, R. E. (2005). Price vector effects in choice experiments: an empirical test. *Resource And Energy Economics*, 27(3), pp. 227-234.

Hanley, N., Wright, R. E. and Alvarez-Farizo, B. (2006). Estimating the economic value of improvements in river ecology using choice experiments: an application to the water framework directive. *Journal of Environmental Management*, 78(2), pp. 183–193.

Hanley, N., Colombo, S., Tinch, D., Black, A. and Aftab, A. (2006b). Estimating the benefits of water quality improvements under the Water Framework Directive: are benefits transferable?. *European Review of Agricultural Economics*, 33(3), pp. 391-413.

Hanley, N., Barbier, E. B. and Barbier, E. (2009). *Pricing Nature: Cost-benefit Analysis and Environmental Policy*. Cheltenham: Edward Elgar Publishing.

Happs, J. (1986). Constructing an understanding of water-quality: Public perceptions and attitudes concerning three different water-bodies. *Research in Science Education*, 16(1), pp. 208–215.

Harley, A. (2008). Bathing Waters Directive. [online] Environment Agency. Available at: http://www.environment-agency.gov.uk/business/1745440/444663/1769279/?version=1&lang=_e [Accessed 14 April 2008].

Harrabin, R. (2016) Countryside faces Brexit anxiety. [online] BBC News. Available at: http://www.bbc.co.uk/news/science-environment-36809570 [Accessed 15 July 2016].

Harwood, V. J., Staley, C., Badgley, B. D., Borges, K. and Korajkic, A. (2014). Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS microbiology reviews*, 38(1), pp.1-40.

Hausman, J. A. (2012). Contingent valuation: from dubious to hopeless. *The Journal of Economic Perspectives*, 26(4), pp. 43-56.

Hausman, J. A., Leonard, G. K. and McFadden, D. (1995). A utility-consistent, combined discrete choice and count data model Assessing recreational use losses due to natural resource damage. *Journal of Public Economics*, 56(1), pp. 1–30.

Havelaar, A. H., Pot-Hogeboom, W. M., Furuse, K., Pot, R. and Hormann, M. P. (1990). F-specific RNA bacteriophages and sensitive host strains in faeces and wastewater of human and animal origin. *Journal of Applied Microbiology*, 69(1), pp.30-37.

Haygarth, P., Granger, S., Chadwick, D., Shepherd, M. and Fogg, P. (2005). *A Provisional Inventory of Diffuse Pollution Losses*. *ADAS Report for Defra.* Wolverhampton: ADAS.

Heberlein, T. A., Wilson, M.A., Bishop, R. C. and Schaeffer, N. C. (2005). Rethinking the scope test as a criterion for validity in contingent valuation. *Journal of Environmental Economics and Management*, 50(1), pp. 1-22.

Heinz, I. (2003). Germany: Searching for Strategies to Provide Pure Water. In: F. Brouwer, I. Heinz and T. Zabel, eds., *Governance of Water Related Conflicts in Agriculture*, 1st ed. Dordrecht: Kluwer Academic Publishers, pp. 85-107.

Helming, J. and Reinhard, S. (2009). Modelling the economic consequences of the EU Water Framework Directive for Dutch agriculture. *Journal of Environmental Management,* 91(1), pp. 114-123.

Hensher, D. A. and Greene, W. H. (2003). The mixed logit model: the state of practice. *Transportation*, 30(2), pp. 133–176.

Hensher, D. A., Rose, J. M., Greene, W. H. (2005). *Applied Choice Analysis: A Primer.* Cambridge: Cambridge University Press.

Hensher, D., Shore, N. and Train, K. (2005). Households' willingness to pay for water service attributes. *Environmental and Resource Economics*, 32(4), pp. 509–531.

Heron, M. (2014). [Interview] Danyel Hampson with Max Heron, Chairman of the Whitlingham Boathouses Foundation, 4 February 2014.

Herriges, J., Kling, C., Liu, C.C. and Tobias, J. (2010). What are the consequences of consequentiality?. *Journal of Environmental Economics and Management*, 59(1), pp. 67-81.

Hertzog, M. A. (2008). Considerations in determining sample size for pilot studies. *Research in Nursing and Health*, 31(2), pp.180-191.

Hess, S. and Daly, A. (2014). *Handbook of choice modelling*. Cheltenham: Edward Elgar Publishing.

Hess, S., Hensher, D. A. and Daly, A. (2012). Not bored yet–revisiting respondent fatigue in stated choice experiments. *Transportation research part A: policy and practice*, 46(3), pp.626-644.

Hess, S., Rose, J. M. and Polak, J. (2010). Non-trading, lexicographic and inconsistent behaviour in stated choice data. *Transportation Research Part D: Transport and Environment*, 15(7), pp. 405-417.

Hibbard, J. H., Slovic, P., Peters, E. and Finucane, M. L. (2002). Strategies for reporting health plan performance information to consumers: evidence from controlled studies. *Health Services Research*, 37(2), pp. 291–313.

Hime, S., Bateman, I. J., Posen, P. and Hutchins, M. (2009). *A transferable water quality ladder for conveying use and ecological information within public* surveys*. CSERGE working paper EDM 09-01.* Norwich: Centre for Social and Economic Research on the Global Environment.

HMSO. (1994). *The Microbiology of Water. Part I: Drinking Water. Reports in Public Health and Medical Subjects, No. 71; Methods for the Examination of Waters and Associated Materials*. London: Her Majesty's Stationery Office.

H. M. Treasury. (2003). The Green Book: Appraisal and Evaluation in Central Government. [online] Available at: http://greenbook.treasury.gov.uk/index.htm [Accessed 28 November 2007].

H.M. Treasury. (2016). Chancellor Philip Hammond guarantees EU funding beyond date UK leaves the EU. [online] Available at: https://www.gov.uk/government/news/chancellor-philip-hammond-guarantees-eu-funding-beyond-date-uk-leaves-the-eu [Accessed 28 August 2016].

House of Commons Environmental Audit Committee. (2016). EU and UK Environmental Policy. Third Report of Session 2015–16, 19 April 2016. [online] Available at: http://www.publications.parliament.uk/pa/cm201516/cmselect/cmenvaud/537/537.pdf [Accessed 20 June 2016].

Houses of Parliament Parliamentary Office of Science and Technology. (2014). Diffuse Pollution of Water by Agriculture. Postnote Number 478 October 2014. [online] Available at: researchbriefings.files.parliament.uk/documents/POST-PN-478/POST-PN-478.pdf [Accessed 8 November 2017].

Hotelling, H. (1949). Letter to the Director of the National Park Service. In: R. A. Prewitt, ed., *The Economics of Public Recreation. The Prewitt Report*. Washington, D.C.: Department of the Interior.

Hoyos, D., Mariel, P. and Fernández-Macho, J. (2009). The influence of cultural identity on the WTP to protect natural resources: some empirical evidence. *Ecological Economics*, 68(8), pp. 2372–2381.

Hunter, P. D., Hanley, N., Czajkowski, M., Mearns, K., Tyler, A. N., Carvalho, L. and Codd, G. A. (2012). The effect of risk perception on public preferences and willingness to pay for reductions in the health risks posed by toxic cyanobacterial blooms. *Science of The Total Environment*, 426, pp. 32–44.

Hutchins, M. (2008a). queries. [email 18 June 2008].

Hutchins, M. (2008b). Re: Tame shapefile/ Sampling Network Data. [email 5 June 2008].

Hutchins, M., Fezzi, C., Bateman, I. J., Posen, P. and Deflandre-Vlandas, A. (2009). Cost-effective mitigation of diffuse pollution: Setting criteria for river basin management at multiple locations. *Environmental Management*, 44(2), pp. 256-267.

Hynes, S., Hanley, N. and Scarpa, R. (2008). Effects on Welfare Measures of Alternative Means of Accounting for Preference Heterogeneity in Recreational Demand Models. *American Journal of Agricultural Economics*, 90(4), pp. 1011–1027.

IGES. (2008). Centre for Research into Environment and Health. [online] University of Aberystwyth. Available at: http://www.ies.aber.ac.uk/en/research/groups/centre-research-environment-and-health [Accessed 5 June 2008].

Institute for European Environmental Policy. (2016). The potential policy and environmental consequences for the UK of a departure from the European Union. [online] Available at: http://www.ieep.eu/assets/2000/IEEP_Brexit_2016.pdf [Accessed 28 July 2016].

ISO. (1998). *Water quality - Detection and enumeration of intestinal enterococci in surface and waste water - Part 1: Miniaturized method (most probable number) by inoculation in liquid medium*. Geneva: International Organization for Standardization (ISO/FDIS 7899-1).

Jenkins, A., Kirkby, M., McDonald, A., Naden, P. and Kay, D. (1984). A process based model of faecal bacterial levels in upland catchments. *Water Science and Technology*, 16, pp. 453-462.

Johnson, F. R., Lancsar, E., Marshall, D., Kilambi, V., Mühlbacher, A., Regier, D. A., Bresnahan, B. W., Kanninen, B. and Bridges, J.F. (2013). Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. *Value in Health*, 16(1), pp.3-13.

Johnston R. J., Boyle, K. J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T. A., Hanemann, W. M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., and Vossler, C. A. (2017). Contemporary Guidance for Stated Preference Studies. *Journal of the Association of Environmental and Resource Economists*, 4, pp.319-405.

Jones, D. L. (1999). Potential health risks associated with the persistence of *Escherichia coli O157* in agricultural environments. *Soil Use and Management*, 15(2), pp. 76-83.

Jones, F. and White, W. R. (1984). Health and amenity aspects of surface waters. *Water Pollution Control*, 83(2), pp. 215–225.

Jones, P. J. and Tranter, R. B. (2008). *Implications of a Nutrition Driven Food Policy for Land Use and the Rural Environment: Modelling the Impact of Different Policy Scenarios on farm Business Management, Land Use and Rural Employment. Work Package 5.2. Interim Report to the Rural Economy and Land Use Programme. August 2008*. Reading: University of Reading.

Jones, P. J., Tzanopoulos, J., Mortimer, S. R. and Traill, B. (2009). “The future for food producing landscapes”, paper presented at the Rural Economy and Land Use Programme Conference, 4 June 2009, London.

JNCC (Joint Nature Conservation Committee). (2010). Council Directive 2000/60/EC establishing a framework for Community action in the field of water policy (Water Framework Directive). [online] Available at: http://jncc.defra.gov.uk/page-1375 [Accessed 8 November 2017].

JNCC (Joint Nature Conservation Committee). (2016). B7. Surface water status. [online] Available at: http://jncc.defra.gov.uk/page-4250 [Accessed 8 November 2017].

Kay, D. (1979). *Factors Affecting Water Quality within Two Pennine Reservoirs under a System of Multiple-Use Catchment Management; PhD Thesis*. Leeds: University of Leeds.

Kay, D. and McDonald, A. (1980). Reduction of coliform bacteria in two upland reservoirs: the significance of distance decay relationships. *Water Research*, 14(4), pp. 305–318.

Kay, D., Jones, F., Wyer, M. D., Fleisher, J. M., Salmon, R. L., Godfree, A. F., Zelenauch-Jacquotte, A. and Shore, R. (1994). Predicting likelihood of gastroenteritis from sea bathing: results from randomised exposure. *The Lancet,* 344(8927), pp. 905-909.

Kay, D., Wyer, M. D., Crowther, J. and Fewtrell, L. (1999). Faecal indicator impacts on recreational waters: budget studies and diffuse source modelling. *Journal of Applied Microbiology, Symposium Supplement*, 85, pp. 70S–82S.

Kay, D., Bartram, J., Prüss, A., Ashbolt, N., Wyer, M. D., Fleisher, J. M., Fewtrell, L., Rogers, A. and Rees, G. (2004a). Derivation of numerical values for the World Health Organization guidelines for recreational waters. *Water Research*, 38(5), pp. 1296-1304.

Kay, D., Wyer, M., Stapleton, C. and Crowther, J. (2004b). "Faecal indicator bacteria from catchment sources: relevance to the Water Framework Directive", paper presented at the Mersey Basin Campaign Integrated River Basin Management in the Northwest: Problems and Solutions Conference, 6 June 2004, Manchester.

Kay, D., Wyer, M., Crowther, J., Stapleton, C., Bradford, M., McDonald, A., Greaves, J., Francis, C. and Watkins, J. (2005a). Predicting faecal indicator fluxes using digital land use data in the UK's sentinel Water Framework Directive catchment: The Ribble study. *Water Research*, 39(16), pp. 3967–3981.

Kay, D., Wyer, M. D., Crowther, J., Wilkinson, J., Stapleton, C. and Glass, P. (2005b). Sustainable reduction in the flux of microbial compliance parameters from urban and arable land use to coastal bathing waters by a wetland ecosystem produced by a marine flood defence structure. *Water Research*, 39(14), pp. 3320-3332.

Kay, D., McDonald, A., Stapleton, C., Wyer, M. and Fewtrell, L. (2006a). Europe: a challenging new framework for water quality. *Proceedings of the Institution of Civil Engineers - Civil Engineering, Water Management.* 159(5), pp. 58–64.

Kay, D., Stapleton, C., Wyer, M., McDonald, A. and Crowther, J. (2006b). Total Maximum Daily Loads: the USEPA Approach to Managing Faecal Indicator Fluxes to Receiving Waters: Lessons for UK Environmental Regulation? In: L. Gairns, C. Crighton and W. A. Jeffrey, eds., *Agriculture and the Environment VI: Managing Rural Diffuse Pollution. Proceedings of the SAC and SEPA Biennial Conference, Edinburgh, 5-6 April 2006*. Edinburgh: International Water Association, Scottish Agricultural College and Scottish Environmental Protection Agency, pp. 23-33.

Kay, D., Edwards, A. C., Ferrier, R. C., Francis, C., Kay, C., Rushby, L., Watkins, J., McDonald, A. T., Wyer, M., Crowther, J. and Wilkinson, J. (2007a). Catchment microbial dynamics: the emergence of a research agenda. *Progress in Physical Geography*, 31(1), pp. 59-76.

Kay, D., Aitken, M., Crowther, J., Dickson, I., Edwards, A. C., Francis, C., Hopkins, M., Jeffrey, W., Kay, C., McDonald, A. T., McDonald, D., Stapleton, C. M., Watkins, J., Wilkinson, J. and Wyer, M. D. (2007b). Reducing fluxes of faecal indicator compliance parameters to bathing waters from diffuse agricultural sources: The Brighouse Bay study, Scotland. *Environmental Pollution*, 147(1), pp. 138-149.

Kay, D., Crowther, J., Stapleton, C. M., Wyer, M. D., Fewtrell, L., Edwards, A., Francis, C. A., McDonald, A. T., Watkins, J. and Wilkinson, J. (2008a). Faecal indicator organism concentrations in sewage and treated effluents. *Water Research*, 42(1), pp. 442-454.

Kay, D., Crowther, J., Stapleton, C. M., Wyer, M. D., Fewtrell, L., Anthony, S., Bradford, M., Edwards, A., Francis, C. A., Hopkins, M., Kay, C., McDonald, A. T., Watkins, J. and Wilkinson, J. (2008b). Faecal indicator organism concentrations and catchment export coefficients in the UK. *Water Research*, 42(10), pp. 2649–2661.

Kay, D. (2008c). Re: FIO reduction caused by Lea Marston purification lakes. [email 24 June 2008].

Kay, D., Crowther, J., Kay, C., McDonald, A. T., Ferguson, C., Stapleton, C. M., and Wyer, M. D. (2012). Effectiveness of best management practices (BMPs) for attenuating the transport of livestock-derived pathogens within catchments. In: A. Dufour, J. Bartram, R. Bos and V. Gannon, eds., *Animal Waste, Water Quality and Human Health - Emerging Issues in Water and Infectious Disease Series*, 1st ed. London: International Water Association Publishing, pp. 95-256.

Kemperman, A. D. and Timmermans, H. J. (2006). Heterogeneity in urban park use of aging visitors: a latent class analysis. *Leisure Sciences*, 28(1), pp. 57–71.

Kessels, R., Goos, P. and Vandebroek, M. (2006). A comparison of criteria to design efficient choice experiments. *Journal of Marketing Research*, 43(3), pp.409-419.

Konishi, Y. and Coggins, J. S. (2008). Environmental risk and welfare valuation under imperfect information. *Resource and Energy Economics*, 30(2), pp. 150–169.

Krinsky, I. and Robb, A. L. (1986). On Approximating the Statistical Properties of Elasticities. *The Review of Economics and Statistics*, 68(4), pp. 715–719.

Krosnick J. A., Holbrook, A. L., Berent, M. K., Carson, R. T., Hanemann, W. M., Kopp, R. J., Mitchell, R. C., Presser, S., Ruud, P., Smith, V. K., Moody, W. R., Green, M. C. and Conaway, M. (2002). The impact of “no opinion” response options on data quality: non-attitude reduction or an invitation to satisfice. *Public Opinion Quarterly*, 66(3), pp. 371-403.

Lancaster, G. A., Dodd, S. and Williamson, P. R. (2004). Design and analysis of pilot studies: recommendations for good practice. *Journal of Evaluation in Clinical Practice*, 10(2), pp.307-312.

Lancaster, K. J. (1966). A new approach to consumer theory. *The Journal of Political Economy*, 74(2), pp. 132–157.

Lancsar, E. and Louviere, J. (2006). Deleting ‘irrational’ responses from discrete choice experiments: a case of investigating or imposing preferences?. *Health Economics*, 15(8), pp. 797-811.

Landscape Institute. (2016). The EU Referendum. [online] Available at: http://www.landscapeinstitute.org/policy/documents/EUReferendumbriefing.pdf [Accessed 20 June 2016].

Langford, I. H., Georgiou, S., Bateman, I. J., Day, R. J. and Turner, R. K. (2000). Public perceptions of health risks from polluted coastal bathing waters: a mixed methodological analysis using cultural theory. *Risk Analysis*, 20(5), pp. 691–704.

Larsen, H. and Ipsen, N. H. (1997). Framework for Water Pollution Control. In: R. Helmer and I. Hespanhol, eds., *Water Pollution Control - A Guide to the Use of Water Quality Management Principles.* London: E. and F. Spon, on behalf of the United Nations Environment Programme, the Water Supply and Sanitation Collaborative Council and the World Health Organization, pp. 25.

Larsen, R. E., Miner, J. R., Buckhouse, J. C. and Moore, J. A. (1994). Water-quality benefits of having cattle manure deposited away from streams. *Bioresource Technology*, 48(2), pp. 113-118.

Lawler, D. M., Petts, G. E., Foster, I. D. L. and Harper, S. (2006). Turbidity dynamics during spring storm events in an urban headwater river system: The Upper Tame, West Midlands, UK. *Science of The Total Environment*, 360(1), pp. 109-126.

Lawlor, J., McCarthy, C. and Scott, S. (2007). Investment in water infrastructure: Findings from an economic analysis of a national programme. *Journal of Environmental Planning and Management*, 50(1), pp. 41-63.

Layton, D. F. (1996). “Rank Ordered, Random Coefficients Multinomial Probit Models for Stated Preference Surveys: An Investigation of Preference Heterogeneity, IIA, and Contingent Ranking Data”, paper presented to the Association of Environmental and Resource Economists Summer Workshop, 2-4 June 1996, Tahoe City, USA.

Leeming, R., Ball, A., Ashbolt, N. and Nichols, P. (1996). Using faecal sterols from humans and animals to distinguish faecal pollution in receiving waters. *Water Research*, 30(12), pp. 2893-2900.

Leroux, B. G. (1992). Consistent estimation of a mixing distribution. *The Annals of Statistics*, 20(3), pp. 1350–1360.

Lines, R. (2014). [Interview] Danyel Hampson with Rob Lines, Chair of Trianglia Recreation Club, 22 September 2014.

Lobley, M. and Potter, C., 1998. Environmental stewardship in UK agriculture: a comparison of the environmentally sensitive area programme and the countryside stewardship scheme in South East England. *Geoforum*, 29(4), pp.413-432.

Loomis, J. B. (2014). Strategies for overcoming hypothetical bias in stated preference surveys. *Journal of Agricultural and Resource Economics*, 39(1), pp. 34–46.

Louviere, J. J. (1996). Relating stated preference measures and models to choices in real markets: calibration of CV responses. In: D. J. Bjornstad and J. R. Kahn, eds., *The Contingent Valuation of Environmental Resources: Methodological Issues and Research Needs,* 1st ed. Cheltenham: Edward Elgar Publishing Ltd., pp. 167–188.

Louviere, J. J., Hensher, D. A. and Swait, J. D. (2000). *Stated choice methods: analysis and applications*. Cambridge: Cambridge University Press.

Lowe, P. and Carter, N. (1994). Environmental Politics and Administrative Reform. *The Political Quarterly*, 65(3), pp. 263-274.

Luce, R. D. (1959). *Individual Choice Behavior a Theoretical Analysis*. Chichester: John Wiley and Sons.

Lusk, J. L. (2003). Effects of cheap talk on consumer willingness-to-pay for golden rice. *American Journal of Agricultural Economics*, 85(4), pp. 840-856.

MAFF. (1991). *Code of Good Agricultural Practice for the Protection of Water*. London: Welsh Office Agriculture Department. MAFF Publications.

MAFF. (2000). *Fertiliser Recommendations for Agricultural and Horticultural Crops Manual RB209*, 7th ed. London: HMSO.

Malpezzi, S. (2003). Hedonic pricing models: a selective and applied review. In: T. O'Sullivan and K. Gibb, eds., *Housing Economics and Public Policy*, 1st ed. Oxford: Blackwell Science Ltd., pp. 67-89.

Maringanti, C., Chaubey, I. and Popp, J. (2009). Development of a multiobjective optimization tool for the selection and placement of best management practices for nonpoint source pollution control. *Water Resources Research*, 45(6), pp.1-15.

Martin, J. R. and Brewin, D. J. (1994). Quality improvements and objectives in the Trent catchment. In: B. H. Rolf, ed., *Wastewater treatment: Evaluation and Implementation,* 1st ed. London: Thomas Telford Ltd.

Martin-Ortega, J. and Berbel, J. (2010). Using multi-criteria analysis to explore non-market monetary values of water quality changes in the context of the Water Framework Directive. *Science of the Total Environment*, 408(19), pp. 3990–3997.

Mattikalli, N. M. and Richards, K. S. (1996). Estimation of Surface Water Quality Changes in Response to Land Use Change: Application of The Export Coefficient Model Using Remote Sensing and Geographical Information System. *Journal of Environmental Management*, 48(3), pp. 263-282.

Mazzotta, M. J. and Opaluch, J. J. (1995). Decision making when choices are complex: a test of Heiner's hypothesis. *Land Economics*, 71(4), pp. 500–515.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In: P. Zarembka, ed., *Frontiers in Econometrics,* 1st ed. New York: Academic Press, pp.105-142.

McFadden, D. (1986). The choice theory approach to market research. *Marketing Science*, 5(4), pp. 275–297.

McFadden, D. and Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15(5), pp. 447–470.

McLachlan, G. and Peel, D. (2004). *Finite mixture models*. Chichester: John Wiley and Sons Ltd.

Meals, D. W. (1996). Watershed-scale response to agricultural diffuse pollution control programs in Vermont, USA*. Water Science and Technology,* 33(4-5), pp. 197-204.

Metcalfe, P. J., Baker, W., Andrews, K., Atkinson, G., Bateman, I. J., Butler, S., Carson, R.T., East, J., Gueron, Y., Sheldon, R. and Train, K. (2012). An assessment of the nonmarket benefits of the Water Framework Directive for households in England and Wales. *Water Resources Research*, 48(3).

Met Office. (2010). *MIDAS: UK Soil Temperature Data* [computer files]. Oxford: NCAS British Atmospheric Data Centre.

Microsoft Corporation. (2006). *Microsoft Office Excel 2007* [computer software]. Redmond: Microsoft Corporation.

Miller, V. (2016). Briefing Paper Number 07213. Brexit: impact across policy areas. [online] UK Parliament. Available at: http://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7213 [Accessed 3 September 2016].

Ministry of Foreign Affairs of Japan. (2016). Japan's Message to the United Kingdom and the European Union. [online] Available at: http://www.mofa.go.jp/files/000185466.pdf [Accessed 4 September 2016].

Mitchell, R. C. and Carson, R. T. (1981). *An experiment in determining willingness to pay for national water quality improvements. Report to the US Environmental Protection Agency.* Washington, DC.: Resources for the Future.

Mitchell, J. (2016). Brexit: the implications for environmental law. [online] Blake Morgan LLP. Available at: https://www.blakemorgan.co.uk/training-knowledge/features-and-articles/brexit-implications-environmental-law/ [Accessed 23 July 2016].

Molinos-Senante, M., Hernández-Sancho, F. and Sala-Garrido, R. (2011). Assessing disproportionate costs to achieve good ecological status of water bodies in a Mediterranean river basin. *Journal of Environmental Monitoring*, 13(8), pp. 2091–2101.

Monaghan, R. M., de Klein, C. A. M. and Muirhead, R. W. (2008). Prioritisation of farm scale remediation efforts for reducing losses of nutrients and faecal indicator organisms to waterways: A case study of New Zealand dairy farming. *Journal of Environmental Management*, 87(4), pp. 609-622.

Moore, D. (2016). Brexit Could Effect Green Infrastructure, Sustainability Professionals Fear. [online] Chartered Institution of Wastes Management. Available at: http://www.ciwm-journal.co.uk/brexit-effect-green-infrastructure-sustainability-professionals-fear/ [Accessed 30 June 2016].

Moran, D. (1999). Benefits transfer and low flow alleviation: what lessons for environmental valuation in the UK? *Journal of Environmental Planning and Management*, 42(3), pp. 425–436.

Morey, E., Thacher, J. and Breffle, W. (2006). Using Angler Characteristics and Attitudinal Data to Identify Environmental Preference Classes: A Latent-Class Model. *Environmental and Resource Economics*, 34(1), pp. 91–115.

Morey, E. R. and Thacher, J. A. (2012). *Using choice experiments and latent-class modeling to investigate and estimate how academic economists value and trade off the attributes of academic positions.* Boulder: University of Colorado.

Mourato, S., Georgiou, S., Ozdemiroglu, E., Newcombe, J. and Howarth, A. (2003). *Bathing water revisions: What are the benefits to England and Wales? A stated preference study. CSERGE working paper ECM 03-12*. Norwich: University of East Anglia.

Murphy, J. J., Stevens, T. and Weatherhead, D. (2005). Is cheap talk effective at eliminating hypothetical bias in a provision point mechanism?. *Environmental and Resource economics*, 30(3), pp. 327-343.

Nakagawa, T. and Osaki, S. (1975). The Discrete Weibull Distribution. *IEEE Transactions on Reliability R,* 24(5), pp. 300–301.

National Technical Advisory Committee. (1968). *Water quality criteria*. Washington, DC.: Federal Water Pollution Control Administration, Dept. of the Interior.

Natural England. (2009a). Environmentally sensitive areas scheme (ESA). [online] Available at: http://www.naturalengland.org.uk/ourwork/farming/funding/closedschemes/esa/default.aspx [Accessed 13 January 2010].

Natural England. (2009b). River Yare Catchment Summary - CSF024. [online] Available at: http://publications.naturalengland.org.uk/publication/9339203 [Accessed 13 January 2010].

Natural England. (2016). *Catchment Sensitive Farming: Anglian River Basin District Strategy 2016 to 2021*. York: Natural England.

Nau, B. (2005). What to look for in regression model output. [online] Duke University. Available at: http://www.duke.edu/~rnau/411regou.htm#validation [Accessed 25 July 2008].

NERC. (2009). *UK National water archive Hydrological Digital Terrain Model (HDTM) of the UK (1 km grid)* [electronic files]. Wallingford: Centre for Ecology and Hydrology Environmental Information Data Centre.

NERC. (2008a). *2000 Land cover map* [electronic files]. Huntingdon: Natural Environment Research Council Centre for Ecology and Hydrology Environmental Information Data Centre.

NERC. (2008b). *1990 Land cover map* [electronic files]. Huntingdon: Natural Environment Research Council Centre for Ecology and Hydrology Environmental Information Data Centre.

Neslen, A. (2016). Brexit would free UK from 'spirit-crushing' green directives, says minister. [online] The Guardian. Available at: http://www.theguardian.com/politics/2016/may/30/brexit-spirit-crushing-green-directives-minister-george-eustice [Accessed 30 May 2016].

Nichols, D. S., Prettyman, D. and Gross, M. (1983). Movement of bacteria and nutrients from pit latrines in the boundary waters canoe area wilderness. *Water, Air and Soil Pollution*, 20(2), pp. 171-180.

Norfolk County Council. (2015a). Annual survey of hours and earnings - resident analysis. [online] Available at: http://www.norfolkinsight.org.uk/dataviews/tabular?viewId=315&geoId=5&subsetId= [Accessed 14 August 2015].

Norfolk County Council. (2015b). 2011 Census - Age structure. [online] Available at: http://www.norfolkinsight.org.uk/dataviews/tabular?viewId=229&geoId=54&subsetId=12 [Accessed 14 August 2015].

Office for National Statistics. (2001a). *2001 Census: Digitised Boundary Data (England, Scotland and Wales)* [computer files]. Edinburgh: ESRC Census Programme, Census Geography Data Unit (UKBORDERS), EDINA, University of Edinburgh.

Office for National Statistics. (2001b). *2001 Census: Standard Area Statistics (England, Scotland and Wales)* [computer files]. Manchester: ESRC/JISC Census Programme, Census Dissemination Unit, Mimas, University of Manchester.

Office for National Statistics. (2002). The definition of urban areas. [online] Available at: http://www.statistics.gov.uk/census2001/pdfs/urban_area_defn.pdf [Accessed 18 May 2008].

Office for National Statistics. (2005). Output Area. [online] Available at: http://www.statistics.gov.uk/census2001/glossary.asp#oa [Accessed 20 May 2008].

Office for National Statistics. (2016). Mid-1971 to Mid-2015 Population Estimates for Regions in England and Wales. [online] Available at: https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland/mid2015/ukandregionalpopulationestimates18382015.zip [Accessed 14 August 2015].

Ofwat. (2008). Protecting consumers, promoting value and safeguarding the future. [online] Available at:

http://www.ofwat.gov.uk/aptrix/ofwat/publish.nsf/Content/protecting_interests280905 [Accessed 8 March 2008].

Ogden, I. D., Hepburn, N. F., MacRae, M., Strachan, N. J. C., Fenlon, D. R., Rusbridge, S. M. and Pennington, T. H. (2002). Long-term survival of *Escherichia coli* O157 on pasture following an outbreak associated with sheep at a scout camp. *Letters in Applied Microbiology*, 34(2), pp. 100-104.

Oliver, D. M., Heathwaite, A. L., Hodgson, C. J. and Chadwick, D. R. (2007). Mitigation and current management attempts to limit pathogen survival and movement within farmed grassland. *Advances in Agronomy*, 93(1), pp. 95-152.

Oliver, D. M., Heathwaite, A. L., Fish, R. D., Chadwick, D. R., Hodgson, C. J., Winter, M. and Butler, A. J. (2009). Scale appropriate modelling of diffuse microbial pollution from agriculture. *Progress in Physical Geography*, 33(3), pp. 358-377.

Ordnance Survey. (2008). *OS Meridian2* [computer files]. Edinburgh: EDINA Digimap collections, University of Edinburgh.

Parsons, G. R. (2003). The Travel Cost Model. In: P. A. Champ, K. J. Boyle and T. C. Brown, eds., *A Primer on Nonmarket Valuation*, 1st ed. Netherlands: Springer, pp. 269–329.

Pearce, D. (1998). Environmental appraisal and environmental policy in the European Union. *Environmental and Resource Economics*, 11(3-4), pp. 489-501.

Pearce, D. (1998). Cost benefit analysis and environmental policy. *Oxford Review of Economic Policy*, 14(4), pp. 84–100.

Pearce, D. and Turner, R. K. (1990). *Economics of natural resources and the environment.* London: Harvester Wheatsheaf.

Pearce, D., Atkinson, G. and Mourato, S. (2006). *Cost-Benefit Analysis and the Environment: recent developments*. Paris: OECD Publishing.

Perkins, J. and Hunter, C. (2000). Removal of enteric bacteria in a surface flow constructed wetland in Yorkshire, England. *Water Research*, 34(6), pp. 1941-1947.

Perman, R., Ma, Y., McGilvray, J. and Common, M. (2003). *Natural resource and environmental economics.* 3rd ed. Harlow: Pearson Education.

Philipp, R., Evan, E. J., Hughes, A. O., Grisdale, S. K., Enticott, R. G. and Jephcott, A. H. (1985). Health risks of snorkel swimming in untreated water. *International Journal of Epidemiology*, 14(4), pp. 624-627.

Pilotto, L. S., Douglas, R. M., Burch, M. D., Cameron, S., Beers, M., Rouch, G. J., Robinson, P., Kirk, M., Cowie, C. T., Hardiman, S. and Moore, C. (1997). Health effects of exposure to cyanobacteria (blue–green algae) during recreational water–related activities. *Australian and New Zealand Journal of Public Health*, 21(6), pp. 562–566.

Pond, K. (2005). *Water Recreation and Disease*. London: IWA Publishing, on behalf of the World Health Organization.

Poor, P. J., Boyle, K. J., Taylor, L. O. and Bouchard, R. (2001). Objective versus subjective measures of water clarity in hedonic property value models. *Land Economics*, 77(4), pp. 482–493.

Posen, P., Hutchins, M., Lovett, A. and Davies, H. (2008). Robust interpolation of agricultural census data to hydrological units and implications for diffuse pollution modelling. In: J. Zhang and M. F. Goodchild, eds., *Proceedings of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, vol. 1: Spatial Uncertainty.* Edgbaston: World Academic Union, pp. 46-53.

Posen, P., Hutchins, M., Lovett, A. and Davies, H. (2009). *Robust interpolation of agricultural census data to hydrological units and implications for diffuse pollution modelling: CSERGE Working Paper EDM 09-08.* Norwich: Centre for Social and Economic Research on the Global Environment, University of East Anglia.

Posen, P., Hutchins, M., Lovett, A. and Davies, H. (2011). Identifying the catchment size at which robust estimations of agricultural land use can be made, and implications for diffuse pollution modelling. *Applied Geography*, 31(3), pp. 919–929.

Postle, M., Fenn, T., Foottit, A. and Salado, R. (2004). *CEA and developing a methodology for assessing disproportionate costs. Final Report for Department for Environment, Food and Rural Affairs (Defra), Welsh Assembly Government (WAG), Scottish Executive (SE) and Department of the Environment in Northern Ireland (DOENI).* London: Risk and Policy Analysts Limited (RPA).

Powe, N. A. and Bateman, I. J. (2004). Investigating insensitivity to scope: a split-sample test of perceived scheme realism. *Land Economics*, 80(2), pp. 258-271.

Prescott, S. C. and Winslow, C. E. (1915). *Elements of Water Bacteriology*, 3rd ed. New York: John Wiley.

Pretty, J. N., Mason, C. F., Nedwell, D. B., Hine, R. E., Leaf, S. and Dils, R. (2003). Environmental costs of freshwater eutrophication in England and Wales. *Environmental Science and Technology*, 37(2), pp. 201–208.

Provencher, B. and Bishop, R. C. (2004). Does accounting for preference heterogeneity improve the forecasting of a random utility model? A case study. *Journal of Environmental Economics and Management*, 48(1), pp. 793–810.

Provencher, B. and Moore, R. (2006). A discussion of ‘using angler characteristics and attitudinal data to identify environmental preference classes: a latent-class model’. *Environmental and Resource Economics*, 34(1), pp. 117–124.

Prüss, A. (1998). Review of epidemiological studies on health effects from exposure to recreational water. *International Journal of Epidemiology*, 27(1), pp. 1-9.

Raphael, K. (1987). Recall bias: a proposal for assessment and control. *International journal of epidemiology*, 16(2), pp.167-170.

Ready, R., Navrud, S., Day, B., Dubourg, R., Machado, F., Mourato, S., Spanninks, F. and Rodriquez, M. X. V. (2004). Benefit Transfer in Europe: How Reliable Are Transfers between Countries? *Environmental and Resource Economics*, 29(1), pp. 67-82.

RELU. (2008). *Safe recycling of livestock manures. Policy and Practice Note No. 4, November 2008.* Newcastle-upon-Tyne: Rural Economy and Land Use Programme.

RELU. (2009). Implications of a Nutrition Driven Food Policy for the Countryside. Policy and Practice Notes No. 6. April 2009. [online] Available at: http://www.relu.ac.uk/news/policy%20and%20practice%20notes/Traill/PPN6%20Traill.pdf [Accessed 9 September 2009].

RERC. (2007). *Reviewing Rural Futures: Population Projections and Migration Across the Urban Rural System*. London: Rural Evidence Research Centre.

Revelt, D. and Train, K. (1998). Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of economics and statistics*, 80(4), pp.647-657.

Ridker, R. G. and Henning, J. A. (1967). The determinants of residential property values with special reference to air pollution. *The Review of Economics and Statistics*, 49(2), pp. 246–257.

Rigby, D., Alcon, F. and Burton, M. (2010). Supply uncertainty and the economic value of irrigation water. *European Review of Agricultural Economics*, 37(1), pp. 97–117.

Robinson, D. (2015). National Water Sports Centre Remains Partially Closed After ‘Spill’ in River Trent. [online] Nottingham Post. Available at: http://www.nottinghampost.com/National-Water-Sports-Centre-remains-partially/story-27756265-detail/story.html [Accessed 18 January 2016].

Robson, C. (2002). *Real World Research*. 2nd. ed. Oxford: Blackwell Publishers Ltd.

Rogerson, P. (2001). *Statistical methods for geography*. London: Sage.

Rolfe, J., Bennett, J. and Louviere, J. (2000). Choice modelling and its potential application to tropical rainforest preservation. *Ecological Economics*, 35(2), pp.289-302.

Rollins, K. and Lyke, A. (1998). The case for diminishing marginal existence values. *Journal of Environmental Economics and Management*, 36(3), pp. 324-344.

Samples, K. C., Dixon, J .A., Gowen, M. M. (1986). Information Disclosure and Endangered Species Valuation. *Land Economics*, 62, pp.306-312.

Scarpa, R., Willis, K. G. and Acutt, M. (2005). Individual-specific welfare measures for public goods: a latent class approach to residential customers of Yorkshire Water. In: P. Koundouri, ed., *Econometrics Informing Natural Resource Management: Selected Case Studies. New Horizons in Environmental Economics Series,* 1st ed. Cheltenham: Edward Elgar Publishing, pp.316-337.

Scarpa, R., Gilbride, T. J., Campbell, D. and Hensher, D. A. (2009). Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European Review of Agricultural Economics*, 36(2), pp. 151-174.

Schultze, W. D., Brookshire, D. S., Walther, E. G. and MacFarland, K. K. (1983). The Economic Benefits of Preserving Visibility in the National Parklands of the Southwest. *Natural Resources Journal*, 23(1), pp. 149-174.

Scottish Executive. (2002). *Costs and benefits of implementation of the EC water framework directive (2000/60/EC) in Scotland. Final Report to the Scottish Executive*. Edinburgh: Scottish Executive.

Scottish Water. (2004). Strategic review of charges 2006-2010: Scottish Water's draft business plan public summary. [online] Available at: http://www.watercommissioner.co.uk/UserFiles/Documents/Scottish%20Water%201st%20draft%20business%20plan%20public%20summary.pdf [Accessed 13 April 2008].

Severn Trent Water. (2005). Minworth sewage treatment works. [online] Available at: http://www.stwater.co.uk/upload/pdf/Minworth.pdf [Accessed 2 December 2007].

Shanks, O., Green, H., Korajkic, A. and Field, K. (2016). Overview of Microbial Source Tracking Methods Targeting Human Fecal Pollution Sources. In: M.

Yates, C. Nakatsu, R. Miller and S. Pillai, eds., *Manual of Environmental Microbiology*, 4th ed. Washington, D.C.: ASM Press, pp. 3.4.3-1-3.4.3-8.

Shepherd and Wedderburn LLP. (2016). Brexit Analysis Bulletin Water and Waste Water May 2016. [online] Available at: https://www.shepwedd.co.uk/sites/default/files/Brexit%20Analysis%20Bulletin%20Water%20and%20Waste%20Water.pdf [Accessed 28 July 2016].

Shonkwiler, J. S. and Shaw, W. D. (2003). A finite mixture approach to analyzing income effects in random utility models: reservoir recreation along the Columbia river. In: N. Hanley, W. D. Shaw and R. E. Wright, eds., *The new economics of outdoor recreation,* 1st ed. Cheltenham: Edward Elgar Publishing, pp. 268-278.

Shrestha, R. K., Seidl, A. F. and Moraes, A. S. (2002). Value of recreational fishing in the Brazilian Pantanal: a travel cost analysis using count data models. *Ecological Economics*, 42(1–2), pp. 289–299.

Shuval, H. (2003). Estimating the global burden of thalassogenic diseases: human infectious diseases caused by wastewater pollution of the marine environment. *Journal of Water and Health*, 1(2), pp. 53-64.

Smith, V. K. and Desvousges, W. H. (1986). *Measuring water quality benefits*. Boston: Kluwer-Nijhoff Publishers.

Smith, G. M. and Fuller, R. M. (2002). Land cover map 2000 and meta data at the land parcel level. In: G. M. Foody and P. M. Atkinson, eds., *Uncertainty in Remote Sensing and GIS*. London: Wiley and Sons, pp.143-153.

SNIFFER. (2006a). Diffuse pollution screening tool: Stage 3, Project FWD77, 2006. [online] Scottish and Northern Ireland Forum for Environmental Research. Available at: http://www.sniffer.org.uk/ [Accessed 20 May 2009].

SNIFFER. (2006b). Provision of a screening tool to identify and characterise diffuse pollution pressures: Phase II, Project FWD19, 2006. [online] Scottish and Northern Ireland Forum for Environmental Research. Available at: http://www.sniffer.org.uk/ [Accessed 20 May 2009].

SPSS Inc. (2006). *SPSS for Windows* (version 15.0) [computer software]. Chicago: SPPS Inc.

State of Maine Department of Marine Resources. (2016). Fecal Coliform. [online] Available at: http://www.maine.gov/dmr/shellfish-sanitation-management/programs/growingareas/coliform.html [Accessed 16 November 2017].

Stapleton, C. M., Sanders, B., Kay, D., Anthony, S., Lyons, H., Wilson, J., Greaves, J., Dunhill, I., Miles, P. and Wither, A. (2003). The effects of destocking on microbial delivery from catchment diffuse sources in an area affected by foot and mouth disease (The Caldew study). In: M. Bruen, ed*., Diffuse Pollution and Basin Management. Proceedings of the 7th International Specialised IWA Conference.* Dublin: International Water Association, pp. 6-29.

Stapleton, C. and Kay, D. (2006). *Assessment of catchment derived and point sources of faecal indicator organisms to the coastal waters off Pwllheli. PRISM - Predictive Irish Sea Models. End of project report.* Aberystwyth: Catchment to Coast Research Centre, University of Wales.

Stapleton, C. M., Wyer, M. D., Kay, D., Crowther, J., McDonald, A. T., Walters, M., Gawler, A. and Hindle, T. (2007a). Microbial source tracking: a forensic technique for microbial source identification? *Journal of Environmental Monitoring*, 9(5), pp. 427-439.

Stapleton, C. and Kay, D. (2007b). “Modelling diffuse sources of faecal indicator organisms: An introduction for ChREAM project members*”,* presentation to ChREAM project members, 18 October 2007, University of East Anglia, Norwich.

Stapleton, C. M., Wyer, M. D., McDonald, A. T., Kay, D., Greaves, J., Wither, A., Watkins, J., and Francis, C. (2008). Quantitative catchment profiling to apportion faecal indicator organism budgets for the Ribble system, the UK’s sentinel drainage basin for Water Framework Directive research. *Journal of Environmental Management*, 87(4), pp. 535-550.

StataCorp, L. P. (2013). *Stata Statistical Software* (version 13.1) [computer software]. College Station, Texas: StataCorp, L. P.

Statistical Innovations. (2014). *Latent GOLD* (version 5.1) [computer software]. Belmont, MA: Statistical Innovations.

Stoner, J. D. (1978). *Water-quality indices for specific water uses. U.S. Geological Survey Circular, 770*. Arlington, VA: Department of the Interior, Geological Survey.

Sukharomana, R. and Supalla, R. (1998). Effect of risk perception on willingness to pay for improved water quality. *Presentations, Working Papers, and Gray Literature: Agricultural Economics,* Paper 46.

Swait, J. (1994). A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. *Journal of Retailing and Consumer Services*, 1(2), pp. 77–89.

Swait, J. and Adamowicz, W. (1996). *The effect of choice environment and task demands on consumer behavior: discriminating between contribution and confusion. Staff paper 96-09.* Edmonton, Canada: Department of Rural Economy, University of Alberta.

Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: an analysis and review. *International Journal of Forecasting*, 16(4), pp. 437-450.

Tasker, J. (2016). Defra minister George Eustice survives reshuffle. [online] Farmers Weekly. Available at: http://www.fwi.co.uk/news/defra-minister-george-eustice-survives-reshuffle.htm [Accessed 22 July 2016].

Thompson, G. C. (1966). The evaluation of public opinion. In: B. Berelson and M. Janowitz, eds., *Reader in public opinion and communication*, 2nd ed. New York: Free Press, pp. 7–12.

Thurston, J. A., Gerba, C. P., Foster, K. E. and Karpiscak, M. M. (2001). Fate of indicator microorganisms, giardia and cryptosporidium in subsurface flow constructed wetlands. *Water Research*, 35(6), pp. 1547-1551.

Tian, Y. Q., Gong, P., Radke, J. D. and Scarborough, J. (2002). Spatial and temporal modeling of microbial contaminants on grazing farmlands. *Journal of Environmental Quality*, 31(3), pp. 860-869.

Train, K. E. (1998). Recreation demand models with taste differences over people. *Land Economics*, 74(2), pp. 230–239.

Train, K. E. (1999). Mixed logit models for recreation demand. In: C. Kling and J. Herriges, eds., *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*. Northampton, M.A.: Edward Elgar Press.

Train, K. E., Hensher, D. and Shore, N. (2005). Households' willingness to pay for water service attributes. *Environmental and Resource Economics*, 32(4), pp. 509–531.

Turner, R. K., Pearce, D. and Bateman, I. J. (1994). *Environmental economics: an elementary introduction.* London: Harvester Wheatsheaf.

Turner, R. K. (2007). Limits to CBA in UK and European environmental policy: retrospects and future prospects. *Environmental and Resource Economics*, 37(1), pp. 253-269.

Tyrväinen, L. (1997). The amenity value of the urban forest: an application of the hedonic pricing method. *Landscape and Urban Planning*, *37*(3), pp. 211–222.

UKTAG. (2005). EU Water Framework Directive. [online] UK Technical Advisory Group. Available at: http://www.wfduk.org/about_wfd/ [Accessed 2 May 2008].

UKTAG. (2008). *Water Framework Directive (2008) Phase 1 Report - Environmental standards and conditions for surface waters.* London: UK Technical Advisory Group.

Upadhyay, N. (2002). *Management of the Lea Marston lakes, PhD Thesis*. Birmingham: University of Birmingham.

US EPA. (1986). *Ambient Water Quality Criteria for Bacteria 1986, EPA-440/5-84-002.* Washington, D.C.: United States Environmental Protection Agency.

US EPA. (2003). *Bacterial Water Quality Standards for Recreational Waters (Freshwater and Marine Waters) Status Report EPA-823-R-03-008.* Washington, D.C.: United States Environmental Protection Agency Office of Water.

US EPA. (2007). Watershed Protection Programs. [online] Available at: http://epa.gov/region02/water/nycshed/protprs.htm [Accessed 5 February 2008].

US EPA. (2012). *Recreational water quality criteria, 820-F-12-058.* Washington, D.C.: United States Environmental Protection Agency Office of Water.

US EPA. (2015). Water Quality Criteria. [online] Available at: http://water.epa.gov/scitech/swguidance/standards/criteria/ [Accessed 22 April 2015].

Van Houtven, G., Powers, J. and Pattanayak, S. K. (2007). Valuing water quality improvements in the United States using meta-analysis: Is the glass half-full or half-empty for national policy analysis? *Resource and Energy Economics*, 29(3), pp. 206–228.

van Teijlingen, E., Hundley, V. (2001) University of Surrey Social Research Update 35, The importance of pilot studies. [online] Available at: http://aura.abdn.ac.uk/bitstream/handle/2164/157/SRU35%20pilot%20studies.pdf?sequence=1&isAllowed=y [Accessed 15 July 2017].

Vaughan, W. J. (1981). The Water Quality Ladder. In: R. C. Mitchell and R. T. Carson, eds., *An experiment in determining willingness to pay for national water quality improvements. Report to the US Environmental Protection Agency*. Washington, DC: Resources for the Future, Inc., appendix II.

Veisten, K., Hoen, H. F., Navrud, S. and Strand, J. (2004). Scope insensitivity in contingent valuation of complex environmental amenities. *Journal of Environmental Management*, 73(4), pp. 317-331.

Vinten, A. J. A., Lewis, D. R., McGechan, M., Duncan, A., Aitken, M., Hill, C. and Crawford, C. (2004). Predicting the effect of livestock inputs of *E. coli* on microbiological compliance of bathing waters. *Water Research*, 38(14), pp. 3215-3224.

Vossler, C.A., Doyon, M. and Rondeau, D. (2012). Truth in consequentiality: theory and field evidence on discrete choice experiments. *American Economic Journal: Microeconomics*, 4(4), pp. 145-171.

Vrhovsek, D., Kukanja, V. and Bulc, T. (1996). Constructed wetland for industrial waste water treatment. *Water Research*, 30(10), pp. 2287-2292.

Warn, A. E. (1987). SIMCAT-A catchment simulation model for planning investment for river quality. In: M. B. Beck, ed., *Systems Analysis in Water Quality Management: Proceedings of a Symposium Sponsored by the IAWPRC Specialist Group.* Oxford: Pergamon Press, pp. 211-218.

Water Industry Network Wales. (2007). Cost Effectiveness Analysis of Implementing the Water Framework Directive. [online] Available at: http://waterindustrynetworkwales.blogspot.com/2007/10/cost-effectiveness-analysis-of.html Accessed 6 March 2008].

Waters, N. M. (1997). Spatial Interpolation. [online] University of Calgary. Available at: http://www.geog.ubc.ca/courses/klink/gis.notes/ncgia/u40.html [Accessed 20 May 2008].

Wattage, P., Mardle, S. and Pascoe, S. (2005). Evaluation of the importance of fisheries management objectives using choice-experiments. *Ecological Economics*, 55(1), pp.85-95.

Watts, S. and Stenner, P. (2005). Doing Q Methodology: theory, method and interpretation. *Qualitative Research in Psychology*, 2(1), 67–91.

Watts, S. and Stenner, P. (2012). *Doing Q methodological research: Theory, method and interpretation*. London: Sage Publications, Ltd.

Wedel, M. and Kamakura, W. A. (2012). *Market segmentation: Conceptual and methodological foundations* (Vol. 8). New York: Springer Science and Business Media.

Weikard, H. P. (2002). “The existence value does not exist and non-use values are useless”, paper presented at the Annual Meeting of the European Public Choice Society, 6 April 2002, Belgirate, Italy.

Whitelegg, J. (1993). *Transport for a sustainable future: the case for Europe*. London: Belhaven Press.

WHO. (2003). *Guidelines for safe recreational water environments. Volume 1: Coastal and fresh waters*. Geneva: World Health Organization.

Wilkinson, J., Jenkins, A., Wyer, M. and Kay, D. (1995). Modeling faecal coliform dynamics in streams and rivers. *Water Research*, 29(3), pp. 847-855.

Wilkinson, J., Kay, D., Wyer, M. and Jenkins, A. (2006). Processes driving the episodic flux of faecal indicator organisms in streams impacting on recreational and shellfish harvesting waters. *Water Research*, 40(1), pp. 153-161.

Wither, A., Greaves, J., Dunhill, I., Wyer, M., Stapleton, C., Kay, D., Humphrey, N., Watkins, J., Francis, C., McDonald, A. and Crowther, J. (2005). Estimation of diffuse and point source microbial pollution in the Ribble catchment discharging to bathing waters in the north west of England. *Water Science and Technology*, 51(3-4), pp. 191-198.

Woods, D. R., Green, M. B. and Parish, R. C. (1984). Lea Marston Purification Lake: Operational and River-Quality Aspects. *Water Pollution Control*, 83(2), pp. 226-242.

Wyer, M., Kay, D., Stapleton, C., Francis, C., Watkins, J. and Rushby, L. (2001). *Faecal Indicator Organism Sources and Budgets for Irvine Bay, Ayrshire: A report to West of Scotland Water and the Scottish Environment Protection Agency, South Ayrshire Council and North Ayrshire Council*. Aberystwyth: Centre for Research into Environment and Health.

Yatchew, A. and Griliches, Z. (1985). Specification error in probit models. *The Review of Economics and Statistics*, 67(1), pp. 134–139.

Young, K., Morse, G. K., Scrimshaw, M. D., Kinniburgh, J. H., MacLeod, C. L. and Lester, J. N. (1999). The relation between phosphorus and eutrophication in the Thames catchment, UK. *Science of the Total Environment*, 228(2), pp. 157–183.