VALUING HEALTH AND AIR QUALITY USING STATED PREFERENCE METHODS

By

ALAN DIENER, B.B.A., M.A.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy (Economics)

McMaster University

VALUING HEALTH AND AIR QUALITY
USING STATED PREFERENCE METHODS

DOCTOR OF PHILOSOPHY (1999) McMaster University
(Economics) Hamilton, Ontario

AUTHOR: Alan Diener, M.A. (Queens' University)

THESIS COMMITTEE: R. Andrew Muller (Supervisor)
Bernie O'Brien
A. Leslie Robb
Gregory Stoddart

NUMBER OF PAGES: xiii, 192

In memory of my parents, Leon and Golda Diener

ABSTRACT

The objectives of this thesis were to examine the use of stated preference methods in the valuation of health and environmental attributes. The first objective was to investigate the current state of research of non-market valuation of health care interventions. The second objective was to examine the potential for using conjoint analysis to value the health and non-health benefits of environmental improvements.

A case study was conducted in which conjoint analysis was employed in order to value the attributes of air quality changes in the Regional Municipality of Hamilton-Wentworth. The objectives of the case study were (i) to obtain accurate information about the willingness-to-pay and the health trade-offs of the attributes, and (ii) to gather evidence of the nature and effect of cognitive difficulties and to explore the use of ranking information.

In the case study, a total of 1908 surveys were mailed out to households. Four attributes were employed - health effects, black fallout, bad odours, and poor visibility. A fifth attribute - property taxes - was used so that willingness-to-pay could be estimated. Three levels of each alternative were employed: The current situation, one-third better, and one-third worse. Respondents were asked to rank nine choice sets each consisting of four alternatives.

Of the 515 surveys were returned, only 115 respondents completely ranked all of the choice sets without choosing a dominated alternative (a dominance violation). Hence, the task was cognitively difficult. Incorporating both first and second ranked choices increased the precision of the results.

Conjoint analysis ranked data can provide useful estimates of the trade-offs between attributes of air quality. More research is necessary to evaluate the sources of cognitive difficulties so that policy makers can be more confident with the results. The results from a conjoint analysis study can provide a valuable starting point in any economic evaluation of public goods, particularly health and the environment.

ACKNOWLEDGEMENTS

This thesis could not have been accomplished without the help and participation of many people. I am grateful to my supervisor R. Andrew Muller for taking me on as graduate student and particularly for initiating my involvement with the Hamilton-Wentworth Air Quality Initiative (HAQI) who provided the funding for this survey which was the focus of part of this research.

I am indebted to the members of my thesis committee, A. Leslie Robb, Bernie O'Brien, and Gregory Stoddart. They along with professor Muller provided insightful guidance at all stages of the thesis. They were always available to discuss my work. Many hours were spent in Professor Muller's office along with Professor Robb thoroughly planning and discussing the theory and the methods that were employed and I thank them for that. The entire committee reviewed several drafts of the thesis and provided valuable suggestions on how to best organize the material.

I am also thankful to Amiram Gafni, one of the co-authors, along with Professor O'Brien on the paper that became Chapter 3 of this thesis. We had many enlightening conversations on the use of the contingent valuation method in health care that contributed to my understanding of the topic.

The survey employed in this thesis was funded by the Ontario Ministry of the Environment and Energy and the Regional Municipality of Hamilton-Wentworth through the Hamilton-Wentworth Air Quality Initiative. I am indebted to the members of the Aesthetic, Odours, and Economics Workgroup of HAQI for providing comments on the questionnaire. The members included George Vance, Eric Loi, Frank Dobroff, Susan Elliot, Robert Hall, John Jarvie, and Chris Cincar. Others who I am indebted to in this regard include Don Cole, Paul Kreuger, Douglas Frost, Walter Bramberger and Jim Rice. Michelle Vickers helped coordinate the mailing of the questionnaires.

My colleagues and staff at The Department of Preventive and Societal Medicine at the University of Nebraska Medical Center where I was employed during the latter months of completing this thesis must be acknowledged. In particular I would like to thank Joan Penrod for being available to discuss several issues, Laura Bashus, Fred Ullrich, and Angella Bowman for providing technical support, and Brandi Shay for proofreading the entire document.

While conducting the research for Chapter 3, Alan Diener was supported by a summer studentship from the Centre for Health Economics and Policy Analysis at McMaster University. Dr. O'Brien is supported by a career award in health sciences from the Medical Research Council and Prescription Drug Manufacturers Association of Canada. We are grateful for comments on an earlier draft of this paper from members of the Polinomics seminar at McMaster University, Stephen Birch of McMaster University, Cam Donaldson and other participants in the Health Economists Study Group meeting at Liverpool, UK, January 1997, and two anonymous reviewers.

I would like to thank my entire family for all of their unconditional support. This project could not have been completed without the devotion and support of my wife Shirley. She experienced the anxiety of completing a thesis and helped to keep me focussed. Her work as a unofficial research assistant and her many trips to the library to photocopy "urgently needed" papers is greatly appreciated. Finally, I would like to thank Zoe, just for being.

PREFACE

Chapter 3 of this thesis contains a previously published paper which was co-authored with Bernie O'Brien and Amiram Gafni of McMaster University. All of the research in the paper was conducted by Alan Diener. The original draft was written by Alan Diener who ultimately had the final responsibility with respect to what research should be included and how the paper should be presented.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

LIST OF TABLES (CONTINUED)

Chapter 1

INTRODUCTION AND OVERVIEW

1.1. Introduction

Many public programs and policies have consequences which may be complicated to evaluate. These include the assessment of new or existing health care interventions (such as new technologies, pharmaceuticals, etc.), and programs that affect the environment. Environmental programs may have multiple outcomes including both health effects and non-health outcomes which can be as diverse as loss of wetlands to bad odours. In order to assess such public programs, analysts can turn to economic evaluation. Economic evaluation can be a valuable tool in deciding the optimal allocation of societies' scarce resources.

In an economic evaluation the costs of a program must be compared to its consequences. In one form of economic evaluation, cost-benefit analysis (CBA) both the costs and consequences of the program are measured in monetary units. However, assigning dollar valuations to the benefits is not a simple task. One approach is to use stated preference methods in which people are surveyed and asked questions relating to the value that they place on the outcomes[1]. These survey approaches, which include the contingent valuation method (CVM) and

---

[1]Various taxonomies are used in different disciplines. For the purposes of this dissertation stated preference methods refer to survey-based methods in which individuals are asked specific questions relating to their preferences.

conjoint analysis (CA) are generally employed in order to infer estimates of respondents' willingness to pay (WTP) (or willingness-to-accept, WTA) for the program or the attributes of the program in question.

1.2. Objectives and overview

Contingent valuation is a stated preference technique that has become a popular tool in the evaluation of health, health care, and the environment (Diener et al, 1998, Carson, 1995). More recently another technique, conjoint analysis, has been implemented as a method that can be used to value the multiple outcomes or attributes of programs simultaneously. Examples of programs with multiple attributes include a pharmaceutical that causes several side-effects or an environmental program that may result in a decrease in hospital admissions for asthma, as well as decreases in dustfall, bad odours, and poor visibility.

This thesis is an investigation of using stated preference methods for valuing the benefits of the non-marketed attributes of health and environmental programs for use in cost-benefit analysis. Since, in general, there exists no explicit market for these attributes we must rely on other methods such as contingent valuation and conjoint analysis in order to value them.

The objectives of this thesis are (a) to investigate the current state of research of non-market valuation of health care interventions and (b) to investigate the potential for using conjoint analysis for valuing health and non-health benefits of environmental improvements. The first objective was achieved by the comprehensive literature survey in chapter 3. This analysis uncovered a number of potential limitations - namely with respect to the valuation of multi-attribute goods. The second objective was achieved by conducting a case study applying these techniques to the valuation of air quality changes in the Regional Municipality of Hamilton-Wentworth.

While there is evidence that poor air quality in the Region has lead to an increase in hospitalizations for cardio-respiratory diseases, there is also evidence that the "nuisance" effects such as black particulate deposition (black fallout) and bad odours are of concern to the residents (HAQI, 1997; Elliott et al., 1997). Hence, any evaluation of a program that would result in a better level of air quality in Hamilton-Wentworth should focus on the multi-attribute facets of the benefits. The specific objectives of the case study were (i) to obtain as accurate information as possible about health trade-offs and willingness-to-pay for the attributes of air quality improvements in Hamilton-Wentworth, taking into account the limited budget for the study, and (ii) to contribute to the methodology of the conjoint analysis method by gathering evidence on the nature and effect of cognitive difficulties as well as exploring the use of ranking information.

The next chapter will briefly review the methodology of economic evaluation and will then go on to review the economic foundations behind the willingness-to-pay techniques that are employed. This will include a discussion of the consumer surplus concepts of compensating variation and equivalent variation. Finally, non-market valuation methods, which include contingent valuation and conjoint analysis, are reviewed.

Chapter 3 reports a comprehensive review and classification of all CVM studies that were undertaken to analyse health care interventions for use in CBA in the period 1984 to June 1996.

Chapter 4 presents a detailed exposition of conjoint analysis and how it can be used to value multi-attribute goods in general and the attributes of air quality in particular. This includes a discussion about how CA experiments are designed and the economic models that are used in order to estimate willingness-to-pay. The theoretical underpinnings of the technique that allowed discrete rankings to be employed are then presented before the chapter concludes with a review of conjoint analysis studies pertaining to health and the environment.

Chapters 5, 6, and 7 focus on the conjoint analysis survey that was conducted in order to estimate WTP and other trade-offs for the attributes of improved air quality in the Regional Municipality of Hamilton Wentworth - namely the health effects, and the number of days with bad odours, black fallout and poor visibility. Chapter 5 provides the details of the study design and presents the economic models that were employed. Chapter 6 presents the general survey results including tests of cognitive difficulty. The willingness-to-pay and health trade-off results are reported in Chapter 7 which also explores the gains of employing multiple rankings and includes tests of construct validity.

Finally, Chapter 8 concludes with a discussion of the overall findings of the dissertation and presents areas of possible future research.

Chapter 2

ECONOMIC EVALUATION OF HEALTH AND THE ENVIRONMENT

2.1. Economic evaluation and cost-benefit analysis

Economic evaluation is an important tool which can be used by decision makers in determining the appropriate uses of societies' scarce resources. For instance, an economic evaluation can help to decide whether or not to approve a new health care intervention, to choose the best amongst a group of alternative health care therapies or alternative programs, or to decide whether or not to implement a new program that will increase local air quality. Note that the two types of public programs that are being discussed here can both cause health effects and environmental effects. It is clear that a public health program (health intervention, etc.) will yield health effects. Environmental, or air quality programs (i.e. policies that affect the physical environment), can yield both health effects and environmental effects. Moreover, health interventions can create further health effects via environmental effects. For example, the incineration of certain hospital wastes may produce pollutants that cause further health effects. It should be noted that this latter effect has not been studied in depth nor will it be here. The focus of this dissertation is on the health effects of health interventions and the health and environmental effects of environmental programs.

An economic evaluation is concerned with both the inputs and the outputs of a course of action such as a public project or program (or group of projects). The inputs are, in general, the costs of the resources employed in undertaking the specific program. The outputs are the consequences of the program. In one form of economic evaluation, cost-benefit analysis (CBA), both the costs and consequences (or benefits) of the program or group of programs are measured in monetary terms.

Other forms of economic evaluation include cost-effectiveness analysis (CEA) and cost-utility analysis (CUA). Costs will be the same in both of these alternatives. In CEA outcomes are measured in their natural units resulting in measures such as quantity of dollars needed to reduce serum cholesterol by 10%. Hence, it is difficult to compare programs with different outcomes using CEA. In CUA outcomes are converted to utility measures such as quality adjusted life years (QALYs) or healthy year equivalents (HYEs) resulting in measures such as quantity of dollars needed to obtain 10 QALYs. Hence, the comparison of health interventions that have different outcomes is more easily facilitated. The limitation of CUA, however, is that it is limited to only morbidity and mortality outcomes. For more on economic evaluation and health care see Drummond *et al.* (1997).

One advantage of cost-benefit analysis is that since both the costs and benefits of any program are measured in dollars it is relatively straightforward to compare programs that have distinctly different outcomes. For instance, a health care intervention that decreases one's risk of coronary heart disease can be compared with a new anti-depressant that reduces the number of days of a side effect such as coughing. Alternatively, these interventions can be compared to a program that will make highways safer resulting in a reduction in auto injuries. Some programs, such as those that improve air quality, will have several distinct outcomes. For example, in the Regional Municipality of Hamilton-Wentworth air pollution has been responsible for health outcomes such as increases in cardio-respiratory diseases as well as non-health outcomes including black fallout, bad odours, and poor visibility (HAQI, 1997). One drawback of CBA is that defining and calculating the benefits of such programs are not clear-cut tasks. In the remainder of this chapter the theoretical underpinnings of CBA will be reviewed and the empirical methods of calculating WTP values will be presented.

2.2. Measures of consumers surplus

The conceptual foundation of CBA is in welfare economics; specifically the Kaldor-Hicks criterion as a hypothetical compensation test between the value of utility gains (to gainers) from a program compared to the utility losses (to losers) (Mishan, 1988). While the costs of a health care or environmental program are usually straightforward to calculate, valuing the benefits can be an onerous task. In general the notion of consumers surplus is used as the measure of the benefits or utility.

> The consumer's surplus is just the same sort of hypothetical magnitude; it involves the question "what is the maximum amount which the consumer would be willing to pay for the particular quantity of the particular commodity if he were given the choice between having the quantity on such terms or not at all?" (Hicks, 1941).

The two different monetary measures of consumer surplus - compensating variation and equivalent variation - can be estimated using either willingness-to-pay (WTP) or willingness-to-accept (WTA) (Hicks, 1943; Johansson, 1993). The main difference between the two is that compensating variation (CV) is evaluated at the initial (pre-program) level of welfare, while equivalent variation (EV) is evaluated from the new (post program) level of welfare. The compensating and equivalent variations can be formally derived by examining an individuals indirect utility function.[1] The compensating variation is the amount of money such that

$$\begin{aligned} WTP&: \quad V(Q^1, Y^0 - CV) = V(Q^0, Y^0); \; Q^1 > Q^0 \\ WTA&: \quad V(Q^1, Y^0 + CV) = V(Q^0, Y^0); \; Q^1 < Q^0 \end{aligned} \tag{1}$$

where Q^0 and Y^0 denote the initial levels of the environmental and/or health goods

---

[1]The individual chooses the quantity of private goods, x, that maximizes her utility function U(x,q), (where x is a vector of private goods and q is a vector of public goods which could be interpreted as the quality of the environment or the level of health care offered.) given an exogenous level of income and prices and level of public goods. Substituting x(p,y,q), the optimal choice, into the utility function we obtain the individual's indirect utility function.

$$V = U[x(p,y,q),q] = V(p,y,q)$$

and income respectively, and Q^1 and Y^1 the final levels[2] (Johansson, 1993). If the new level of the good is greater than the initial level, the compensating variation gives the maximum amount of money that the individual (or household) can give up while leaving it just as well off as it was before the improvement in environmental quality or health[3]. In other words, the compensating variation is the willingness-to-pay for the improvement in utility that the greater amount of the good will bring. On the other hand, if the new level of the good is less than the initial level, the compensating variation is the minimum amount of money that must be given to the individual (or household) to compensate for the loss in utility. In the latter case the compensating variation measures willingness-to-accept (see figure 2.1[4]). The other monetary measure that can be estimated, equivalent variation, is the amount of money such that

$$\begin{aligned} WTA&: \quad V(Q^1,Y^0)=V(Q^0,Y^0+EV);\; Q^1>Q^0 \\ WTP&: \quad V(Q^1,Y^0)=V(Q^0,Y^0-EV);\; Q^1<Q^0 \end{aligned} \qquad (2)$$

The main difference between compensating and equivalent variation is that the compensating variation uses the original utility level as its base, while the equivalent variation uses the new utility level as its base. With an increase in the quantity of the good, the equivalent variation is the minimum amount of money

---

[2]Prices are assumed equal to 1 for simplification. It is also assumed that more of the good is better.

[3]Hicks (1943) refers to this as the *quantity* compensating variation.

[4]All figures and tables can be found at the end of each chapter.

that must be given to the individual to make it as well off as they could have been with the increase in utility obtained from the greater quantity of the good. In other words it states the willingness-to-accept in order to not proceed with the increase in level of the good or service. With a decrease in the level of the good or service, the equivalent variation states the maximum amount of money that the individual would be willing to give up in order for the intended decrease not to occur, or the willingness-to-pay (see figure 2.2).

2.3. Non-market valuation methods

Programs that result in improvements in health or the environment generally produce outcomes that are not easily valued. For example, the attributes of air quality including improvements in health status or decreases in black fallout are not goods that are purchased in a market setting, with their prices determined by supply and demand. Hence, we must rely on non-market methods in order to value them. Non-market valuation methods can be broadly classified as falling into two categories - stated preference and revealed preference. In general, revealed preference, or indirect, methods focus on inferring willingness-to-pay from data on actual, or observed, behaviour. For example, it is possible to infer some measure of willingness-to-pay for improved air quality indirectly through market phenomena such as housing prices. Under this method, known as hedonic pricing, we assume that individuals who live in areas that are less desirable due to high levels of air pollution, for example, will pay less for their housing (Freeman, 1995). Other examples include the travel cost method and the wage-risk method.

The travel cost method can be used for a non-market commodity, such as outdoor recreation, where the cost of consuming the commodity includes the travel costs to the recreational site (Bockstael, 1995). In wage-risk studies the goal is to examine the relationship between particular health risks associated with a hazardous job and the wage rates that individuals require to accept the job (Marin and Psacharopolous, 1982). All of these methods attempt to measure values indirectly, through related market mechanisms.

Recently, however, stated preference approaches have become increasingly popular. While different disciplines may define stated preference in various ways, in the context of this dissertation stated preference refers to those methods which use surveys to estimate WTP values. The two relevant stated preference approaches in the field of economic evaluation are the contingent valuation method (CVM) and conjoint analysis (CA).

CVM is a well-established technique in the areas of economics of safety, transport, and the environment (Bishop *et al.*, 1995; Jones-Lee *et al.*, 1985; Mitchell and Carson, 1989) and is being used with growing frequency in health economics (Diener *et al.*, 1998). In a typical CV survey respondents are presented with a carefully described but hypothetical choice between two alternatives. They are then asked directly how much they would be willing to pay for their preferred alternative, or conversely, how much they would be willing to accept to forego their preferred alternative.[5] For example, an environmental CVM study might ask WTP questions for the health benefits of cleaner air due to some program: The exercise proceeds on the hypothetical *contingency* that such a market exists to determine what consumers would be willing to pay. The technique, however, is not without controversy though there seems to be a general opinion that carefully designed studies can reveal useful information about the order of magnitude of environmental benefits.[6] Chapter 3 will provide more details about this method, specifically with regard to its application in valuing health care programs.

Conjoint analysis generally refers to techniques that are used to establish the impacts of individual attributes on the overall utility of a good or service. Generally, in a conjoint analysis survey, respondents are asked to rank or rate a group of alternatives, each defined by various levels of several attributes. If a payment vehicle is included as one of the attributes, willingness-to-pay can be estimated. The application of CA in the fields of both health and environmental economics has, however, been limited. In the field of health economics, it has been used, amongst other applications, to measure consumer preferences and utility in the provision of health care services (Ryan, 1996). It has also been used in environmental economics to measure the value of wilderness areas

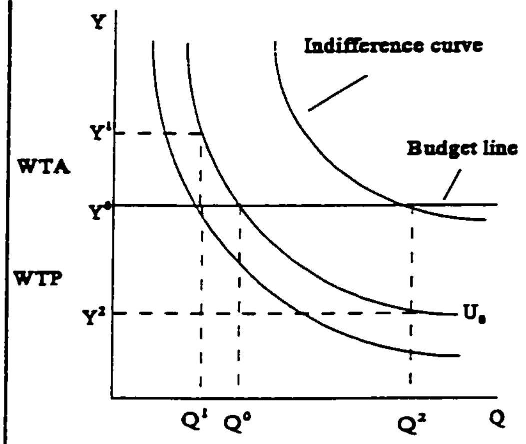
---

[5]There are several different methods that can be used to pose such questions. The example here is of an open-ended approach. Other approaches include bidding-games, payment cards, and discrete choice (take-it-or-leave-it). These methods are discussed in more detail in Chapter 3.

[6]For more on the debate of the validity of contingent valuation see Portney 1994, Hanneman 1994, and Diamond and Hausman 1994.

(Adamowicz *et al.*, 1994). Chapter 4 will provide an overview of CA, including a brief history of the methodology as well as the steps involved in conducting such an analysis.

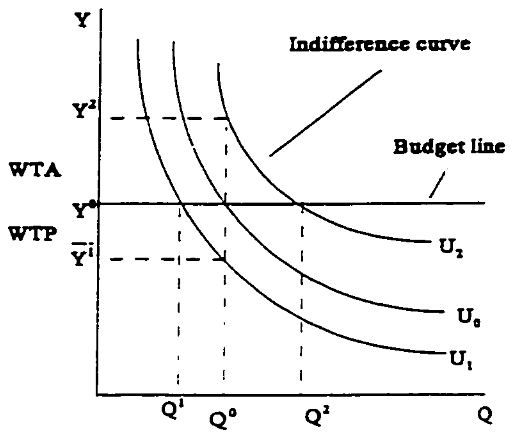
Figure 2.1. Compensating variation of a quantity change



Q is increased from Q^0 to Q^2
- the CV, measured by WTP is Y^0-Y^2
- i.e. the amount of money such that the individual can achieve the initial level of utility U_0

Q is decreased from Q^0 to Q^1
- the CV, measured by WTA is Y^1 -Y^0

Figure 2.2. Equivalent variation of a quantity change



Q is increased from Q^0 to Q^2
- the EV, measured by WTA is Y^2-Y^0
- i.e. the amount of money such that the individual can achieve the new level of utility U_2, without having Q change

Q is decreased from Q^0 to Q^1
- the EV, measured by WTP is Y^0 -Y^1

Chapter 3

HEALTH CARE CONTINGENT VALUATION STUDIES: A REVIEW AND CLASSIFICATION OF THE LITERATURE*

3.1. Introduction

A number of CVM studies have now appeared in the health care program evaluation literature with valuations in a range of diseases and treatments including hypertension, in-vitro fertilization [44], screening for cystic fibrosis [12,38,39] , and choice of anti-depressant drugs [47]. Although some reviews of the health care CVM literature have been published [19,41,45], these reviews have not been comprehensive and have mainly sought to highlight some methodologic points by reference to selected published studies.

In this study we undertake a comprehensive review and classification of health care CVM studies published in a broad range of journals covering economics, medicine and health services research. Why is such a survey necessary and useful? As indicated in an earlier study [45], there appears to be large variation

*This chapter is excerpted from:
Diener, Alan, O'Brien, Bernie, Gafni, Amiram. 1998. Health care contingent valuation studies: a review and classification of the literature. *Health Economics* (7)4: 313-326

The references follow the formatting that appeared in the original article.

among health care CVM studies in terms of the type of questions posed, in what format, and the sampling of respondents. Our aim is to help those who wish to read or design a CVM study gain a better understanding of how and, when possible, why published CVM studies vary in terms of methods employed. While some sources of variation may be due to pragmatic measurement issues (eg. use of a mail survey versus in-person interview) other variation is more fundamental (eg. use of willingness-to-accept versus willingness-to-pay). Classifying studies according to such design features will yield a quantitative summary of the literature and provide a basis for critical appraisal of the CVM method and how it is being used in health care.

3.2. Methods

3.2.1. Search criteria and methods

We sought to identify studies published in the health or economics literature in the period 1984-96 in which an attempt had been made to measure either willingness-to-pay or accept in the context of the evaluation of a health care program. Although such WTP/WTA studies might go by the contemporary collective description of contingent valuation method (CVM) studies this label is relatively new. For this reason we employed a broad search strategy of computerized databases in medicine (MEDLINE, CINAHL) and economics (ECONLIT) using keywords of both 'willingness-to-pay' and 'willingness-to-accept' in addition to 'contingent valuation'. We searched both keywords and textword strings within the title and abstract. Other databases searched included the Science Citation Index, Social Science and Humanities Citation Index, Current Contents, and Bibliography of Bioethics. Further studies were found by examining other bibliographies [6] and review papers, as well as personal bibliographies and contacts. The journals from which papers were retrieved are listed in the appendix. Some working papers have also been included in the analysis. Such papers were found by personal contact with authors that had completed CVM studies in health care previously or were presented at conferences. These papers were all recently completed and may yet appear in journals. Also included were technical reports or government publications.

We defined the time period of search from 1984 onwards because prior to this date very few health care CVM studies had been published and in this year Thompson and colleagues published an important paper on WTP in arthritis [59]. Although we recognize that earlier health care CVM papers were published, the current flow of papers began in that year. Excluded from our review are papers that address theoretical issues concerning CVM methods that offer no empirical estimates of WTP or WTA, using original survey data.

3.2.2. Primary classification of all CVM studies retrieved

Retrieved papers were initially classified using five descriptive criteria: (1) was the study designed to address a methodological question or concerned with the estimation of WTP or WTA for various reasons?; (2) in what type of journal (eg. health economics or clinical) was the study published?; (3) did the study elicit responses using in-person interviews, by mail survey or phone survey?; (4) what was the content area of the study, both in terms of the disease area and the form of the intervention?; (5) was the (stated or implicit) objective of the study to assist in pricing and demand forecasts for a product or to assist with program evaluation (i.e. CBA)?

3.2.3. Secondary classification of 'CBA' studies by criteria of O'Brien and Gafni

The final criterion in the preceding paragraph is important for defining the group of studies that were eligible for a second classification exercise. O'Brien and Gafni [45] recently suggested a set of methodologic criteria applicable to health care CVM studies that were performed for the purpose of program evaluation or CBA. We used these criteria to further classify the CVM studies where the underlying objective was stated (or thought) to be program evaluation and CBA. These questions and considerations are reproduced here as Table 3.1. We explain the relevance and meaning of each criterion as we present results of the classification exercise in the next section.

3.3. Results

3.3.1. Characteristics of retrieved studies

Initial screening of the literature using our keywords and text strings identified just over 200 papers for potential inclusion. The majority of these papers were not included in our review because they were either theoretical essays with no empirical content or CVM applications to outcomes of programs other than health care. A total of 48 studies met our inclusion criteria and used some variant of the contingent valuation method to value a health care intervention.

General characteristics of the 48 studies are presented in Table 3.2. In order to detect any time trends we divided the time frame of interest into two six-year periods; 1984-1989 and 1990-1996. Growth in the application of CVM methods is evident with the majority of studies (38;79%) being published in the period 1991-96. The distribution of empirical studies by primary motivation of either estimation or methodological enquiry was similar with 19 (40%) of the papers motivated by methodological issues [4,9,10,13,25,26,33-36,46-50,54,57-59]. This distribution does not appear to change over the surveyed period. We also found that in many cases we had to infer the motivation for the study because objectives were not clearly stated. Furthermore, we found that a number of studies could be classified either way – e.g., Stalhammer [57], conducted an original contingent valuation survey, but the study's main motivation was to explore a specific methodological issue, which was obvious from the title.

We grouped the journals from which studies were retrieved into five general areas (see Table 3.3). The majority of CVM studies were published in health services research (17;35%) or clinical journals (13;27%), with a small minority in economic (non-health) journals (3; 6%). Although there has been growth in these applied CVM studies in the health economics literature the six reports in these journals constitute only 13% of the literature over the time period studied.

The majority of studies were mail surveys (25;52%) [4,8,10,12,13,16-18,22-

26,31,33-35,38,39,42,44,49,52,54,56] or in-person interviews (18;38%) [1-3,7,9,14,20,21,37,46-48,50,51,55,57-59]. Only four of the studies were telephone surveys [15,32,36,53]. For one of the studies retrieved it was unclear how the respondents were contacted [30].

Some of the studies did not address a particular disease, but rather examined the value of general health care services. Examples would include Birdsall, who valued the benefits of a new health worker in a rural village [3], and Eckerlund which examined the optimal value of the Swedish health care budget [15]. There were a total of 11(23%) such studies [3,8,10,15,17,20,21,42,50,53,55]. Of the remaining studies we collected data on the disease area studied and the type of intervention, either medical, surgical or pharmaceutical.

A variety of disease areas were studied including such far-ranging items as heart attack risk, depression, and chronic lung disease. We broadly grouped the studies as addressing respiratory diseases (3;6%) [4,48,49] and cardiovascular diseases(9;19%) [7,26,30,31,33-37,37], and cystic fibrosis screening (3;6%) [12,38,39], as these were the only diseases addressed by more than two of the studies. The rest of the studies are captured by the 'other' classification. Some of the studies further decomposed the health impacts into symptoms or side-effects, and valued them all separately [1,4,47,49,54]. The types of intervention employed was 31 (64%) medical [1-3,8-10,12-18,20,21,24,25,30,34,35,37-39,42,44,48,50, 52,53,55,56] , 2(6%) surgical [23,51], and 12(25%) pharmaceutical [4,22,26,30, 31,33,36,46,47,49,54,57]. A few of the studies, just conveyed that the change in health status would be due to a hypothetical treatment and did not explicitly state the type of intervention (4;8%) [7,32,58,59]. Some studies compared the results of two types of interventions for the treatment of the same disease [30,31].

3.3.2. Secondary classification of 'CBA' studies

From the total of 48 CVM papers retrieved a subset of 42 was identified where the stated or implied context for the study was CBA. Although very few of these studies were complete cost-benefit analyses, we refer to these as CBA studies because this is the presumed motivation for money measurement. Given that many reports were not explicit in their reasons for conducting a CVM survey, we had to infer a study's motivation from its context, methods, results or interpretation. On this basis we judged that six of the retrieved papers were performed for the purpose of pricing or marketing of a product and these were excluded from further consideration [14,16,18,22,42,54]. The remaining forty-two 'CBA' papers were then classified using the framework of O'Brien and Gafni [45]. The results of this secondary classification are presented in Table 3.4.

<u>What question do we want to answer?</u>

Whether the context for a CVM study is the introduction or removal of a program has implications for the type of money measurement used and the values elicited from consumers. Moreover, a comprehensive design for a CVM survey recognizes that not all consumers will necessarily be better off from the introduction of a new health care program; similarly, the removal of an existing program may not be a source of disutility to all.

The majority of the papers valued currently existing programs (24;61%) [2,9,10,12,13,15,17,20,23,24,26,30,31,33-35,38,39,44,50-53,56], while seventeen of the studies (37%) valued the creation of a possible new program [1,3,4,7,8,21, 25,32,36,46-49,55,57-59]. The latter included either a new technology or a new drug. In some cases it may have been a proposed, or even hypothetical, new drug or program.

Only 4 (9%) studies, in this subset, allowed a loss in utility from the intervention whether a program was being introduced or removed [2,10,12,37]. Lindholm [37], which was the only paper to examine the removal of a program, asked respondents for the minimum money payment they would accept in order for a health care program to be removed (thereby assuming that everyone would experience a reduction in utility if the program was removed). The only paper that allowed both a loss and a gain in utility due to the intervention was Berwick and Weinstein [2]. Their study attempted to elicit WTP for ultrasound information. Respondents were allowed to state values to both *avoid* and *obtain* the information. Some studies, while not allowing a loss in utility may have first asked the respondents whether they preferred one program over another and then asked them to state their WTP to have the preferred rather than the less preferred program [10].

What type of measure can we use?

The measurements of willingness-to-pay and willingness-to-accept can be used to estimate two different monetary measures of change in consumer surplus: compensating variation (CV) and equivalent variation (EV). CV is evaluated from the pre-program level of welfare, in contrast to EV, which is evaluated from the post-program level of welfare. In short, while CV and EV are both monetary measures of the utility change, WTP and WTA are two possible valuation methods that can be used to estimate either one of these monetary measures. More detail on these concepts is beyond the scope of this paper and readers are referred elsewhere [45].

Only seven of the 42 studies were explicit in stating whether the utility change they sought to value was CV or EV [4,7,21,33,48,49,53]. Inferring this criterion for the remaining studies we found that, overall, the vast majority of studies (40 of 42) conducted money measurement in the context of compensating variation. All of the studies were clear in stating the money measurement technique employed – either WTP or WTA and only two of the 42 papers, used WTA questions [23,37]. Cross-tabulations [23,52] of the consumer surplus measure being estimated (CV or EV) and the technique of money measurement (WTP or WTA) are presented in 3.5. From this table it is clear that the vast majority of studies have been CV-WTP designs. Of note, we found no study that had attempted an EV-WTA design. This is, perhaps, not surprising given the cognitive demand such questions would place upon respondents – i.e., stating the minimum compensation they would require to forego introduction of a proposed health care program.

What do we need to ask of whom?

A CVM study for a health care intervention can sample from among persons with the disease of interest, those at risk of the disease, those not at risk of the disease, or some combination of all these groups. The choice of who is surveyed, and why, should be based upon the conceptual foundation of how the resulting survey values can be used in CBA. In principle the scope of the survey should be to capture utility impacts arising out of the introduction or removal of the program irrespective of whether persons are users or non-users of the program. In other words, for CBA, we are interested in capturing both use and non-use values.

A total of 23 (55%) studies involved respondents who were either currently diseased or undergoing the treatment/intervention [1,2,7-9,15,20,23,24, 26,30,31,34-37,47,48,51,55,57-59]. Fifteen (36%) of the studies garnered responses from non-users -- those at future risk of obtaining the disease or in need of the intervention [3,10,12,13,17,25,33,38,39,44,46,49,50,52,56], -- while 4 (10%) of the studies asked respondents from the general population, which may include respondents from all of the aforementioned groups [4,21,32,53]. The study by Osmond [51] regarding different methods of wound closure for pediatric facial lacerations in children was based on WTP responses from parents and we viewed this as a household consumption decision.

Previous papers have discussed the sampling implication of two general valuation perspectives for health care CVM studies, the ex-post user-based perspective and an ex-ante insurance-based approach [19,45]. O'Brien and Gafni defined the ex-post user-based perspective as one "...where the respondent is asked to assume that he or she is at the point of consuming some unit of the program being evaluated," while the ex-ante insurance-based perspective assumes that "...a person at risk of a disease and consuming the treatment program, are asked for some variant of his or her maximum WTP as an insurance premium to have the program available." The perspective was whether or not the respondent was suffering from the disease of interest.

We initially attempted to classify the studies according to the aforementioned valuation perspectives. However, it became clear that a slight redefining of the valuation perspectives was in order. For example, eleven of the studies (27%) ascertained WTP *after* the entire intervention had taken place, and it was unclear where these studies would fall under this classification [1,8,9,13,20, 24,26,30,31,34,35]. Hence, we redefined the valuation perspective to refer to whether or not the intervention had taken place. In other words, an ex-ante study refers to one in which respondents were interviewed before undergoing the intervention, while an ex-post study refers to one in which the respondents were interviewed after they had experienced the intervention. Ex-ante studies can be further decomposed into ex-ante insurance based in which the respondents are assumed to be at risk of contacting a disease that would require the intervention, and ex-ante user-based in which the respondents already have the disease, but have yet to undergo the intervention. While the majority (32;75%) of the studies employed an ex-ante perspective [1-4,7,9,10,12,15,17,21,23,25,32,33,36-39,44, 46-53,55-59], only six (14%) of these used the ex-ante insurance-based approach [21,32,46,50,53,55].

What characteristics of the program are important for determining how it is valued?

Among the numerous special characteristics of health care as a commodity there are two aspects that are particularly important for how valuation scenarios for CVM are constructed. The first is the issue of whether program outcomes have been presented as being certain or uncertain. As argued by Gafni [19], an important advantage of presenting CVM valuation scenarios with associated uncertainty is that money valuations elicited will include risk preference. The second issue concerns whether CVM valuation scenario and payment vehicle frame the problem in terms of a public or private goods market, and, if private, whether the good exhibits consumption externalities (which could be physical or "caring"). There are few, if any, health care interventions that could be described as pure public goods, although a number of public health programs (such as water fluoridation and air quality improvement) generate benefits that are largely non-excludable and non-rival. When valuing such a good, it becomes necessary to frame the vehicle in such a way that the potential for "free riding" is recognized and the respondent knows whether or not other consumers would be required to finance the program. More commonly, there are a number of health care interventions, such as vaccination against communicable diseases, where consumption externalities exist and use of a purely private good payment scenario would lead to an underestimate of the total value of the program, if some citizens derived benefit from having others vaccinated.

The majority of papers, (32;76%) valued certain outcomes, and valued goods from a purely private market perspective (36;86%) -- i.e., without accounting for the possibility of consumption externalities. We classified only six studies as adopting some form of political market to frame the valuation scenario where factors such as externalities might be taken into account. A good example is the study by Lindholm, *et al*, [37] where WTP was elicited for a community-based cardiovascular disease prevention program using a referendum format with tax implications. The authors argue that this form of "tax-voting" payment scenario was both more realistic for respondents given that Swedish health care is tax-financed (thus reducing the so-called 'hypothetical bias' [40]) and the method permits expressions of value due to externalities with citizen A being willing to pay for health benefits to citizen B.

What question formats minimize bias and increase precision?

Four of the studies valued decomposed outcomes. Such an approach would focus on the separate components of a program and value them separately. For example, three of the papers broke down the outcomes to be evaluated into different side-effects from an intervention [1,47], while the other two, both conducted by the same authors, decomposed the outcome into drug safety versus drug efficacy [4,49]. Few of the side effects are common to the two papers. It would be interesting to examine what the differences may be in WTP for similar changes in side effects, from different sources.

Examining the survey question type employed in the analysis, we found that all of the question types were employed almost equally. The take-it-or-leave-it approach, in which each respondent is presented with a price and asked to state whether or not they are willing to pay that value, was used in 26% (11 studies) [4,15,17,26,32,34-37,49,56] of the studies overall, but over the latter half of the time frame it was employed in 36% of the studies. While the sample is rather small, it appears to be becoming a more popular approach. The open-ended approach may be the easiest question type to employ. In such a survey respondents are asked to state their maximum WTP (or WTA) without being given any suggestions. This method was employed 38% of the time (16 studies) [2,7,8,10,13,20,23,25,26,35,38,39,47,51,58,59].

3.4. Discussion

3.4.1. General Results

In this study we have reviewed the application of CV methods to health care questions by classification of the published literature. Our ability to classify studies is clearly conditional upon the extent to which authors have labelled or otherwise indicated their methods and study objectives. A key finding was the lack of explicit statements regarding study objectives and the methods employed. This lack of clarity in study presentation necessitated a number of judgments on our part to infer methods and objectives. While such inference introduces the potential for classification error, the possibility of contacting all study authors to clarify issues was beyond the scope of our study. Furthermore, our impression was that clarity in reporting of CV studies has not been improving over time.

This study showed that the ratio of methodological to empirical studies remained constant over the time period, indicating that researchers are still searching to improve upon their methods and trying to derive the best possible ways of employing the contingent valuation method in health care resource allocation decisions. Johannesson cautions that: "Much more methodological work is needed before the results of the CVM method can be used with any confidence in decisions about the allocation of health care resources" [27], and our review attests that many methodological questions continue to be addressed in the literature.

Noteworthy, is that while 37 of the papers retrieved were published in journals, most of these journals only published one CVM study in the past 12 years which may raise concerns regarding whether these journals have the switchable mechanisms enabling them to properly review such studies. Medical Decision Making published 4 papers in the past 10 years, which was the greatest number published amongst all of the journals. Medical care, International Journal of Technology Assessment in Health Care, and the Journal of Health Economics each published three papers. Two studies each appeared in Drug Intelligence and Clinical Pharmacy and Journal of Pediatrics. This leaves 22 papers that appeared in different journals. While the number of papers published by each journal type is small there appears to be a trend towards the majority of the papers being published in health or medical journals (62%). Given the economic basis of the analysis it is surprising that the field has not been covered to a greater degree in the economic journals.

3.4.2. The 'CBA' Papers

Distinguishing papers where the underlying motivation or context was CBA from pricing demand studies was difficult because of the limited clarity in reporting discussed earlier. Only two papers explicitly noted that their perspective was one of marketing. Reardon [54] (which appeared in The Journal of Health Care Marketing) examined market segmentation and market research while Eastaugh [14] explicitly states that his perspective was industrial marketing in the abstract.

Although Pennie [52] conducted a CVM study in order to ascertain a demand curve, the perspective was one of program evaluation and this study was included in the CBA subset. Pennie examined the use of hepatitis B vaccine by students in health care disciplines. It was clear that the authors wanted all students to be able to have access to the program, but since the public system would not pay for it, there would have to be user fees. Hence, in the case of a public program with a private component, a study may wish to evaluate a demand curve in order to ascertain when the number of people requesting the program which can then be subsidized, in order to meet those needs.

Although in the majority of studies the program being valued was already in existence, only one paper examined the removal of a program [37] using a WTA question and implying that a compensating variation was being measured. Examining the money measurement employed (CV/EV) and the direction of this measurement of the utility change employed (WTA/WTP), the studies performed rather poorly with respect to explicitly defining what exactly was being measured. When applying such economic analyses, authors' should be clear as to what methods are being used and what exactly is being measured. While it may be possible to infer the money measure it should be made clear and stated explicitly in the study. In this way the reader of such a study can be assured that the study correctly measured what the authors intended to measure.

A plausible explanation for the finding that no papers employed EV/WTA is that asking someone about how much compensation would be required to forgo an increase in future utility may be cognitively challenging, as conceptually, this requires a subject to think about a situation that has not yet occurred. Such cognitive difficulties with WTA were found in studies by O'Brien and Viramontes [48] and in a recent study by Donaldson *et al.* (1997) [11]. As discussed in O'Brien and Gafni [45] there are concerns that WTA valuations may be higher than WTP due to the fact that WTP values are bounded by ones income while WTA is not. Given that WTA can be measured in the context of CV or EV it is important that methods are clearly defined so that readers can be clear as to what exactly is being valued.

With respect to studies that explicitly stated health status outcomes, studies that employed the user-based approach may sometimes be more appropriate than the ex-ante insurance approach because patients are likely to have greater knowledge about the health consequences of a treatment [28] and hence the cognitive burden of the task may be lower than for non-patients. Furthermore, the two approaches can be linked: Johannesson showed that the user-based approach (what he terms expected WTP) will be a lower bound for ex-ante insurance-based willingness-to-pay under conditions of risk aversion with respect to income and if marginal utility of income does not vary with health status [28] .

Our review indicates that 55% of contingent valuation studies in health care surveyed persons currently diseased or undergoing treatment. If the purpose of the contingent valuation survey is to assess WTP to be used in a societal resource allocation decision (via CBA) it is more appropriate to ask the question of a representative sample of the general population (i.e., those who finance the health care through taxes or insurance) including both non-diseased and diseased persons. The latter group would be represented in the sample proportionate to the prevalence of the disease or service utilization in question. Such sampling recognizes that the goal of the WTP/WTA exercise is to quantify (in money terms) all the gains and losses in utility arising from the program change. Hence if CBA operates on the sum of the compensating variations, restricting the sampling to diseased persons will under-estimate the trial value because it excludes option values (those at future risk of the disease) and externalities (those who are not at risk but value consumption by others).

3.5. Guidelines and NOAA

In 1993 the National Oceanic and Atmospheric Administration (NOAA) published a set of general guidelines on the design and implementation of contingent valuation studies [43] for environmental damage assessment in the context of CBA. Although a detailed appraisal of the NOAA guidelines is beyond the scope of this study, it is instructive to briefly review whether these guidelines might be applicable to health care CVM studies, or whether the 'special' characteristics of health and health care call for modification of this guidance.

The NOAA report recommended employing WTP questions as opposed to WTA questions in order to elicit valuations, as the former would produce a more "conservative" estimate. A consistent empirical finding (in health and elsewhere) is that WTA exceeds WTP [40]. There are many competing hypotheses as to why this is observed [45] but the main reason is simply that WTP valuations are bounded by people's budget constraints while WTA are not. While the NOAA guidance not to use WTA may yield "conservative" estimates we have argued in this review that a comprehensive CVM needs to capture utility gains and losses from consumers and some assessment of WTA may be unavoidable. Furthermore, in circumstances where the primary policy under evaluation is the removal of a health care program, a WTA frame may be more appropriate [45].

With respect to the valuation elicitation method the NOAA report recommended the take-it-or-leave-it approach. It is clear that this method is gaining popularity in the health care field. After not being used at all in the first half of the time-frame examined, it become the most popular choice in the latter half. It is difficult to assess how much of this trend may have been attributed to the NOAA report. The difficulties with this approach include it's sample size requirements and computational complexity [5].

The NOAA panel "believes it unlikely that reliable estimates of values could be elicited with mail surveys". Hence, they recommended that in-person interviews be employed. However, given that take-it-or-leave-it type questions have been recommended by the panel, it will be difficult and costly to employ both of these recommendations together. The potential disadvantage of other elicitation formats such as "bidding games" is the threat of a framing effect known as starting point bias. The evidence on starting point bias in health care WTP studies is equivocal – some studies have confirmed it's presence [57] while others have not [46]. We believe that an important future direction for health care CVM studies is through the use of computer-based interviewing where bidding games and other elicitation methods can be used with randomized starting bids and with full multi-media presentation of valuation scenarios.

An important recommendation from the NOAA panel was that CVM studies should include tests of construct validity. The NOAA panel refers to these as 'scope tests' where the analyst needs to show an association with the scope of the environmental damage (or benefit) and the respondents' WTP. Although we did not formally review all construct validation tests in health care CVM studies, we found that 50% of studies reviewed did examine the association between WTP and income, for example, and more recent studies are examining the relationship between expected health benefit or prior risk of disease and WTP as scope tests [36 ,46]. We believe such validation is an important item for the research agenda of health care CVM studies; better yet, would be data from observed market choices ('actual WTP') against which to validate hypothetical statements. (For more on NOAA guidelines and health see Johannesson [29]).

While it appears that the NOAA report may not have had a major impact on contingent valuation studies employed in health care evaluations, it must be remembered that these guidelines did not appear until 1993. Given the large time lags from the development of a survey to a published paper it may still take some time for the impacts of the report to be noted. However, with clear references to the NOAA report in some recent papers [4,56], and attempts by the same to incorporate their recommendations it is clear that the NOAA report is beginning to have an impact outside of the environmental field.

3.6. Where do we go from here?

In summary, one of the major findings of this review is the lack of consistent and transparent reporting methods employed in many of the studies. If the rationale for a contingent valuation study is not made explicit, it is difficult for the reader to assess the validity of the data reported. As is attested by the number of methodological issues still being explored the above statement is even more significant. As methods are still being debated in the literature, it is key that reporting of the methods needs to be more clear and transparent. On another similar note, if one wants to compare results across studies, without explicit reporting this task becomes more difficult.

At a minimum any study should explain and defend the methods employed. Given the number of decisions needed to be made when conducting a CVM study, from the type of question asked to the choice of the population to sample it is important that authors be clear as to why particular methods were chosen. An important area for future research is the ongoing validation of data from CVM studies, either in terms of standardized approaches to construct validation (i.e., scope tests) by associations with income and other factors, or by comparison with actual consumer choices rather than hypothetical survey scenarios.

Table 3.1. Questions and considerations for a contingent valuation study of a health care program*

	QUESTION		CONSIDERATIONS
1	What question do we want to answer	A	Problem definition: 1. Pricing and demand studies 2.Project appraisal for resource allocation
		B	Current status of program 1.Program currently exists 2.Program does not currently exist
		C	Utility and disutility of program to respondent 1.Gain in utility from program 2.Loss in utility from program
2	What type of measure can we use?	D	Money measure of utility change: 1.Compensating variation (CV) 2.Equivalent variation (EV)
		E	'Direction' of measurement: 1.Willingness-to-pay (WTP) 2.Willingness-to-accept (WTA)
3	What do we need to ask of whom?	F	Externality and option value 1. Currently diseased 2. Currently non-diseased; at future risk 3. Currently non-diseased; not at future risk
		G	Framing of program consumption and payment: 1. Ex-post user-based question 2. Ex-ante insurance-based question
4	What characteristics of the program are important for determining how it is valued?	H	Program Outcome description: 1. Certain outcomes 2. Uncertain outcomes
		I	Nature of the 'market' for valuation scenario: 1. Private goods market 2. Political market
5	What question formats minimize bias and increase precision?	J	Valuation scenario 1. Holistic versus decomposed scenarios 2. Realism of scenarios
		K	Value elicitation method (e.g. 1-5 below): 1. Open-ended questions 2. Bidding Games 3. Payment Cards 4. Take-it-or-leave-it 5. Take-it-or-leave-it (with follow-up)

*Adapted from O'Brien and Gafni[15]

Table 3.2. Characteristics of all health care CVM papers retrieved (n=48) by time period of publication

	Number of studies (column %):		
	1984-89	1990-96	Total
Number of Papers	10 (100)	38 (100)	48 (100)
Type of Empirical Study:			
Methodology	3 (30)	16 (42)	19 (40)
Estimation	7 (70)	22 (58)	29 (60)
Objective of Study:			
Pricing and demand studies	3 (30)	3 (8)	6 (13)
Program evaluation (CBA)	7 (70)	35 (92)	42 (88)
Journal Type*:			
Health Economics	0 (0)	6 (16)	6 (13)
Economics (non-health)	1 (11)	1 (3)	3 (6)
Clinical	4 (44)	9 (24)	13 (27)
Health Services Research	3 (22)	14 (39)	17 (35)
Other	2 (22)	6 (18)	9 (19)
Survey Method:			
In-person	5 (50)	13 (34)	18 (38)
Mail	5 (50)	20 (53)	25 (52)
Telephone	0 (0)	4 (11)	4 (8)
Can't Tell	0 (0)	1 (3)	1 (2)
Content Area of Study:			
Disease:			
Health Care Programs / Services	3 (30)	8 (22)	11(23)
Respiratory Disease	0 (0)	3 (6)	3 (6)
Cardiovascular Disease	0 (0)	9 (24)	9 (19)
CF screening	0 (0)	3 (6)	3 (6)
Other	7 (70)	15 (38)	22(45)
Intervention type:**			
medical	8 (80)	23 (60)	31 (64)
surgical	0 (0)	2 (5)	2 (4)
pharmaceutical	0 (0)	12 (32)	12 (25)
hypothetical	2 (20)	2 (5)	4 (8)

*See Table 3.3 for journals that are represented by each journal type

** some papers employed more than one intervention type

Table 3.3: Journals from which studies were retrieved

Clinical (13)

Acta Obstetrica et Gynaecologica Scandinavica
Archives of Family Medicine
British Medical Journal
Drug Intelligence & Clinical Pharmacy (2)
Journal of Clinical Pharmacy & Therapeutics
Journal of Hypertension
Journal of Internal Medicine
Journal of Medical Genetics
Journal of Medical screening
Journal of Pediatrics (2)
Leadership in Health Services

Health Services Research / Health Policy (14)

Canadian Journal of Public Health
Health Policy
International Journal of Technology Assessment in Health Care (3)
Journal of Health Care Marketing
Journal of Nursing Administration
Medical Care (3)
Medical Decision Making (4)

Health Economics (6)

Health Economics
Journal of Health Economics (4)
Pharmacoeconomics

Economic (non-health) (3)

Applied Economics (2)
International Journal of Social Economics

Table 3.4. Classification of CBA studies (n=42) according to criteria of O'Brien and Gafni*

Questions and Considerations		1984-89 (n=8)	1990-96 (n=34)	Total (n=42)§
		Number of studies (column %):		
What question do we want to answer				
Current status of program	Program currently exists	3 (38)	21 (62)	24 (61)
	Program: Introduction	5 (62)	12 (35)	17 (37)
	Program: Removal	0 (0)	1 (3)	1 (2)
Utility/disutility of program to respondent	Gain in utility from program	7 (88)	31 (91)	38 (90)
	.Loss in utility from program	0 (0)	1 (3)	1 (2)
	Gain or loss possible	1 (12)	2 (6)	3 (7)
What type of measure can we use?				
Money measure of utility change	Compensating variation (CV)			
	Explicit	0 (0)	6 (18)	6 (15)
	Inferred	8 (100)	26 (82)	34 (85)
	Equivalent variation (EV)			
	Explicit	0 (0)	1 (3)	1 (2)
	Inferred	0 (0)	2 (6)	2 (5)
Valuation method employed	Willingness-to-pay (WTP)	8 (100)	32 (94)	40 (95)
	Willingness-to-accept (WTA)	0 (0)	2 (6)	2 (5)
What do we need to ask of whom?				
Externality and option value	Currently diseased	6 (75)	16 (47)	22 (55)
	At future risk	2 (25)	14 (41)	16 (36)
	Not at future risk	0 (0)	0 (0)	0 (0)
	General population	0 0	4 (12)	4 (10)
Framing of program consumption and payment:	Ex-post user-based	3 (38)	8 (24)	11 (26)
	Ex-ante user based	5 (63)	21 (67)	26 (62)
	Ex-ante insurance-based	0 (0)	6 (15)	6 (14)
What characteristics of the program are important for determining how it is valued?				
Program Outcome description	Certain outcomes	7 (88)	25 (74)	32 (76)
	Uncertain outcomes	1 (12)	9 (26)	10 (14)
Nature of the 'market' for valuation	Private goods market	7 (88)	29 (85)	36 (86)
	Political market	1 (12)	5 (17)	6 (14)
What question formats minimize bias and increase precision?				
Valuation scenario	Holistic	8 (100)	30 (88)	38 (90)
	Decomposed	0 (0)	4 (12)	4 (10)
Value elicitation method	Open-ended questions	6 (75)	10 (29)	16 (38)
	Bidding Games	0 (0)	12 (35)	13 (29)
	Payment Cards	1 (13)	7 (21)	8 (19)
	Take-it-or-leave-it	0 (0)	12 (36)	11 (26)
	Other/can't tell	1 (13)	0 (0)	1 (2)

*O'Brien and Gafni, 1995

§Total may be greater than 42 as some studies employed more than one method (the percentages are based on totals of 8, 34 and 42)

Table 3.5. Valuation method and money measurement

		Number of studies (column %):		
		WTP	WTA	Total
Compensating Variation	explicit	6 (17%)	0 (0%)	6 (16%)
	inferred	37 (78%)	2 (100%)	39 (79%)
Equivalent Variation	explicit	1 (0%)	0 (0%)	1 (0%)
	inferred	2 (6%)	0 (0%)	2 (5%)
	Total	46	2	8

WTP = Willingness-to-pay; WTA = Willingness-to-accept
Note: the total adds up to 48 as some papers employed two methods.

References

1. Appel LJ: Risk reduction from low osmolality contrast media: What do patients think it is worth? *Medical Care* 1990;28:324-327.

2. Berwick DM, Weinstein MC: What do patients value? Willingness to pay for ultrasound in normal pregnancy. *Medical Care* 1985;23:881-893.

3. Birdsall N, Chuhan P: *Willingness to pay for health and water in rural Mali: Do WTP questions work?*, Population Health and Nutrition Dept., World Bank; 1987:

4. Blomquist GC, O'Conor RM: *Eliciting preference for safety and efficacy of asthma treatments from nonasthmatic individuals*, Kentucky, Department of Economics, University of Kentucky; 1996:

5. Boyle KJ, Bishop RC: Welfare measurements using contingent valuation; A comparison of techniques. *American Journal of Agricultural Economics* 1988;20-28.

6. Carson RT, Wright J, Carson N, Alberini A, Flores N: *A bibliography of contingent valuation studies and papers*, La Jolla, California, National Resource Damage Assessment Inc.; 1995:

7. Chestnut LG, Keller LR, Lambert WE, Rowe RD: Measuring heart patients' willingness to pay for changes in angina symptoms. *Medical Decision Making* 1996;16:65-77.

8. Culbertson VL, Arthur TG, Rhodes PJ, Rhodes RS: Consumer preferences for verbal and written medication information. *Drug Intelligence & Clinical Pharmacy* 1988;22:390-396.

9. Donaldson C: Willingness to pay for publicly-provided goods: A possible measure of benefit? *Journal of Health Economics* 1990;9:103-118.

10. Donaldson C, Hundley V, Mapp T: *Willingness to pay: A new method for measuring patients' preferences? HERU Discussion Paper 03/95*, Aberdeen, Health Economics Research Unit, University of Aberdeen; 1995:

11. Donaldson C, Mapp T, Farrar S, Walker A, Macphee S: Assessing community values in health care: Is the willingness to pay method feasible. *Health Care Analysis* 1997;5:7-29.

12. Donaldson C, Shackley P, Abdalla M, Miedzybrodzka Z: Willingness to pay for antenatal carrier screening for cystic fibrosis. *Health Economics* 1995;4:439-452.

13. Donaldson C, Thomas R, Torgerson DJ: Validity of open-ended versus payment scale approaches to eliciting willingness to pay. *Applied Economics* 1996;29:

14. Eastaugh SR: Valuation of the benefits of risk-free blood. Willingness to pay for hemoglobin solutions. [Review]. *International Journal of Technology Assessment in Health Care* 1991;7:51-57.

15. Eckerlund I, Johannesson M, Johansson P-O, Tambour M, Zethreaus N: Value for money? A contingent valuation study of the optimal size of the Swedish health care budget. *Health Policy* 1995;34:135-143.

16. Einarson TR, Bootman JL, McGhan WF, Larson LN, Gardner ME, Donohue M: Establishment and evaluation of a serum cholesterol monitoring service in a community pharmacy. *Drug Intelligence & Clinical Pharmacy* 1988;22:45-48.

17. Ellison P, Marr J: Willingness to pay for unfunded health services in a family practice clinic. *Leadership in Health Services* 1994;3:14-17.

18. Fisher M, Marks A, Trieller K, Brody R: Are adolescents able and willing to pay the fee for confidential health care? *Journal of pediatrics* 1985;107:480-483.

19. Gafni A: Willingness-to-pay as a measure of benefits. Relevant questions in the context of public decision making about health care programs. *Medical Care* 1991;29:1246-1252.

20. Gafni A, Feder A: Willingness to pay in an equitable society: The case of the kibbutz. *International Journal of Social Economics* 1987;14:16-21.

21. Golan EH, Shechter M: Contingent valuation of supplemental health care in Israel. *Medical Decision Making* 1993;13:302-310.

22. Gore PR, Madhavan s: Consumers' preference and willingness to pay for pharmacist counselling for non-prescription medicines. *Journal of Clinical Pharmacy & Therapeutics* 1994;19:17-25.

23. Granberg M, Wikland M, Nilsson L, Hamberger L: Couples' willingness to pay for IVF/ET. *Acta Obstetrica et Gynaecologica Scandinavica* 1995;74:17-25.

24. Grimes D: Value of a negative cervical smear. *BMJ* 1988;296:1636-1636.

25. Hill JW: *Valuing reductions in risk of breast cancer mortality: A comparison of revealed preference and contingent valuation methods. PhD Dissertation*, Madison, University of Wisconsin; 1988:

26. Johannesson M: Economic evaluation of hypertension treatment. *International Journal of Technology Assessment in Health Care* 1992;8:506-523.

27. Johannesson M: The contingent valuation method -- appraising the appraisers. *Health Economics* 1993;2:357-359.

28. Johannesson M: A note on the relationship between ex-ante and expected willingness to pay for health care. *Social Science and Medicine* 1996a;42:305-311.

29. Johannesson M: The contingent valuation controversy in environmental economics and its relevance to health services research. *J.Health Serv.Res.Policy* 1996b;1:116-117.

30. Johannesson M, Aberg H, Agreus L, Borgquist L, Jonsson B: Cost-benefit analysis of non-pharmacological treatment of hypertension. *Journal of Internal Medicine* 1991;230:307-312.

31. Johannesson M, Fagerberg B: A health economic comparison of diet and drug treatment in obese men with mild hypertension. *Journal of Hypertension* 1992;10:1063-1070.

32. Johannesson M, Johansson P-O: To be or not to be, that is the question -- an empirical study of the WTP for an increased life expectancy at an advanced age. *Journal of Risk and Uncertainty* 1997;Forthcoming:

33. Johannesson M, Johansson P-O, Kristrom B, Borgquist L, Jonsson B: Willingness to pay for lipid lowering: A health production function approach. *Applied Economics* 1993;25:1023-1031.

34. Johannesson M, Johansson P-O, Kristrom B, Gerdtham U-G: Willingness to pay for antihypertensive therapy: Further results. *Journal of Health Economics* 1993;12:95-108.

35. Johannesson M, Jonsson B: Willingness to pay for antihypertensive therapy: Results of a Swedish pilot study. *Journal of Health Economics* 1991;10:461-474.

36. Kartman B, Andersson F, Johannesson M: Willingness to pay for reductions in angina pectoris attacks. *Medical Decision Making* 1996;16:248-253.

37. Lindholm L, Rosen M, Hellsten G: Are people willing to pay for a community-based preventive program. *International Journal of Technology Assessment in Health Care* 1994;10:317-324.

38. Miedzybrodzka Z, Semper J, Shackley P, Abdalla M, Donaldson C: Stepwise or couple antenatal carrier screening for cystic fibrosis? Women's preferences and willingness to pay. *Journal of Medical Genetics* 1995;32:282-283.

39. Miedzybrodzka Z, Shackley P, Donaldson C, Abdalla M: Counting the benefits of a screening: A pilot study of willingness to pay for cystic fibrosis carrier screening. *Journal of Medical Screening* 1994;1:82-83.

40. Mitchell RC, Carson RT: *Using surveys to value public goods: The contingent valuation method*, Washington, D.C., Resources for the Future; 1989:

41. Morrison GC, Gyldmark M: Appraising the use of contingent valuation. *Health Economics* 1992;1:233-243.

42. Nathan RG, Bont GM, Minz RB: Patient interest in receiving audiotapes of information presented by their physicians. A survey of patients awaiting treatment in university and private practice settings. *Archives of Family Medicine* 1994;3:509-513.

43. National Oceanic and Atmospheric Administration: Report of the NOAA panel on contingent valuation. *Federal Register* 1993;58:4607-4614.

44. Neumann PJ, Johannesson M: The willingness to pay for In Vitro Fertilization: A pilot study using contingent valuation. *Medical Care* 1994;32:686-699.

45. O'Brien B, Gafni A: When do the dollars make sense? Toward a conceptual framework for contingent valuation studies in health care. *Medical Decision Making* 1996;16:288-299.

46. O'Brien B, Goeree R, Gafni A, Torrance GW, et al: Assessing the value of a new pharmaceutical: A feasibility study of contingent valuation in managed care. *Medical Care* 1997;forthcoming:

47. O'Brien B, Novosel S, Torrance G, Streiner D: Assessing the economic value of a new antidepressant. *PharmacoEconomics* 1995;8:34-45.

48. O'Brien B, Viramontes JL: Willingness to Pay: A valid and reliable measure of health state preference? *Medical Decision Making* 1993;14:289-297.

49. O'Conor RM, Bloomquist GC: *Measurement of consumer-patient preferences using a hybrid contingent valuation method* , Stockholm, Centre for Health Economics, School of Economics, Stockholm; 1996:

50. Olsen JA, Donaldson C: *Willingness to pay for public sector health care programmes in northern Norway. HERU Discussion Paper 05/93*, Aberdeen, Health Economics Research Unit, University of Aberdeen; 1993:

51. Osmond MH, Klassen TP, Quinn JV: Economic comparison of a tissue adhesive and suturing in the repair of pediatric facial lacerations. *Journal of pediatrics* 1991;230:307-312.

52. Pennie RA, O'Connor AM, Garvock MJ, Drake ER: Factors influencing the acceptance of hepatitis B vaccine by students in health disciplines in Ottawa. *Canadian Journal of Public Health* 1991;82:12-15.

53. Phillips KA, Homan RK, Luft HS, et al: Willingness to pay for poison control centers. *Journal of Health Economics* 1997;16:343-357.

54. Reardon G, Pathak DS: Segmenting the antihistamine market: An investigation of consumer preferences. *Journal of Health Care Marketing* 1990;10:23-33.

55. Reis J, Sherman S, Macon J, Friedman B: Care for the underinsured: Who should pay? *Journal of Nursing Administration* 1990;20:16-20.

56. Ryan M, Ratcliffe J, Tucker J: Using willingness to pay to value alternative models of antenatal care. *Social Science and Medicine* 1997;44:371-380.

57. Stalhammar N-O: An empirical note on willingness to pay and starting point bias. *Medical Decision Making* 1996;16:247-247.

58. Thompson MS: Willingness to pay and accept risks to cure chronic disease. *American Journal of Public Health* 1986;76:392-396.

59. Thompson MS, Read JL, Liang M: Feasibility of willingness-to-pay measurement in chronic arthritis. *Medical Decision Making* 1984;4:195-215.

Chapter 4

CONJOINT ANALYSIS IN THE VALUATION OF HEALTH AND ENVIRONMENTAL ATTRIBUTES:

4.1. Introduction

One of the disadvantages of the contingent valuation method is its inability to satisfactorily handle programs with multiple outcomes. This was reinforced by the findings of the comprehensive literature review in Chapter 3. The measurement of multi-attribute outcomes was not addressed in that CVM literature. Clearly, one of the benefits of conjoint analysis is its ability to incorporate several levels of multiple attributes in one analysis. This allows for a richer description of attribute trade-offs that individuals are willing to make. This is particularly important in cases where different programs may affect different attributes. In deciding how to allocate resources, knowing what the trade-offs are between attributes can allow for a more efficient allocation. Hence conjoint analysis is an important route to explore in the valuation of the attributes of health and the environment.

This chapter provides a review of conjoint analysis including a brief history, the steps involved in conducting a CA study, and a survey of recent applications of the technique. The next section provides a background to the methodology including a comparison of the CA and CVM techniques and under what circumstances CA may be more applicable. The chapter then proceeds to review the specific application of the methodology that was employed in the case study - namely the multinomial logit random utility model. Following that is a presentation of the steps involved in undertaking such a CA study, with an emphasis on the discrete choice and ranking method that was employed in the case study. The methods that were proposed by Chapman and Staelin (1982) that allow extra information to be used if rankings are obtained are then presented. The chapter concludes with a review of some of the CA studies that have been conducted in the field of environmental and health economics.

4.2. Background

> The term 'conjoint analysis' defies a simple definition.
>
> Unfortunately, the term "conjoint analysis" coined by Green and Srinivasan (1978) was implicitly defined to include any technique used to estimate attribute utilities based on subjects' responses to combinations of multiple decision attributes. This suggests that there is a general technique called conjoint analysis that one uses when one wants to model consumer decision making and develop measures of consumers' utilities. However, different "conjoint" paradigms have different assumptions, methods of analysis, and experimental procedures (Louviere, 1988A).

Conjoint analysis methodology originated in the market research field where it has been used to establish what factors influence the demand for different commodities and what combinations of attributes would maximize sales (Cattin and Wittink, 1982). The method has also been used extensively in the field of transport economics in the areas of preference evaluation, demand analysis, and forecasting (Kroes and Sheldon, 1988). Generally, in a conjoint analysis survey, respondents are asked to rank or rate a group of alternatives each defined by various levels of several attributes. If a payment vehicle is included as one of the attributes, willingness-to-pay can be estimated. WTP applications of the techniques include estimating the monetary value of health care programs such as time spent on NHS waiting lists in the U.K. (Propper, 1991; 1995), miscarriage management (Ryan and Hughes, 1997) and environmental services such as wilderness recreation (Adamowicz *et al.*, 1994), and toxic dump sites (Schulze *et al.*, 1995).

The significant difference between conjoint analysis and contingent valuation is that in a conjoint analysis study one is able to incorporate several levels of multiple attributes. While a CVM study can examine several attributes they must either be separated with a different question employed for each attribute and for each level of the attribute. Alternatively all of the attributes must be included together in one question with all of the attributes changing together. In CA study techniques allow for the trade-offs between goods to be estimated. As such conjoint analysis may be more appropriate in the case where the good to be valued is composed of several distinct attributes.

There is a parallel between the conjoint analysis methodology and multi-attribute utility theory (MAUT). Multi-attribute utility theory is a method of cost-utility analysis in which utility scores for various health states are determined (Drummond *et al.*, 1997; Torrance *et al.*, 1996). The method has been used to develop pre-scored multi-attribute health status classification systems such as the

Quality of Well Being (QWL), Health Utilities Index (HUI) and EuroQol (EQ-5D). The method is based on multi-attribute preference theory (Keeney and Raiffa, 1993) in which an equation is determined that expresses the overall utility as a function of the single attribute scores.

The particular conjoint paradigm that this study shall focus on is the choice theoretic model developed by McFadden which combined the evaluation of alternatives and random utility maximization (McFadden, 1974; 1986). In a random utility model, utility of an available alternative is a linear function of the attributes of the alternative and the characteristics of the individual. The random utility model that is employed in this study is the multinomial logit model (MNL), or more specifically McFadden's conditional logit model (McFadden, 1974).

4.3. The random utility model

The random utility model assumes that an individual's utility from any given alternative is specified as a linear function of the characteristics of the individual and the attributes of the alternative plus an error term. Basically each individual i (i= 1,...,n) faces a choice set C_i which consists of J alternatives (j=1,...,J). Each choice set is characterized by various levels, M, of each attribute A. The utility derived by individual i from choice j can be expressed as

$$U_{ij} = \mu_{ij} + e_{ij} \quad (3)$$

where μ_{ij} is the observable, or systematic, component and e_{ij} is the unobserved, or random, component of total utility. While both terms are known to the individual, the e_{ij}s are unobserved by the researcher, and are thus considered random variables. Facing several alternatives within a choice set the individual will choose the alternative that yields the greatest utility. The probability of individual i choosing alternative j is

$$P_i(j) = P(U_{ij} > U_{is}) \quad \forall s \in C_i,\ j \neq s \tag{4}$$

where C_i is the specific choice set facing the individual. The systematic component of utility, μ_{ij}, is assumed to have the following linear form.

$$\mu_{ij} = \beta' X_{ij} = \beta_1 x_{ij1} + \beta_2 x_{ij2} + \beta_3 x_{ij3} + \cdots + \beta_k x_{ijk} \tag{5}$$

The utility function, μ, can in essence be considered an individual's indirect utility function with the variables included in the X matrix being the specific attributes in question and the individuals income. If the error terms, e_{ij}, are independently and identically distributed with a Weibull distribution, it can be shown that (McFadden, 1974)

$$P_i(j) = \frac{e^{\beta' X_{ij}}}{\sum_{j=1}^{J} e^{\beta' X_{ij}}} \tag{6}$$

The above form of the random utility model is referred to as the multinomial logit model as it is the multiple choice generalization of the binary logit model. Maximum likelihood techniques can then be used to estimate the values of the parameters in the linear model specified above.

4.4. The steps of a CA study

4.4.1. Defining the attributes and their levels

The first step in any conjoint analysis involves establishing the attributes that are to be valued. This can be done by various methods including using focus groups, literature reviews, or even by a specific (pre-defined) policy question such as a government decision maker wishing to know the trade-off residents may be willing to make between increased taxes and the attributes of air quality (the study of this dissertation). Adamowicz *et al.* (1994), for example, included attributes such as types of fish available, fish size, types of camping facilities, existence of beaches and others in attempting to value recreational areas. On the other hand, Ryan and Hughes (1997) used attributes such as level of pain, time in hospital cost of treatment and complications following treatment in assessing preferences for miscarriage management. If the CA study is used to estimate willingness-to-pay values a monetary attribute must be included. This can include the cost of the program or a change in property or income taxes.

Once the attributes to be used have been defined, different levels must be assigned to them. It is important that these levels are plausible and capable of being traded off. As with defining the attributes, the levels can be ascertained from various sources and are dependent upon the particular research question.

4.4.2. Design of experiments

Once the attributes and the levels have been defined, scenarios that will be presented to the individuals must be developed. The scenarios, referred to as choice sets, are comprised of two or more alternatives which involve different levels of each attribute. Given that one may wish to study several attributes that may take on several levels, the total number of possible alternatives that can be included in the choice sets can become rather large. For instance, with 3 attributes and employing 2 levels of each, there are a total of 2^3, or 8, possible combinations. More generally the total number of possible alternatives is defined as

$$J = M^{A} \qquad (7)$$

where J is the total number of alternatives, A is the number of attributes, and M is the number of levels of each attribute. Clearly with larger numbers of attributes and levels, the total number of alternatives can quickly become quite large. In such cases it is infeasible to employ the full factorial design, and thus researchers often rely on some form of a partial factorial block design. In a fractional factorial design a subset of all the alternatives are employed. Researchers often choose to use an orthogonal design in which each level of one attribute occurs with each level of another attribute with equal or at least proportional frequencies. Louviere and Woodworth (1983) have suggested the use of orthogonal main effects fractional factorial designs for both practical and academic applications. Main effects designs include no interaction effects and may be easier to implement as they allow less alternatives to be used. Designs are available, however, that allow for interaction effects to be included (Petersen, 1985).

Once the alternatives to be used are obtained, they are often grouped together in what are referred to as blocks. This is done in order to simplify the choice process of the respondent as they only have to consider a few alternatives at a time. These "blocks" can then be used as the choice sets to be presented to the respondents, either on their own, or supplemented by a "base" alternative. It has been recommended that a constant, or base, alternative be added to each choice set. This may provide a meaningful and important origin such as choosing the status quo (Louviere and Woodworth, 1983; Louviere, 1988B).

4.4.3. Elicitation of responses

Two methods that are available for eliciting responses which are consistent with the random utility model, MNL, framework are obtaining discrete choices or rankings. In the discrete method, respondents are asked to state their most preferred choice from each of the alternatives within a choice set. Alternatively, respondents could be asked to rank the alternatives within each choice set (Chapman and Staelin , 1982; Chapman , 1984; Hensher and Louviere, 1983). The ranking method requires the following assumptions of the individual's rankings: (a) the multinominal logit choice model (MNL) is a good approximation to the unobserved choices implied by the rankings; (b) the individual is perfectly transitive in the unobserved choice sets implied by the rankings; and (c) the individual is perfectly consistent in his/her ranking behaviour in the unobserved choice sets implied in the rankings.

4.5. "Exploding" the data

Chapman and Staelin (1982) propose a method in which models are estimated from simulated choices based on the "explosion" of the rankings which can enhance the estimation of parameters of the random utility model. Namely, if we get a complete rank ordering of each choice set (as opposed to just first choices) we can exploit this information in the previous setting. That is, we can relate ranking behaviour to choice behaviour. Using Luce and Suppes (1965) Ranking Choice Theorem, it can be shown that

$$Pr(a,b,c,...) = Pr(a \mid C) \cdot Pr(b,c,.....) \qquad (8)$$

where Pr(a,b,c,...) is the prob of observing the rank order of Alternative *a* being preferred to alternative *b* being preferred to alternative *c* and so on, in which a,b,c,... are the alternatives available in the choice set C. Pr (a|C) is the probability of Alternative *a* being chosen from the choice set C. The theorem enables the prob of a ranking event Pr(a,b,c,....) to be decomposed into two probabilities - the prob of a choice event and the prob of a subranking event. Thus, by successively applying the theorem to the subranking events, one can derive the probability expression for the ranking event which is the product of J-1 choice events.

$$Pr(a,b,c,...)=Pr(a|C)\cdot Pr(b|C-\{a\})\cdot Pr(c|C-\{a,b\})......... \quad (9)$$

Where C-{a} is the set of alternatives excluding Alternative A. Equation (9) states that the probability of the joint ranking event of J alternatives is composed of J-1 statistically independent choice events. Applying the Ranking Choice Theorem to the random utility model, it follows that:

$$prob(U_{i1}\geq U_{i2}\geq \cdots \geq U_{iJ})=\prod_{j*=1}^{J} prob(U_{ij}.\geq U_{ij}\forall j=(j*,...,J)) \quad (10)$$

The LHS of equation (10) is the joint probability that alternative 1 is preferred to alternative 2, etc., up to alternative J for decision maker i. The RHS can be interpreted as the statistical definition of the independence of the events:

$$(U_{i1}\geq U_{ij},j=1,2,...J),(U_{i2}\geq U_{ij},j=2,3,...J).... \quad (11)$$

The statistical independence condition of the above equation leads to the notion of the "explosion" process. More formally the rank order explosion rule states (Chapman and Staelin, 1982 pg. 291) :

> Given rank ordered choice set data, the original rank ordered observation that decision maker i prefers the alternatives in the order 1,2,...J (i.e. that $U_{i1}\geq U_{i2}\geq .....U_{iJ}$) can be exploded (decomposed) into J-1 statistically independent choice observations of the form $(U_{i1}\geq U_{ij}, j=1,2,...J)$, $(U_{i2}\geq U_{ij}, j=2,3,...J)$....$(U_{iJ-1}\geq U_{ij}, j=1,2,...J)$

Hence, by obtaining ranking data, we have additional choice observations for analysis at the estimation stage. These observations will in general lead to estimates of the random utility model that have a lower sampling variance (if the model is true).

One of the advantages of this method is that since more data can be garnered from the same number of people, sampling costs can be reduced. Nevertheless, at the same time, it may be a more difficult task for the respondents, which they may not want, or be able to complete. In theory each ranking should be statistically independent, which is the equivalent of obtaining them from a set of independent decision makers. However, since the task is difficult, people may give less thought to choices associated with lower ranked alternatives (as opposed to only their first ones) which may lead to larger variances for the lower ranked alternatives. It must be noted, however, that this latter view is not consistent with the underlying assumptions that lead to the development of the explosion rule, as we assumed that the errors were IID in the derivation of the McFadden probability ratio. Also, theoretically, Luce and Suppes Ranking Choice Theorem assumes that the respondents will rank the alternatives from best to worst in that order. Clearly, more research is required to ascertain whether respondents do give equal thought to all of the rankings. Ideally, the ranking exercise would be conducted in a lab setting. Once the respondent gives their most preferred choice we would remove that from the choice set and get them to then give their most preferred choice from the remaining alternatives, and so on. Chapman and Staelin's (1982) assessment of these data suggest that "...the statistical worth of more than the first couple of rank ordered choice set alternatives is marginal at best (because of reliability and noisy data problems)..."

4.6. A note on multiple observations

When conducting a choice-based CA study, not only is the design of the choice sets an important procedure, so is deciding how many of the choice sets each respondent will face[1]. It is often the case that each respondent may be presented with several choice sets that they are asked to consider. However, repeated measurements from one respondent may result in estimation difficulties. For instance, one assumption is that the error terms are independent, both *within* an individual respondent across the set of alternatives, and *between* respondents. It has been suggested that the result of ignoring these "repeated measurements" is likely to lead to an under-estimation of the standard errors of the parameters. Louviere and Woodworth (1983) recommend that the estimated variance matrix should be multiplied by the number of choice sets. The reasoning behind this conservative approach is that the choice sets do not constitute separate pieces of information, and should therefore not contribute to the calculation of the degrees of freedom.

> However, a more rigorous justification would seem to imply that the error terms pertaining to the (choice sets) for any given individual are completely correlated, and this does not seem at all likely. It seems much more likely that the variance of the error term differs between individuals....The general principles of estimation suggest that the consequence will be a decrease in the efficiency of the estimation of the coefficients, but without leading to bias - this provides some consolation in relation to estimation. (Bates, 1988)

---

[1]The number of choice sets may be dependant on the fractional factorial design employed.

Bates also notes that Louviere and Woodworth's suggestion is likely an over-correction. As well, the consequences of this source of heteroscedasticity do not cause problems with respect to estimation, but it does create problems when the goal of conjoint analysis is one of forecasting. When the aim of the study is to estimate marginal rates of substitution, as is the interest of the current study, explicit specification of the errors may be unimportant. There are some techniques available that can take into account multiple observations from individuals. These include Goldstein's multilevel modelling technique (Goldstein, 1995) and Hsaoi's random-effects panel data models (Hsaoi, 1986).

4.7. CA applications in health and the environment

While conjoint analysis techniques have been used a number of times the technique has rarely been used to explicitly estimate WTP. According to Layton (1995), the first contingent ranking application was Beggs *et al.* (1981). That study ranked 16 different hypothetical electric car models which differed over 9 attributes, including price. In this section some of the applications of the method to environmental and health care services are reviewed.

The first application of the contingent ranking technique to value an environmental good was Rae (1983). In that study visibility at Mesa Verde and Great Smokey National Parks was valued. The attributes included were visibility, waiting times for a tour, visitor centre activities, and park admission fee. While this study did not explicitly estimate WTP, they did find that price had a significant, negative effect on utility.

Smith & Desvouges (1986) considered water recreation on the Monongahela river. That study included both a contingent valuation and a contingent ranking component. In the contingent ranking component, only two attributes were included: the level of water quality described according to safe for drinking, swimming, boating, fishing, and an annual payment. Each respondent was asked to rank the same four alternatives, with higher levels of water quality associated with higher annual payments. Smith & Desvouges found that both the payment and water quality measure were statistically significant and properly signed using various model specifications.

Lareau and Rae (1989) estimated the WTP of reduced exposure to diesel vehicle odours. Attributes included were weekly occurrences of two types of odour and control costs paid by the households on a yearly basis. Each respondent was asked to rank one choice set which included four alternatives. They estimated WTP for a reduction of one odour exposure per week was \$5.86/year and for odour B \$15.53/yr. Two specification tests were conducted to examine whether the results were stable: (i) does the unconstrained set of specifications based on ranking each alternative separately differ from the constrained model specification that uses rankings on all four choices? and (ii) does the model estimated from using only the first alternative differ from the model estimated using all four ranked choices? A chi-square test was used, where the likelihood ratios were compared. The null hypothesis was rejected both times - i.e. there was no difference between the models.

Opaluch *et al.* (1993) ranked potential noxious facility sites, in Rhode Island, U.S. The authors were interested in evaluating the social trade-offs associated with the siting of such facilities. A paired comparison survey was conducted, in which respondents chose between two hypothetical landfill sites described in terms of attributes such as quality of underlying groundwater, wildlife habitat, number of houses in the vicinity, and acreage of woodlands, wetlands, and farmland. Each respondent was asked to rank eleven paired comparisons. While they did not explicitly measure WTP, cost per household was included as an attribute.

Adamowicz *et al.* (1994) combined revealed preference data in conjunction with stated preference data in order to value recreational site choices. Both models were based on random utility theory which allowed the joint estimation. Attributes included driving distance, fishing amenities, camping facilities, water quality, presence of boating, beach and swimming facilities, and type of terrain. Each respondent was asked to state their most preferred of three alternatives, in a total of sixteen choice sets. The main contribution is that they show that the underlying preferences reflected in the revealed preference and stated preference models are similar. This lends support to the use of this stated preference technique, at least in the measurement of values.

Schulze *et al.* (1995) estimated WTP for alternative cleanup options associated with hazardous waste sites, known as Superfund sites, in the U.S. The attributes employed included unpleasant odour, health risk (deaths per 10 million), and cleanup rating. This is the only study that was found that included both environmental and health attributes. Each respondent was asked to rank five alternatives, ranging from a complete cleanup to no action at all. Different survey versions were used in order to include different price combinations. The authors then estimated WTP using multinomial logit techniques utilizing the ranking data as proposed by Chapman and Staelin (1982).

Magat, Viscusi, and Huber (1988) conducted a paired-comparison survey in order to value morbidity risk reductions associated with making products safer. Specifically, they examined risk reductions from the use of safer chemical products including bleach and drain openers. Respondents were required to rate their preferences between the current product and a new one described by the cost per year and the injury level using a computer program. Viscusi, Magat, and Huber (1991) used a paired-comparison approach to measure marginal rates of substitution between chronic bronchitis risk reduction and both risk of automobile accidents and dollars. The former was referred to as a risk-risk trade-off and the latter as risk-dollar trade-off. The methodology of this study was similar to that of the authors' previous study, including the use of an iterative computer program to present the choice sets to the respondents.

Propper (1991 and 1995) estimated WTA for increased time on NHS (National Health Service, UK) waiting lists. Each respondent faced fourteen pairwise comparisons. The attributes included were waiting time, money cost, and uncertainty of date of admission. The study incorporated a Probit estimation model with WTP calculated as the ratio of the cost and waiting time coefficients (marginal rate of substitution). The estimated value of waiting list time was on average 38 pounds per month for a unit reduction in waiting list time (in months). Propper also calculated results based on segmenting the sample and found that those with higher incomes expressed a higher value of waiting time as did full-time employed versus part-time employed and housewives.

Ryan and Hughes (1997) used a conjoint analysis study in order to assess preferences for the management of miscarriages amongst women. The attributes included in the study were the level of pain, time spent in hospital, complications of treatment, length of time required to return to normal household activities, and cost of the treatment. Each individual was required to rank 12 or 13 pairwise comparisons each consisting of one surgical treatment alternative and one medical treatment alternative.

Desvouges *et al.* (1996) conducted a conjoint analysis study in order to estimate the value of the health effects of air pollution. The study focussed only upon health effects and did not include any of the other attributes of air quality. Both morbidity and mortality effects were employed. (The mortality effects are also reported in Johnson *et al.*, 1998. The study employed a modified version of the health state descriptions used in the Quality of well Being (QWB) index and used four attributes (symptom, number of episodes, daily activity level, and cost). Note that the source of the health conditions were not stated in the questionnaire employed. Hence the valuations were more generally for the actual health conditions than for a specific change in air quality. Respondents were asked to rate paired comparisons using a scale from 1-7 and an ordered probit was used to analyse the results. WTP results for the entire sample for all symptoms were not estimated as the authors' note that the WTP results were preliminary and illustrative. For the English-speaking sample mean WTP estimates ranged from $990 for decreases in chest or arm pain to %790 for cough, fever, and ache. It is unclear whether these vales are for a one-time payment or for a yearly payment. WTP for a year of additional longevity beyond the average expected longevity averaged about $14,000 (Canadian dollars).

Of note is that only one of these studies (Schulze *et al.* 1995) combined health and non-health related attributes. The majority of these studies used pairwise comparisons and discrete choice rather than rankings in the elicitation of responses.

4.8. Conclusions

The use of conjoint analysis methods to value public goods appears to be a promising stated preference approach. While the technique has had limited application in the field of value elicitation, this approach does have advantages over the more popular contingent valuation approach, particularly when the good is composed of several attributes. It allows the valuation of health care programs to include process outcomes, such as waiting time, along with the health outcomes. In the evaluation of environmental policies, such as those that reduce air pollution, it allows the inclusion of the discernible outcomes that may occur.

The feature of conjoint analysis allowing it to focus on multiple attributes is a significant one particularly when used in the application of examining the optimal allocation of public resources. For instance, in a collectively funded public health system with a finite number of resources, dollars spent on one program are no longer available to be spent on other programs or policies. Hence, it will be important to know what value members of the community attach to each program or each outcome concerned. Similarly there is a fixed budget with respect to the provision of environmental programs and hence the same principles apply. One of the recommendations of the NOAA Panel was that respondents be made to consider their budget constraint and must be reminded of substitute goods. This recommendation should be extended so that respondents are reminded about other goods and attributes from a societal point of view[2]. Obtaining values for substitute goods, or trade-offs, would be valuable when undertaking an economic evaluation in a collectively funded system. Conjoint analysis methods allow this to be undertaken and hence proves to be a valuable tool when examining multi-attribute outcomes.

The conjoint analysis methodology has both advantages and disadvantages. It can reduce the number of protest responses, which can occur when scenarios involve potential tax increases. As well, the additional information obtained by eliciting a complete ranking of choices enables us to obtain more efficient estimates of WTP. On the other hand, describing the scenarios convincingly requires considerable effort and it is not clear that the

---

[2]It must be noted that the NOAA Panel was not evaluating the use of CVM in resource allocation decisions in a collectively funded system. They were mandated to examine the applicability of CVM in the context of natural resource damage assessment. Hence it is not a shortcoming of their analysis that such a recommendation was not included.

average respondent is able to fully understand the implicit trade-offs. This question will be further addressed in Chapter 5.

The CA methodology constructs total value for an option based on marginal choices between options. This feature may address the issues of embedding and warm-glow (Kahneman and Knetsh, 1992). Respondents may feel some warm-glow by supporting the program in question, and thus may overstate their true value. For example, while each alternative, within a choice set, may yield warm-glow, the marginal value estimated between the options will likely net out any value derived from the warm-glow phenomenon (Schulze *et al.*, 1995).

Another advantage of conjoint analysis is that we may be able to obtain more information from fewer respondents, if a complete ranking of several alternatives is obtained, thus reducing the costs of obtaining such information. On the down side, however, requiring a complete ranking may be too cognitively difficult for respondents to yield valid and reliable results. Given its potential advantages, the conjoint analysis approach, as a method of estimating willingness-to-pay values, needs to be examined further, in order to assess its feasibility.

In the next chapter the methodology employed in a conjoint analysis study that was conducted in order to value the health and other attributes of air quality in the Regional Municipality of Hamilton-Wentworth is presented.

Chapter 5

WILLINGNESS-TO-PAY FOR HEALTH AND THE OTHER ATTRIBUTES OF AIR QUALITY IN HAMILTON-WENTWORTH: SURVEY DESIGN AND ADMINISTRATION

5.1. Introduction

Although air pollution is often viewed primarily as a health problem, other aspects also require attention. Poor air quality exacerbates cardiac and respiratory conditions and may increase hospitalization rates and mortality (Dockery and Pope, 1994; Burnett *et al.* 1995). There is evidence, however, that the so-called nuisance or aesthetic effects of poor air quality are also important to the general population. For example, 70% of respondents living in the industrialized north end of Hamilton reported a concern with black particulate deposition (black fallout) (Elliott *et al.*, 1997). Of that same sample, the most frequently cited effects of black fallout concerned "lifestyle disruption", including property damage to cars and furniture. Although almost half of those complaining about black fallout considered it a health problem, current research indicates that the particulates most dangerous to health are much smaller than those implicated in black fallout, so the damage done by this type of air pollution is unlikely to be reflected in medical statistics. Thus studies which infer the benefits of air pollution control from avoided health risk are likely to underestimate the true benefits and must be supplemented by some estimate of willingness-to-pay for other dimensions of improved air quality

In order to estimate resident's valuations of the attributes of air quality in Hamilton-Wentworth a conjoint analysis study was undertaken. The study was conducted in order to obtain information about the trade-offs and the willingness-to-pay for the attributes and to examine the nature and effect of cognitive difficulties of using conjoint analysis and to explore the use of ranking information.

Following the steps presented in Chapter 5 respondents were presented with well-defined alternative scenarios involving four attributes of air quality (health effects, bad odour, black fallout and poor visibility) and one payment vehicle (property taxes or rent) and asked to rank order the alternatives. The next section presents an introduction to the region of concern. The chapter proceeds to review the methods used in the survey, including the general design of the questionnaire, the derivation of the attribute levels, and sample selection. Section 5.7 presents the econometric model that was used to estimate willingness-to-pay.

5.2. Background

The Regional Municipality of Hamilton-Wentworth is located in southern Ontario on the southwest tip of Lake Ontario on a sheltered natural harbour. The Municipality is composed of The City of Hamilton, The City of Stoney Creek, the Township of Glanbrook, the Town of Ancaster, the Town of Dundas, and the Town of Flamborough, and has a total population of over 468,000 (1991 Census).

The Niagara escarpment (referred to as "the Mountain") surrounds the harbour and cuts through the Region creating geographically defined communities. The City of Hamilton can be divided into Lower and Upper Hamilton based on this geography with Upper Hamilton referring to the portion of the city above the escarpment and Lower Hamilton being below the escarpment. Figure 5.1 shows a map of the Region.

Two of the largest steel companies with associated secondary industries are located in the region with the bulk of industrial activity located in the northeast end of the Region (lower Hamilton). Hence, the north end of Hamilton is associated with the poorest air quality of the Region. However, local sources are not the only cause of air pollution in the Region.

According to the Summary Report of the Hamilton-Wentworth Air Quality Initiative (HAQI, 1997) there are three main air pollution problems in Hamilton. The first source is the industrial impact on nearby residential areas including black fallout and bad odours. The second problem pertains to short term episodes of pollution that build up throughout the city during temperature inversions. These episodes occur primarily in the spring and the fall and are the main source of poor visibility episodes in the area. The third problem is attributable to long range transport of pollutants from outside the region.

5.3. Questionnaire design

The questionnaire comprised three sections.[1] The first was intended to measure respondents' concerns about air quality relative to other social issues, to familiarize them with the air quality attributes being evaluated, and to elicit information about their past experiences with the attributes. It began by asking respondents to compare their air quality to that of the rest of southern Ontario, followed by four questions asking how concerned the individual was about each of the air quality attributes of odour, black fallout, poor visibility and health. Since these had not yet been defined or described, the responses should be unbiased by framing effects. Further questions described each attribute in detail and asked for the respondent's experience with that attribute. Featured in this section was a colour photograph of a hazy day. The section concluded with a pie chart illustrating the share of various services, such as education and policing, in the Regional property tax bill for a typical household (Figure 5.2). The second section, described in detail below, contained the scenarios used in the CA study and the final section elicited demographic information.

Two pilot studies were conducted. The first study consisted of mailing questionnaires to a random sample of thirty residents in the City of Hamilton. This study demonstrated that, despite the complexity of the choice sets, we might expect a response rate of about 30%. In the second pilot study a questionnaire was administered to a sample of 189 undergraduate economic students at McMaster

---

[1] A copy of the questionnaire is included in Appendix A.

University. This study demonstrated the feasibility of the computational technique employed to estimate the willingness-to-pay for improvements in the air quality attributes. The two pilot studies also led to a revision of the wording and placement of several questions in the survey.

5.4. The attributes

The air quality attributes selected for the study were (i) health effects, described as increased mortality and hospital admissions due to cardio-respiratory diseases, (ii) number of days of black fallout, (iii)number of days of bad odour, and (iv)number of days of poor visibility. The attributes were decided upon through research and consultation with epidemiologists and environmental consultants. Three levels of each attribute were employed. The middle level of each attribute corresponded to the status quo. The other two levels were a one-third reduction and a one-third increase in the current level of the attribute. Table 5.1, extracted from the questionnaire, shows how each of these levels was defined.

The health effects were described as hospital admissions for cardiac and respiratory diseases in the Region of Hamilton-Wentworth and increased mortality. The number of hospital admissions for cardio-respiratory diseases was calculated using the results of Burnett *et al.*'s (1995) study of the relationship between daily cardiac and respiratory admissions to 168 hospitals in Ontario with daily levels of particulate sulfates. They found that current levels of sulphate exposure led to increases of 2.5% and 3.5% in cardiac admissions (for those under and over 65, respectively) and 3.2% in respiratory admissions, compared with levels expected under zero exposure attributable to sulphate exposure. Since Burnett *et al.* only considered sulfates, the actual estimate for cardiac and respiratory hospital admission attributable to all air pollution would be larger. Table 5.2 indicates the diseases and ICD9[2] codes that were used by Burnett *et al.*

Applying Burnett *et al.*'s findings to annual data for Hamilton-Wentworth it was estimated that approximately 128 (2.5% of 5108) cardiac admissions and 77 (3.2% of 2407) respiratory admissions were attributable to sulfates for a total of 205 admissions per year (Dickson, 1995). For the purposes of the questionnaire this was rounded to 18 hospital admissions per month. Because Burnett *et al.* only considered exposures to sulphates, our estimate for hospital admissions from all types of air pollution is conservative.

As previously stated, many of the studies that have attempted to value the health effects of air pollution have focussed on personal effects only (e.g. Hall *et al.*, 1992). Given that people may, and do, care about the well-being of others in society it is important to frame questions in such a way that these externalities are accounted for. Thus, rather than estimate the willingness-to-pay for a reduction in personal health risk it may be more appropriate to frame questions from a societal perspective. For example, "How much would you be willing to pay to decrease society's health risk?". In other words, the health effects were actually *public* health effects. This allowed for externalities to be accounted for. The changes in the attributes were all specified in certain terms and in actual changes, rather than

---

[2]International Classification of Diseases, 9th revision

percentage reductions. This was done in order to avoid the added complexity of understanding how the respondents may have reacted to changes in risk (Schulze *et al.*, 1995).

The estimate of bad odour days was based on occurrences of total reduced sulphur (TRS) at the Beach Blvd. monitoring station. The five-year average of number of days with one or more hours over 10 ppb (49 days per year) was divided by 12 and rounded off to yield an estimate of 4 bad odour days per month under the current situation[3].

The estimate of poor visibility days was based on the five-year average of the number of days with one or more hours of the Air Pollution Index (API) being 25 or greater at the Elgin/Kelly monitoring station (14 days per year). Since these events generally only occur during the five months of April, May, June, September and October, it was expressed as 3 days per month. The months of July and August may also experience poor visibility days due to factors not captured in the API. It was estimated that these summer poor visibility days occur at the same rate as the spring and summer poor visibility days and for purposes of this analysis have grouped them together.

Black fallout (BFO) is not monitored in the Region and the only available proxy was a monthly dustfall measure. Since this estimate cannot be used to ascertain the number of days of occurrences the estimate of high BFO days was

---

[3]The level of 10 ppb of TRS and an API of 25 that were used as the criteria for bad odour days and poor visibility days were suggested by experts at the Ontario Ministry of the Environment.

based on complaints registered with the Ministry of Environment and Energy from the Beach strip area. The five year average of 36 days of complaints annually was expressed as 3 black fallout days per month.

Because they are frequently used to finance local services, property taxes was chosen as a payment vehicle. Since not all respondents would be homeowners, the payment vehicle was described as a change in monthly property taxes *or* rent. People were told that increased taxes may be necessary in order to reduce pollution. A pie chart (see Figure 5.2) was used to illustrate the average level and functional breakdown of the monthly tax payments on a typical property (City of Hamilton, 1996; Regional Municipality of Hamilton-Wentworth, 1996)[4].

5.5. The choice sets

In Part B of the questionnaire each respondent was presented with nine choice sets. Each choice set consisted of four alternative choices that were to be ranked by each respondent. Figure 5.3, extracted from the questionnaire, shows the Sample Choice Set used to explain the procedure. As in the actual choice sets, Choice A describes the status quo. The remaining choices were created by specifying differing levels for each of the five attributes. As noted, three levels were used (status quo, one-third better and one-third worse) for each of the air

---

[4]According to the City of Hamilton Current Estimates, 1996 and the Regional Municipality of Hamilton-Wentworth 1996 Budget the average household in the city paid approximately $2400 per year in property taxes. Of this amount, $557 went towards Regional services including $209 on police services, $137 on roads, and $9 on public health. $1018 was spent on education services. These amounts were then divided by 12 to obtain a monthly values.

quality attributes. A total of 13 levels of change in monthly taxes was used. (no change plus increases or decreases of $5, $10, $15, $20, $25, and $50). The range of values employed were partially based on the responses obtained in the pilot studies.

There are 3^4x13 or 1053 possible ways of combining the air quality attribute and tax levels to form choices. The set employed in the analysis was restricted in two ways. First, it was decided not to combine different increments in taxes within the same choice set. That is, the tax attribute within each choice set was restricted to three levels: an increase or decrease of the same magnitude or no change.[5] This left 3^5 or 243 choices from which to create choice sets. Following Petersen (1985) a one-ninth partial factorial design in 9 blocks of 3 was designed so as to permit estimation of all first and second order interactions of the factors (See Appendix B for more on the factorial design that was employed). These blocks of 3 were appended to Choice A (the status quo) to form the nine choice sets, of 4 alternatives each, that were used in the questionnaire.

Six versions of each choice set, one for each of the six possible increments in property tax ($5, $10, $15, $20, $25, and $50) were developed. Finally, six versions of the final questionnaire were created in such a way that every choice set was presented using each of the tax increment levels and every respondent faced each of the tax increment levels at least once. Table 5.3 reports the tax levels that were used in each version of the questionnaire. In all other respects the

[5]For example, Choices B, C, and D in Figure 5.2 all involve a $5 change in taxes or rent.

questionnaires were identical.

This procedure for generating choice sets produced some choices which were weakly dominated by others in the same set, in the sense that the level of at least one of its attributes was inferior to another choice and none were better. (It was assumed that improvements in air quality attributes and reductions in taxes would always be preferred). As a result, 9 of the 54 possible pairwise comparisons in the nine choice sets involve a dominance relationship[6]. These included Choice Set One (3 dominated alternatives), Choice Set 2 (2 dominated alternatives), Choice Set Three (1 dominated alternative) and Choice Set Nine (3 dominated alternatives). This was viewed as a virtue of the design, because it allowed a test for rationality of responses which is reported in the next chapter.

5.6. Sample selection

The survey was administered by mail because the complexity of the ranking exercises made it infeasible to conduct the survey by telephone and the cost of personal interviews would have been prohibitive. The final version, of the questionnaire, was sent to 1908 households within the Regional Municipality of Hamilton-Wentworth.[7] The sample was randomly chosen from the regional tax

---

[6]There are 4 alternatives in each of the nine choice set resulting in 54 possible pairwise comparisons.

[7]Questionnaire types A, C, D, and E were each sent to 333 residents. Questionnaire type B was sent to 243 residents, and questionnaire type F was sent to 323 residents. Due to a clerical error, the same number of each type of questionnaires were not sent out. There were enough of each type returned such that the results were not significantly affected.

assessment rolls as of December 1995. The survey packages, which included a copy of the questionnaire, a cover letter, and a prepaid return envelope, were mailed out on February 10, 1997 to the person who appeared on the tax assessment roll.[8] In order to encourage responses, recipients were informed that each respondent who returned a completed questionnaire by February 28, 1997 would be entered into a draw for one of ten $25 gift certificates to the local department store, or mall, of their choice. Following Dillman (1978), one week after the initial mailing, a follow-up postcard was sent reminding those who had not yet sent in a survey to please to so and thanking those who already had. Due to financial considerations further follow-ups were not undertaken.

5.7. The model and willingness-to-pay

As presented in Chapter 3, a random utility model is employed, and estimated using multinomial logit analysis. Four main models were estimated. The first model (Model 1) was a linear model where the variables employed were the five main attributes. Model 2 was the linear model used in Model 1 supplemented by a squared tax term. The latter term was included in order to allow taxes to enter the model with some curvature as it was assumed that larger tax changes might affect utility differently than would smaller changes. Model 3

---

[8]Note that by conducting the survey during the Winter months some of the attributes associated with air quality may not have been in the forefront of respondents minds. Environmental issues had not been the subject of recent new stories. A few months after the survey was conducted there was a fire at a local plastics factory which may have affected survey results if it had been conducted after that date.

was a full quadratic model. Dropping subscripts (that appeared in Chapter 5)for convenience, utility in the three models was defined as follows.

Model 1:

$$\mu = \beta_1 H + \beta_2 O + \beta_3 F + \beta_4 V + \beta_5 T \qquad (12)$$

Model 2:

$$\mu = \beta_1 H + \beta_2 O + \beta_3 F + \beta_4 V + \beta_5 T + \beta_6 (T^2) \qquad (13)$$

Model 3:

$$\begin{aligned}\mu =& \beta_1 H + \beta_2 O + \beta_3 F + \beta_4 V + \beta_5 T + \beta_6 (H^2) + \beta_7 (H \cdot O) + \beta_8 (H \cdot F) + \beta_9 (H \cdot V) \\ &+ \beta_{10}(H \cdot T) + \beta_{11} O^2 + \beta_{12}(O \cdot V) + \beta_{13}(O \cdot F) + \beta_{14}(O \cdot T) + \beta_{15} F^2 + \beta_{16}(F \cdot V) \\ &+ \beta_{17}(F \cdot T) + \beta_{18} V^2 + \beta_{19}(V \cdot T) + \beta_{20} T^2\end{aligned} \qquad (14)$$

where H = health risk
O = bad odour days
F = black fallout days
V = poor visibility days
T = property taxes or rent

Each of the air quality attributes was coded 0 for the initial situation, and +1 or -1 for one-third improvements or reductions respectively. Taxes were coded as deviations from the status quo, measured in dollars.

If WTP is calculated from the initial level of each attribute (H=O=F=V= T=0), and it is assumed that all of the other variables remain constant, the following WTP estimates are obtained by totally differentiating equation (12), equation (13) or equation (14)[9].

$$
\begin{aligned}
WTP_H &= dT/dH = -\beta_1/\beta_5 \\
WTP_O &= dT/dO = -\beta_2/\beta_5 \\
WTP_F &= dT/dF = -\beta_3/\beta_5 \\
WTP_V &= dT/dV = -\beta_4/\beta_5
\end{aligned}
\qquad (15)
$$

In addition to calculating willingness-to-pay values, the other trade-offs (HTO) between the attributes was calculated. Using health as the base, the health trade-off was calculated. This figure provides the value placed on the non-health attributes as a percentage of the value they place on the health attribute. Once again assuming that H=O=F=V= T=0, the HTOs are as follows[10].

$$
\begin{aligned}
HTO_O &= dH/dO = -\beta_2/\beta_1 \\
HTO_F &= dH/dF = -\beta_3/\beta_1 \\
HTO_V &= dH/dV = -\beta_4/\beta_1
\end{aligned}
\qquad (16)
$$

The fourth model (Model 4) which was estimated was a variation of the full quadratic model that also included individual specific, or demographic, characteristics of the respondents. These variables were incorporated by

[9]This value states the change in income that would be required in order to leave utility unchanged. This can be related to the compensating and equivalent variations derived in Chapter 2.

[10]The health trade-off for taxes is merely the inverse of the WTP for health, i.e. $HTO_T = dH/dT = (WTP_H)^{-1}$

interacting them with the attribute information[11]. Incorporating the respondent's age, employment status, income, and experience with cardio-respiratory diseases, results in the following form, which was estimated:

Model 4:

$$\begin{aligned}\mu =& \beta_1 H + \beta_2 O + \beta_3 F + \beta_4 V + \beta_5 T + \beta_6 (H^2) + \beta_7 (H \cdot O) + \beta_8 (H \cdot F) + \beta_9 (H \cdot V) \\ &+ \beta_{10}(H \cdot T) + \beta_{11} O^2 + \beta_{12}(O \cdot V) + \beta_{13}(O \cdot F) + \beta_{14}(O \cdot T) + \beta_{15} F^2 + \beta_{16}(F \cdot V) \\ &+ \beta_{17}(F \cdot T) + \beta_{18} V^2 + \beta_{19}(V \cdot T) + \beta_{20} T^2 + \beta_{21}(I \cdot H) + \beta_{22}(I \cdot 0) + \beta_{23}(I \cdot F) \\ &+ \beta_{24}(I \cdot V) + \beta_{25}(I \cdot T) + \beta_{26}(E \cdot H) + \beta_{27}(E \cdot O) + \beta_{28}(E \cdot F) + \beta_{29}(E \cdot V) + \beta_{30}(E \cdot T) \\ &+ \beta_{31}(A \cdot H) + \beta_{32}(A \cdot O) + \beta_{33}(A \cdot F) + \beta_{34}(A \cdot V) + \beta_{35}(A \cdot T) + \beta_{36}(C \cdot H) \\ &+ \beta_{37}(C \cdot O) + \beta_{38}(C \cdot F) + \beta_{39}(C \cdot V) + \beta_{40}(C \cdot T)\end{aligned} \tag{17}$$

Where I = income (in 1000's)
E = employment variable (1 if unemployed, zero otherwise)
A = age (in years)
C = experience with cardio-respiratory disease (1 if the respondent or a household member is currently diagnosed with a cardio-respiratory disease)

Estimating willingness-to-pay in the same way as in Model 1 the following is obtained:

[11]Note that any individual-specific characteristic cancel out of equation (4) unless it interacts with the attributes. Greene's (1993) suggestion of multiplying demographic variables by dummy variables for each choice within a set was impractical because separate dummies would have been required for each of the 27 (9x3) distinct alternatives to the status quo faced by each individual. Instead it was decided to interact the demographic variables with each level of the attribute.

$$
\begin{aligned}
WTP_H &= dT/dH = -(\beta_1+\beta_{21}I+\beta_{26}E+\beta_{31}A+\beta_{36}C)/(\beta_5+\beta_{25}I+\beta_{30}E+\beta_{35}A+\beta_{40}C)\\
WTP_O &= dT/dO = -(\beta_2+\beta_{22}I+\beta_{27}E+\beta_{32}A+\beta_{37}C)/(\beta_5+\beta_{25}I+\beta_{30}E+\beta_{35}A+\beta_{40}C)\\
WTP_F &= dT/dF = -(\beta_3+\beta_{23}I+\beta_{28}E+\beta_{33}A+\beta_{38}C)/(\beta_5+\beta_{25}I+\beta_{30}E+\beta_{35}A+\beta_{40}C)\\
WTP_V &= dT/dV = -(\beta_4+\beta_{24}I+\beta_{29}E+\beta_{34}A+\beta_{39}C)/(\beta_5+\beta_{25}I+\beta_{30}E+\beta_{35}A+\beta_{40}C)
\end{aligned} \quad (18)
$$

Note that now WTP varies with the demographic characteristics in this model.

The average value of each of these variables was used in the calculation of WTP.

The Health trade-offs for Model 4 were derived as follows.

$$
\begin{aligned}
HTO_O &= dT/dO = -(\beta_2+\beta_{22}I+\beta_{27}E+\beta_{32}A+\beta_{37}C)/(\beta_1+\beta_{21}I+\beta_{26}E+\beta_{31}A+\beta_{36}C)\\
HTO_F &= dT/dF = -(\beta_3+\beta_{23}I+\beta_{28}E+\beta_{33}A+\beta_{38}C)/(\beta_1+\beta_{21}I+\beta_{26}E+\beta_{31}A+\beta_{36}C)\\
HTO_V &= dT/dV = -(\beta_4+\beta_{24}I+\beta_{29}E+\beta_{34}A+\beta_{39}C)/(\beta_1+\beta_{21}I+\beta_{26}E+\beta_{31}A+\beta_{36}C)
\end{aligned} \quad (19)
$$

The next two chapters present the results from this survey, including the WTP and HTO estimates that were derived above from the three models.

Figure 5.1. The Regional Municipality of Hamilton-Wentworth

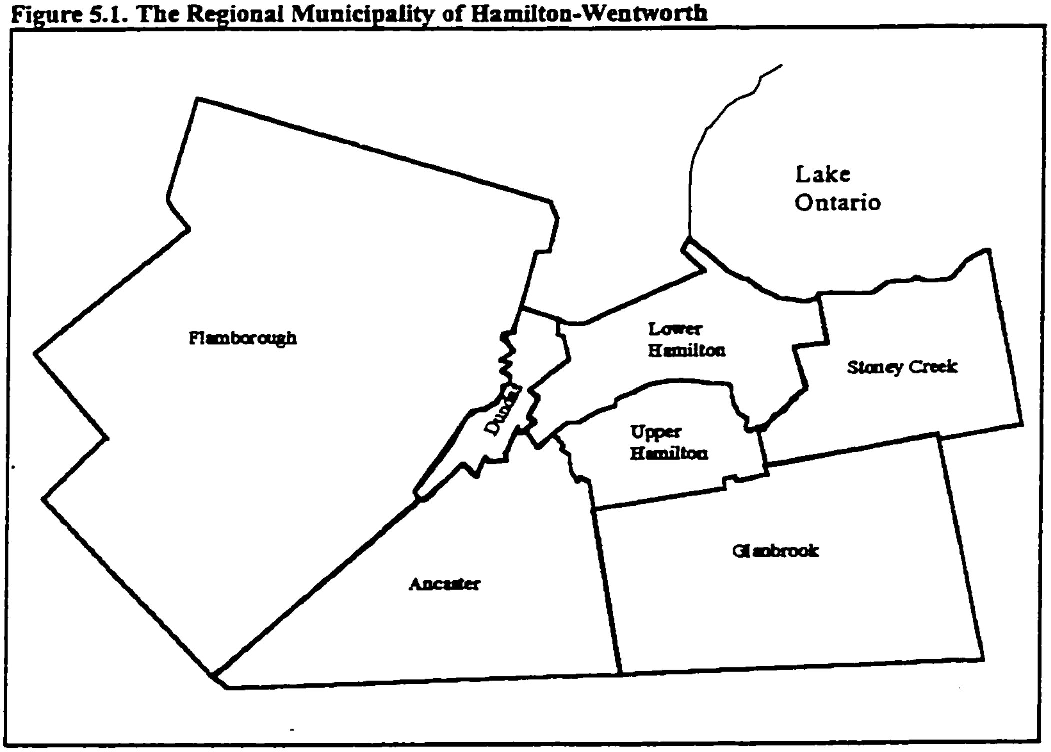


Figure 5.2. Chart showing allocation of property taxes used in questionnaire

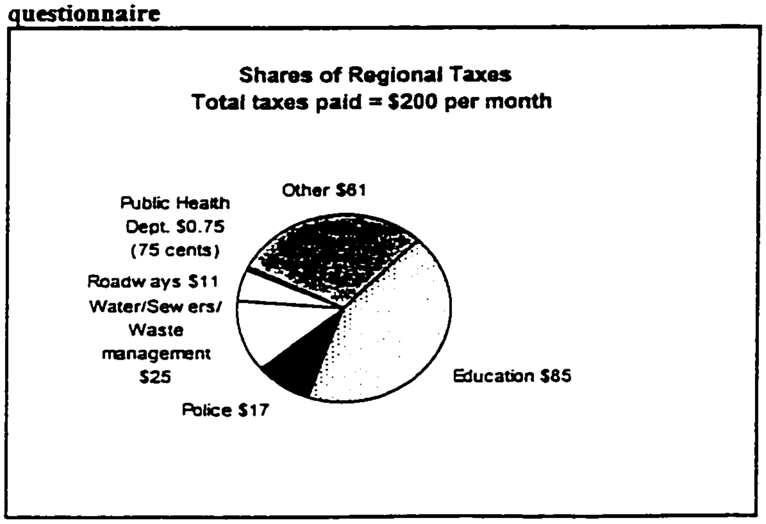


Figure 5.3. Sample Question Taken from the Questionnaire

	State of the Environment			
Attribute	Choice A (Current situation)	Choice B	Choice C	Choice D
Bad Odour	4 days of bad odour per month	One-third Worse	One-third Better	Same
Black Fallout (BFO)	3 days of BFO per month	One-third Worse	One-third Better	Same
Poor Visibility	3 days of poor visibility per month	One-third Worse	One-third Better	Same
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	One-third Worse	One-third Better	Same
Monthly Property Taxes or Rent	$200 per month for typical household	$5 more taxes/rent	$5 less taxes/rent	$5 more taxes/rent
Rank	*2*	*4*	*1*	*3*

Please put a 1 under the choice you like BEST.
Please put a 4 under the choice you like LEAST.
Consider the remaining choices. Put a 2 under the one you like best. Put a 3 under the other one.

A Sample Question

Suppose that you like choice C the best. You would put a 1 under choice C as shown.
Suppose that you like choice B the least. You would put a 4 under choice B as shown.
Choice A and D remain. Suppose you like Choice A better than Choice D. You would put a 2 under that Choice A and a 3 under Choice D, as shown.

Table 5.1. Levels of the attributes employed

Attribute	Current Situation	About One-third Better	About One-third Worse
Bad Odour	4 days of bad odour per month	3 days of bad odour per month	5 days of bad odour per month
Black Fallout (BFO)	3 days of black fallout per month	2 days of black fallout per month	4 days of black fallout per month
Visibility	3 days of poor visibility per month	2 days of poor visibility per month	4 days of poor visibility per month
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	12 extra hospital admissions per month and 1 extra death per month, compared to perfectly clean air	24 extra hospital admissions per month and 3 extra deaths per month, compared to perfectly clean air

Table 5.2. Estimation of cardio-respiratory hospital admissions

Disease and associated ICD9* codes	Hospital discharges of Hamilton-Wentworth residents (1990)
cardiac:	
all ischemic, 410-414	3107
all other heart disease 390-398, 415-429..	1951
total cardiac	5108
respiratory:	
pneumonia, 480-486	961
copd**, 491-496	978
asthma, 493	468
total respiratory	2407

*International Classification of Diseases, Ninth Revision
source: Dickson *et al.*, 1995
**chronic obstructive pulmonary disease

Table 5.3. Increase or decrease in property tax by choice set and version

	Version					
Choice Set	A	B	C	D	E	F
1 (Q23)	$50	$5	$10	$15	$20	$25
2 (Q24)	$15	$20	$25	$50	$5	$10
3 (Q25)	$5	$10	$15	$20	$25	$50
4 (Q26)	$20	$25	$50	$5	$10	$15
5 (Q27)	$5	$10	$15	$20	$25	$50
6 (Q28)	$25	$50	$5	$10	$15	$20
7 (Q29)	$10	$15	$20	$25	$50	$5
8 (Q30)	$10	$15	$20	$25	$50	$5
9 (Q31)	$15	$20	$25	$50	$5	$10

The amount shown is the size of the property tax increase or decrease specified in each case. For example, question 23 of version B (shown in Appendix A) contains either an increase or decrease of $5 as the tax change.

Chapter 6

GENERAL RESULTS

6.1. Introduction

This chapter reports some of the general results of the survey including the response rates and results to sections A and section C of the survey - respondents' general attitudes towards air quality in Hamilton-Wentworth - and the demographic characteristics of the sample respectively. Results were stratified according to the six cities and townships that compromise the Regional Municipality of Hamilton-Wentworth when applicable. Where possible[12], the City of Hamilton was decomposed into Hamilton Lower and Hamilton Upper, characterized according to being either above or below the escarpment. At the conclusion of the chapter the effects of cognitive difficulty, as measured by whether respondents were able to properly undertake the ranking exercise, are addressed. The chapter concludes with an examination of the characteristics of those respondents that exhibited cognitive difficulties, including multivariate analysis.

---

[12]When using census data this was not possible, for example.

6.2. Sample characteristics

The final version of the questionnaire was mailed to 1,908 addresses, obtained from tax assessment rolls, on February 10, 1997. 246 questionnaires were returned because the addressee had moved and thirteen others were removed from the sample because the recipients had died, were incapable of answering, or had received a misprinted questionnaire. 515 completed surveys (31% of the remaining 1649) were returned by March 19, 1997 (Table 6.1). Surveys were returned from all of the regions and there was proportionate representativeness of each region. The crude response rates varied form 24% to 33% for each region. However, given the small populations in some regions, the number of respective responses was small. For instance, only 13 questionnaires were received from Glanbrook residents. Response rates within Ancaster were the highest, while Hamilton exhibited the lowest response rate (Table 6.2).

The median age of the respondents was 42 years which is slightly lower than the median age (48) for those over the age of 19 in the entire Regional Municipality of Hamilton-Wentworth (Table 6.3). The survey sample displayed a higher percentage of those with university degrees (20% vs 10%) and a higher percentage of homeowners (84% vs 62%) than is found in the entire region. There was also a lower proportion of low-income households in the survey sample, a higher median income, and a lower rate of unemployment. With respect to the health characteristics of the respondents a higher percentage stated that they suffered from respiratory problems, while a similar proportion stated that they had heart or circulatory problems. There were also significant differences within the sample itself when the characteristics of the separate cities or townships are examined (Table 6.4). The most striking finding was the range in median household income - from a high of $85,0000 in Ancaster, to a low of $30,734 in Lower Hamilton. In fact, the respondents of Lower Hamilton exhibited the lowest median income, highest percentage of low-income households, highest median age, and the lowest percentage of those with university degrees. The area of lower Hamilton encompasses the industrial area of the city and exhibits poorer air quality than the rest of the region. The mean income of the entire sample was $43,550 which was slightly higher than the mean of the Region which was $41,213.

6.3. Experience and attitudes towards air quality

The survey elicited information regarding people's experience with each of the four attributes. With respect to bad odour, black fallout and poor visibility, the respondents were first given a description of the attribute and then asked to state their experience with it. With respect to cardio-respiratory disease, each respondent was asked whether they or another family member had been diagnosed with either asthma, heart disease, angina, emphysema or chronic bronchitis, or any other breathing or heart disease. The results to these questions are presented in Table 6.5. Lower Hamilton had the greatest percentage of respondents reporting experience with bad odour and black fallout. Poor visibility was experienced by 94% of all respondents, and 97% of those from Upper Hamilton.

Visibility is less of a local phenomenon, however, and that may explain this finding. Black fallout was clearly more of a problem for those of Lower Hamilton, with 81% of the respondents experiencing the phenomenon, compared to the mean of 66% for the entire sample.

Examining the results to selected questions from the section A of the questionnaire, it is clear that the respondents perceived air quality in Hamilton-Wentworth to be poor (Table 6.6). The majority of the respondents (58%) thought the air quality in their neighbourhood was somewhat worse or much worse than the rest of Southern Ontario. Only 15% of the respondents thought that air quality in their neighbourhood was better than the rest of Southern Ontario.

The majority of the respondents were either very concerned or extremely concerned about each of the four air quality attributes. Respondents were most concerned with health effects, with 81.2% of the respondents reporting they were very concerned or extremely concerned. Concern for the remaining attributes was also high, with 70.4%, 58.0% and 56.0% of respondents reporting they were very concerned or extremely concerned about black fallout, bad odour and poor visibility respectively. Table 6.7 breaks these results down by cities and townships. Recall that these attributes had not yet been described when the questions about level of concern were posed.

As a social issue, air quality ranks slightly below crime, unemployment, and the quality of the educational system, distinctly above the level of taxes, and far above the quality of snow removal. Health effects were ranked first in personal importance by 85% of the respondents, while black fallout, bad odour, and poor visibility were ranked first by 9%, 7%, and 4% of the respondents, respectively (Table 6.8). Of note is that the respondents from Lower Hamilton ranked bad odour, black fallout and poor visibility higher than the average for the entire region. This factor was probably attributed to the higher percentage of experience of the attribute amongst those respondents.

6.4. Cognitive difficulty

Some respondents had difficulty in ranking the choice sets in Section B of the questionnaire. In some cases respondents left a choice set entirely unranked, in others they gave only a partial ranking or gave the same ranking to more than one choice. Hence if a respondent only provided their first 3 rankings within a choice set, that was considered to be a partial ranking. All of these cases will be referred to as an incomplete ranking. 74 of the 515 respondents did not attempt the ranking exercise at all. A further 61 respondents committed at least one incomplete ranking. Hence, a total of 135 of the respondents committed at least one incomplete ranking. Table 6.9 reports the breakdown of the number of incomplete rankings that occurred by respondents.

Recall that 9 of the 54 possible pairwise comparisons in the choice sets involved a dominance relationship. A situation in which a respondent ranked one of the dominated alternatives higher than dominant alternative will be defined as a dominance violation. It turns out that of the 441 respondents who attempted the ranking exercise, 209 violated at least one dominance relationship (see Table

6.10). Note that some respondents may have committed both a dominance violation and an incomplete ranking. A total of 48 respondents did both. A total of 188 respondents provided a complete ranking for each choice set and committed no dominance violations. Table 6.11 presents these results. Figure 1 provides an accounting of the entire survey sample starting with the original mailing and a breakdown of those who completed the survey according to whether a complete ranking was provided and whether or not a dominance violation had occurred.

Table 6.12 reports results about incomplete rankings and dominance violations according to choice sets. Note that only four choice sets included dominated alternatives. There was not a clear relationship between the number of possible dominance violations within a choice set and the number of violations that did occur. For instance 140 respondents committed a dominance violation in Choice Set 4 which included only one dominance situation, whereas only 52 violations occurred in Choice Set 2 which included two possible violations. Nor did there appear to be any relationship between dominance violations and the placement of the Choice Set. One hypothesis is that fatigue set in as the ranking exercise was being undertaken which would result in more dominance violation occurring in the latter choice sets. This, however, was not observed. Both Choice Sets 1 and 9 included three possible violations with 161 instances of dominance violations occurring in Choice set 1 and 126 dominance violations occurring in Choice Set 9. However, there does appear to be positive relationship between the number of incomplete rankings per choice set and the timing of the choice set. For instance while Choice Set 1 had 6 instances of an incomplete ranking, there were

31 incomplete rankings associated with Choice Set 9.

It is unclear whether the number of choice sets employed or the number of alternatives used may have been the cause of some of the cognitive difficulties. It is possible that if respondents were only presented with pairwise comparisons that they may have committed less dominance violations as the differences between the alternatives would have been more explicit. As well with pairwise comparisons only the first ranked alternative would need to be considered also simplifying the task. However, if as many alternatives would have been incorporated in the study it would have required 54 choice sets for each respondent which may have also made the task appear to be more onerous. Unfortunately the design of the survey does not allow to test whether the size of the choice sets did affect the results.

The respondents who committed ranking or dominance violation were on average distinctly older (52 yrs vs 45 yrs), poorer (median household income of $31,300 vs $51,600 and 36% low-income households vs 14%) and less well educated than the remaining respondents (12% with university degrees vs 21%). These results are presented in Table 6.13.

In order to further examine what individual specific characteristics may predict whether a respondent committed either a dominance violation or an incomplete ranking a multivariate analysis was conducted. Using a logistic regression, the entire sample was split into two sub-samples depending on whether or not a violation had been committed.[13] The dependant variable was discrete and equal to 1 if the respondent was classified as "clean."[14] The regression results from this analysis are displayed in Table 6.14. The multivariate analysis indicated that those respondents who had experience with one of the cardio-respiratory diseases listed in the questionnaire (either themselves or a family member), were younger, had a post-secondary education, had no children and were female were likely more likely to not have committed any violations in completing the ranking exercise. The positive influence of age peaks at age 40. After that age had a negative effect on whether or not a violation had occurred.[15] While it may have been a-priori expected that those who were younger and better educated would have fared better on the ranking exercise, it is unclear how the other variables may have affected one's ability to complete the task properly. One hypothesis is that those who had experience with cardio-respiratory diseases were more diligent in completing the task as they believed it to be important, while those who had no children may have had more time available to carefully complete the task.

The cognitive difficulties experienced in the application of the survey was not a unique outcome. Viscusi *et al.* (1991) included similar tests for cognitive

---

[13]Note that only respondents that attempted the ranking exercise were included in this analysis.

[14]Clean equals 1 for those who did not commit a violation, 0 otherwise.

[15]From the multivariate regression equation the optimum of the age relationship can be calculated as -(estimated coefficient on age)/(2*estimated coefficient on age^2).

difficulty and found similar results. Their study included a consistency check to identify cases in which a respondent switched preferences. Given the design of this study it was not possible to have similar checks. Viscusi *et al.*'s study, however, did include dominance tests, in a fashion similar to the dominance tests included in our study. They found that 20% of respondents committed a dominance violation and 27% failed at least one consistency check. As those authors discuss, however, these results point out an advantage of the conjoint analysis approach over the standard contingent valuation approach.

> These consistency checks distinguish our approach from the usual CV method, in which respondents' answers are taken at face value without such formal tests of whether the subjects understood the valuation task and displayed consistent choices........these checks are an important check on subjects' understanding of the choice task. In contrast the CV approach includes no such systematic checks of task comprehension on a question by question basis and, not surprisingly, allows researchers to use a larger fraction of subjects' responses in the analysis. (Viscusi *et al.* 1991)

Figure 6.1 Survey sample, returned surveys, incomplete rankings and dominance violations.

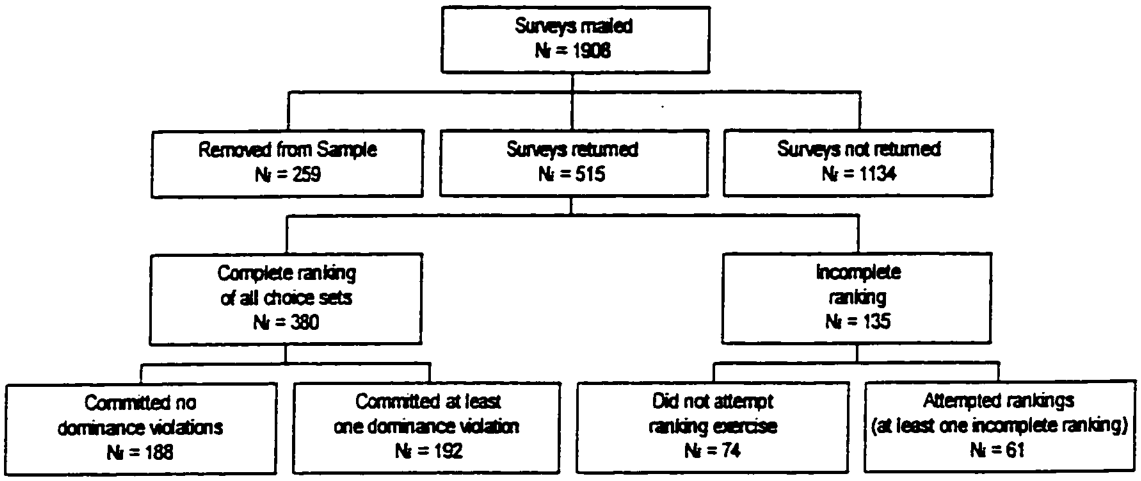


Table 6.1. Response statistics

Total questionnaires mailed	1908
Removed from sample:	
returned undelivered/moved	246
deceased	7
incapable	5
misprint	1
Total removed from sample:	259
Sample remaining	1649
returned completed	515
response rate:	31%*

notes:

- deceased: these were either returned unopened and marked as deceased, or phoned in by a relative after receiving the postcard
- incapable: blind, mentally challenged or elderly, as informed by a relative via letter or phone call
- some of the non-respondents may also belong to one of the above groups.

*515/1649

Table 6.2 Population and response rates by city/township

City/ Township	Pop'n[1]	Random Sample (% of pop'n)		Number Returned	Response Rate[2]
Hamilton	318499	1451	(0.46%)	346	24%
Stoney Creek	49968	205	(0.41%)	61	30%
Ancaster	21988	85	(0.39%)	29	34%
Dundas	21868	95	(0.43%)	26	27%
Flamborough	29616	123	(0.42%)	40	33%
Glanbrook	9726	39	(0.40%)	13	33%
Total	451665	1998	(0.44%)	515	26%

[1] source; 1991 Census

[2] Note: This is the crude response rate. The returned undelivered and invalid questionnaires have not been dropped as this information was not recorded.

Table 6.3. Sample characteristics, survey sample vs region of Hamilton-Wentworth[a]

Characteristic	survey sample	Hamilton-Wentworth
percentage female	54	51[d]
percentage with university degree	15	10[d]
percentage low income households (<$30,000 per year)	28	36[d]
median income[g]	$43,550	$41,213
percentage homeowners	84	62[d]
median age[g]	48	42[d,f]
percentage unemployed	3	7[d,h]
percentage with respiratory problems[b]	11	7
percentage with heart or circulatory problems[c]	6	6

[a]Note that some of the census data and all of the OHS data is based on a sample of the entire Region of Hamilton-Wentworth
[b]Defined as asthma. emphysema, chronic bronchitis, or persistent cough. In the survey persistent cough was not included
[c] Defined as "heart disease" in the survey, "Circulatory and heart disease" in the Ontario Health Survey
[d]Source: 1991 Census
[e]Source: 1990 Ontario Health Survey
[f]median age of the population over the age of 19 in Hamilton-Wentworth
[g]estimated by assuming a uniform distribution within each income and age category
[h]estimated based on the unemployment rate for the Hamilton CMA in December, 1996

Table 6.4. Sample characteristics, stratified by cities and townships

Characteristic	Entire sample	Hamilton Upper	Hamilton Lower	Stoney Creek	Ancaster	Dundas	Flam-borough	Glanbrook
percentage female	54	49	53	60	55	60	63	46
percentage with university degree	20	20	13	13	46	39	26	15
percentage low income households (<$30,000 per year)	28	27	39	26	14	11	13	0
median income ($'s)*	43550	40379	30734	55000	85000	52550	65714	65500
percentage homeowners	84	73	87	90	100	91	98	100
median age*	48	49	51	47	50	45	45	42
percentage unemployed	3	4	6	0	0	0	0	0

*estimated by assuming a uniform distribution within each income and age category

Table 6.5. Experience with attributes of air quality, stratified by cities and townships

Attribute	entire sample	Hamilton Upper	Hamilton Lower	Stoney Creek	Ancaster	Dundas	Flam-borough	Glanbrook
cardio-respiratory disease	54%	53%	55%	57%	48%	46%	55%	54%
bad odour	79%	78%	86%	80%	59%	79%	73%	42%
black fallout	66%	63%	81%	68%	45%	46%	42%	38%
poor visibility	94%	97%	95%	92%	89%	96%	85%	77%

Table 6.6. Responses to Selected Questions in Section A (in percentages)

A) *How would you describe the air quality in your neighbourhood compared to the rest of Southern Ontario? (Question 1)*

Much better than the rest of southern Ontario	Somewhat better than the rest of Southern Ontario	The same as the rest of Southern Ontario	Somewhat worse than the rest of Southern Ontario	Much worse than the rest of Southern Ontario	Don't Know
2.5	12.8	18.8	41.5	16.9	6

B) *How concerned are you about annoying odours/black fallout/poor visibility/ health effects of air pollution in Hamilton-Wentworth? (Questions 2-5)*

	N*	Not at all concerned	A little concerned	somewhat concerned	very concerned	extremely concerned	don't know
annoying odours	511	4.3	11.8	25	36	22	0.4
black fallout	511	4.1	7	15.7	34.5	35.9	2
poor visibility	511	5.2	12.2	24	34	22	1.6
health effects	512	1	4.8	11.4	31.8	50	0.2

C) *Do you think that the issue is a more serious issue, a less serious issue, or about as serious an issue as air quality? (Questions 6-10)*

Issue	N*	more serious issue	less serious issue	equally as serious an issue
the level of crime	509	32.4	11.2	55
the level of taxes	504	17.7	40.3	39.5
snow clearing	504	8.5	53.1	36.1
quality of educational system	505	28.3	16.3	53.3
level of unemployment	507	27.9	20.5	49.8

*some respondents did not answer all of the questions or may have recorded multiple responses and were excluded from these statistics

Table 6.7. Percentage of respondents very or extremely concerned re attributes of poor air quality

Attribute	Entire sample	Hamilton Upper	Hamilton Lower	Stoney Creek	Ancaster	Dundas	Flam-borough	Glanbrook
health effects	82	83	85	83	72	85	75	69
bad odour	58	57	66	44	55	62	45	54
black fallout	70	72	77	69	62	62	48	85
poor visibility	56	55	57	66	45	46	55	62

Table 6.8. Most important attribute, stratified by city or township

	entire sample	Hamilton Upper	Hamilton Lower	Stoney creek	Ancaster	Dundas	Flamb-borough	Glanbrook
	% ranked 1st	% ranked 1st	% ranked 1st	% ranked 1st	% ranked 1st	% ranked 1st	% ranked 1st	% ranked 1st
bad odour	7.2	5.1	10.6	7.1	0	8.7	7.9	0
black fallout	8.9	5.7	16.8	7.4	3.9	0	2.6	0
poor visibility	4.0	3.2	5.6	5.6	0	0	5.3	0
health effects	85.1	89.4	77.8	79.7	96.2	87.5	92.3	100

Table 6.9. Number of incomplete rankings by respondent

# of incomplete rankings	# of respondents
no incomplete rankings	380
1 incomplete ranking	33
2 incomplete rankings	6
3 incomplete rankings	4
4 incomplete rankings	3
5 incomplete rankings	2
6 incomplete rankings	5
7 incomplete rankings	2
8 incomplete rankings	6
9 incomplete rankings	74
total	515

Table 6.10. Number of dominance violations by respondent

# of incomplete rankings	# of respondents
no dominance violations	209
1 dominance violations	82
2 dominance violations	59
3 dominance violations	58
4 dominance violations	16
5 dominance violations	9
6 dominance violations	7
7 dominance violations	0
8 dominance violations	1
9 dominance violations	0
total	441*

*Only those respondents that attempted the ranking exercise were included

Table 6.11. Distribution of respondents with dominance violations and incomplete rankings

	incomplete ranking	no incomplete rankings	total
dominance violation	48	192	240
no dominance violation	87	188	275
total	135	380	515

Table 6.12. Number of incomplete rankings and dominance violations by choice set*

choice set	dominance violation[§]	incomplete ranking
1	161	6
2	52	7
3	-	14
4	140	20
5	-	23
6	-	24
7	-	23
8	-	23
9	126	31

*Only includes those who attempted to rank at least one choice set
[§] Recall that only choice sets 1,2,4 and 9 included dominated alternatives
Choice set 1 - 3 possible dominance violations
Choice set 2 - 2 possible dominance violations
Choice set 4 - 1 possible dominance violation
Choice set 9 - 3 possible dominance violations

Table 6.13. Selected demographic characteristics by dominance violations and incomplete rankings

	entire sample (N=515[a])	no incomplete rankings or dominance violations (N=188[a])	one or more incomplete rankings or dominance violations [c] (N=327[a])
median age	48	45	52
median income[b]	$43,550	$51,600	$31,300
percentage low-income households (<$30,000)	28	14	36
percentage unemployed	3.48	2.78	3.88
percentage with university degree	15	21	12

[a] note: Not all of the respondents answered all of the questions
[b] estimated by assuming a uniform distribution within each income category
[c] also includes those who did not attempt the ranking exercise at all

Table 6.14. Results from multivariate logistic analysis‡

variable*	estimated coefficient	standard error
experience with cardio-respiratory disease †§	0.55	0.25
experience with bad odour †	0.36	0.35
experience with black fallout †	-0.01	0.29
experience with poor visibility †2	-0.85	0.54
age §	0.16	0.07
age-squared §	-0.002	0.0006
post-secondary education †§	0.64	0.27
full-time employment or housework †	0.19	0.28
couple without children † §	0.77	0.32
lone parent with children †	0.44	0.53
non-family, one person † §	1.51	0.48
non-family, more than one person †	0.05	0.59
health professional †	-0.41	0.31
member of environmental organization †	0.09	0.32
male †§	-0.49	0.25
income midpoint	0.01	0.004
own home †	0.62	0.5
live in house †	-0.51	0.6
recently moved †	-1.28	0.71
active exerciser †	-0.21	0.25
reside in Dundas †	0.89	0.69
reside in Ancaster †	-0.12	0.63
reside in Glanbrook †	-0.09	0.78
reside in Stoney Creek †	-0.64	0.52
reside in Hamilton Lower†	-0.22	0.45
reside in Hamilton Upper †	0.11	0.44

‡Dependant variable was equal to 1 if respondent was clean (i.e. committed no violations)

*These are, in general, the demographic information obtained from the questionnaire. See the questionnaire for more details on these variables.

†zero-one dummy variable.

§statistically significant at the 5% level.

Chapter 7

WILLINGNESS-TO-PAY AND HEALTH TRADE-OFFS

7.1. Introduction

This chapter presents and discusses the regression results of the four models that were presented in Chapters 5 - the linear model, the linear model augmented by a squared tax term, the full quadratic model, and the full quadratic augmented with demographic information. Not only was it important to examine the appropriate functional form, but the appropriate data set had to be established as well. Recall, from the previous chapter, that not all of the respondents answered all of the choice sets properly. Hence, it was investigated (using the Chow test for structural change) whether those who made some violations did indeed respond differently than those who made no violation and whether or not such individuals should be included in the analysis. Both willingness-to-pay (WTP) and non-monetary trade-offs - namely the health trade-off (HTO) - are addressed. The chapter then proceeds to examine whether there are any gains to be made from using Chapman and Staelin's (1982) procedure for expanding the data set to include second and third choices in conjunction with respondent's first choices. After examining the construct validity of the WTP results, the chapter concludes with an examination of whether there were any regional differences in the WTP and HTO results.

7.2. Basic regression results

Given the evidence of cognitive difficulties in completing the questionnaire, it was decided to first use only the "cleanest" set of data, that is, the set of 188 respondents who committed neither ranking nor dominance violations. Using only the first ranked choices, the four models introduced in Chapter 5 were estimated. Tests were undertaken in order to ascertain the correct specification. In some instances, particularly in the examination of geographic regions, since some of the regions include small sample sizes, it was not possible to use the larger models due to low degrees of freedom. Table 7.1 presents the general results of the four models that were estimated.

Note that all of the principal variables (health, H; odour, O; black fallout, F; visibility, V; taxes, T) were highly statistically significant in the two linear models and the quadratic model[16]. As well, the squared tax term was also highly statistically significant in all the models in which it appeared[17]. Note, however, that in the quadratic with demographics model, while the black fallout and the visibility variables were not significant at the 95% confidence interval (based on t-tests), the entire group of black fallout and visibility variables (F, V, HF, HV, OF,

---

[16] $p<.001$ for all of the variables.

[17] $p<.001$ in the linear with quadratic taxes model and in the quadratic model. $p<.002$ in the quadratic plus demographics model.

OV, F^2, FV, FT, V^2, VT, IF, IV, EF, EV, AF, AV, CF, CV) as a whole proved to be significant (based on a likelihood-ratio test[18]). A plausible explanation for the finding that F and V individually were insignificant variables was due to the manner in which the demographics were incorporated into the model. It was, however, difficult to avoid this problem as the theoretical framework dictated that the demographics had to be introduced in this way as explained in Chapter 5. Note that the willingness-to-pay values and the health trade-off values do not significantly differ between the full quadratic and the full quadratic model augmented with the demographics. It is indeed these trade-offs in which we are ultimately interested and not the actual coefficients themselves. As well, the standard errors on the estimated values of these trade-offs are all quite small. Hence, there seems to be some justification in using the full quadratic model augmented with the demographics. These trade-offs will be discussed in the next two sections.

The signs on all coefficients were consistent with a priori expectations (Table 7.1.). The signs of the health, odour, black fallout, and visibility coefficients are all positive, implying that the respondents were more likely to choose an alternative that had a better level of each adverse attribute (i.e. a one-third decrease in the number of days) while the tax coefficient had a negative sign associated with it implying that the respondents were less likely to choose an alternative with a higher level of taxes associated with it.

---

[18] $p<.001$

Likelihood ratio tests were conducted in order to see whether adding the variables that were included in the larger models resulted in a better model. Indeed Model 1, the fully linear model ($p<0.001$ on a likelihood ratio test), and Model 2, the linear form augmented by a quadratic term in taxes ($p<0.001$), were both rejected, as was the fully quadratic model (Model 3) in these tests. As a group, the demographic variables were highly significant ($p<0.001$[19]). Hence, the results from the quadratic model with demographics will be used as the reference point for further discussion unless otherwise noted. (As previously noted, there were instances in which it was necessary to use the linear models.) Appendix D (at the end of the thesis) reports the results from the aforementioned likelihood ratio tests.

7.3. Willingness-to-pay and health trade-off

Table 7.2 reports the willingness-to-pay for the four air quality attributes computed for the four aforementioned models, using the formulae developed in Chapter 5. The baseline regression (full quadratic plus demographics) implies that the respondents were willing to pay an additional $39.70 per month in property taxes to achieve a reduction of one-third in the adverse health effects of air pollution (i.e. to reduce excess hospital admissions from 18 to 12 per month and to reduce excess mortality from two deaths to one death per month. They were

---

[19]Note that in conducting the LR test comparing the full quadratic model with the full quadratic plus demographics model the data set for BOTH models was limited to those individuals that stated all the necessary demographic information in the questionnaire.

willing to pay $13.10 per month for a one-third improvement in odour (from 4 to 3 days per month), $12.80 per month for a one-third improvement in black fallout (from 3 days to 2 days per month), and $14.40 per month for a one-third improvement in visibility (i.e. from 3 poor visibility days per month to 2 poor visibility days per month). Note that all of these estimates were statistically significant (i.e. the standard errors are quite low).

In general the magnitude of the WTP and HTO estimates fall as we move to the larger models. This decrease was due to both a reduction in the point estimates of coefficients involving the air quality attributes and to an increase in the point estimate of the tax coefficient. The tax coefficient estimate increases by almost two-fold between the linear model and the full quadratic model augmented by the demographics. Because WTP estimates are so sensitive to the tax coefficient it is particularly useful to examine the health trade-offs which are more stable. Hence, it may be more relevant to examine the health trade-off.

The health trade-off calculates the trade-off between health and each attribute as opposed to willingness-to-pay which calculates the trade-off between dollars and each attribute. By doing so any problems associated with the sensitivity of the tax coefficient are eliminated and it allows for an estimation of the trade-offs between health and the other attributes that is not sensitive to the tax coefficient. Table 7.3 presents estimates of the health trade-off. Note that the health trade-offs (HTO) for the attributes (other than tax) are quite consistent between the models, indicating that it was indeed a more stable measure to ascertain the relative trade-off between attributes.

7.4. Expanding the data set

The sensitivity of these estimates with respect to the quality of the data used in the estimation was also investigated. To address the issue of the quality of the data, the models were estimated using two larger supersets of the data. As a first step in enlarging the data set, all of the completely ranked choice sets that were free of dominance violations were included. Thus, for example, even if an individual had committed a dominance violation in one choice set the information from the remaining eight choice sets was used. This data set, referred to as the "medium data set" included a total of 2756 choice sets from 361 individuals. As a second step, all completely ranked choice sets were included, whether they included a dominance violation or not. This data set, referred to as the "large data set," included a total of 3155 choice sets from 363 individuals. (See Figure 7.1 for a breakdown of these data sets.) Out of the total 515 respondents, 107 individuals did not give all of the demographic information needed for these regressions. Hence, there were a possible 408 individuals, or 3672 choice sets, that could have been included in the analysis. However, of the 3672 choice sets, 517 did not include a complete ranking, leaving a total of 3155 choice sets in the large data set. 399 of those choice sets included a dominance violation, leaving a total of 2756 choice sets in the medium data set. Note that when using first choices only it may have been possible to further expand the data set by using all choice sets that had at least a first ranking.

Table 7.4 reports the main regression results on the full quadratic plus demographics model using each of the three data sets. Table 7.5 reports the WTP and HTO estimates for the respective model and data sets. Examining Table 7.5 it can be seen that the willingness-to-pay estimates are indeed sensitive to the "quality" of the data used. For example, willingness-to-pay for improved health rises from \$39.70 to \$66.50 to \$159.30 per month as the quality of the data set diminishes. However, as previously noted, this feature appears be due to the decline in the size of the tax coefficient. Examining the health trade-off (also in Table 5)it is clear that it is a much more stable estimate of the overall trade-off between attributes.

A question then arose as to which data set was the most appropriate one to employ. One argument that can be put forward for using either the medium or large data set is that they allow us to take advantage of more data points. Also, the estimates, at least those of the health trade-off, are more precise using the larger data sets, as shown by the decrease in the standard errors. Expanding from the most limited data set to include all the choice sets that did not have a dominance violation (i.e. the medium data set) seems reasonable as people may have tired and committed errors in later choice sets, or may have just made an honest mistake on one of the choice sets, but did indeed understand the ranking task.

There are two further compelling arguments for expanding the data set to include all of the usable choice sets (i.e. all of those with a complete ranking), even those in which a dominance violation was committed. The first reason goes back to the theoretical construct of the random utility model which has been employed. Recall that in the random utility model, utility is composed of a systematic and a random component as follows.

$$U_{ij}=\mu_{ij}+e_{ij} \tag{20}$$

While both V_{ij} and e_{ij} are known to the individual, the e_{ij} are unobservable by the researcher and hence are random variables from the researcher's perspective. Hence, it is possible that the respondent does indeed have rational preferences but we are observing what appears to be an irrational choice due to the random component of the utility function[20]. Dropping the choice sets hence assumes that the e_{ij}s are close to zero.

Secondly, those respondents who are making the dominance violations may just not really care about the attributes and decided to act randomly. This behaviour is consistent with a large variance in e_{ij}. These individuals may not care about the environment in general and are expressing their preferences by behaving in this manner. It was found that those who committed dominance violations were poorer and less educated than those who did not commit any violations. It may be

[20]Rational preferences are defined as those in which an individual prefers a decrease in the number of days of each air quality attribute and a decrease in monthly property taxes.

that the environment is a luxury good and as such low-income respondents may not be willing to pay very much for improvements in it. Hence, it would be desirable to include their responses. Note, however, that according to the multivariate analysis discussed in section 5.4. belonging to an environmental organization was not a significant explanatory variable for being a "clean" individual. By limiting the data employed in the analysis to only those respondents who committed no violations, the representativeness of the data may be diminished.

These arguments, based on qualitative issues, suggest employing the medium or large data sets. However, before making the decision to use one of these supersets of data it seemed important to conduct an objective, or quantitative examination, as to whether there was any difference in these data sets. Hence, it was decided to run a series of Chow tests of structural change (Greene, 1993). These tests were able to inform us as to whether the incremental choice sets resulted in statistically different responses than the choice sets included in the small data set. As a first step, the small data set was compared to the medium data set and the large data set respectively using all of the four models previously discussed. (The medium and large data sets could not be compared as the incremental sample between the two data sets was too small to allow for enough degrees of freedom.)

In each case, the null hypothesis that the estimated coefficients from the "additional" respondents are the same as those from the initial set of respondents was rejected. In other words, those respondents that committed at least one ranking or dominance violation had responses that were different than those who made no such violations. One possibility for this result was that the models were not correctly specified with respect to the demographic variables that were significant in explaining the differences between "clean" and "not-clean" respondents. Hence, the Chow tests were conducted using two more models. Model 5 included the full quadratic model augmented by age, age-squared, and dummy variables for sex, post-secondary education, experience with cardio-respiratory disease, and no children. Model 6 was an augmentation of Model 5 including the income variable that was included in Model 4[21]. The null hypothesis

---

[21]Model 5:

$$\begin{aligned}\mu = {} & \beta_1 H + \beta_2 O + \beta_3 F + \beta_4 V + \beta_5 T + \beta_6 (H \cdot O) + \beta_7 (H \cdot V) + \beta_8 (H \cdot T) + \beta_9 (H \cdot F) \\ & + \beta_{10} H^2 + \beta_{11} O^2 + \beta_{12} (O \cdot V) + \beta_{13} (O \cdot F) + \beta_{14} (O \cdot T) + \beta_{15} V^2 + \beta_{16} (V \cdot F) \\ & + \beta_{17} (V \cdot T) + \beta_{18} F^2 + \beta_{19} (F \cdot T) + \beta_{20} T^2 + \beta_{21} (A \cdot H) + \beta_{22} (A \cdot 0) + \beta_{23} (A \cdot F) \\ & + \beta_{24} (A \cdot V) + \beta_{25} (A \cdot T) + \beta_{26} (A^2 \cdot H) + \beta_{27} (A^2 \cdot O) + + \beta_{28} (A^2 \cdot F) + \beta_{29} (A^2 \cdot V) \\ & + \beta_{30} (A^2 \cdot T) + \beta_{31} (C \cdot H) + \beta_{32} (C \cdot O) + \beta_{33} (C \cdot F) + \beta_{34} (C \cdot V) + \beta_{35} (C \cdot T) + \beta_{36} (P \cdot H) \\ & + \beta_{37} (P \cdot O) + \beta_{38} (P \cdot F) + \beta_{39} (P \cdot V) + \beta_{40} (P \cdot T) + \beta_{41} (M \cdot H) + \beta_{42} (M \cdot O) + \beta_{43} (M \cdot F) \\ & + \beta_{44} (M \cdot V) + \beta_{45} (M \cdot T) + \beta_{46} (NK \cdot H) + \beta_{47} (NK \cdot O) + \beta_{48} (NK \cdot F) + \beta_{49} (NK \cdot V) + \beta_{50} (NK \cdot T)\end{aligned}$$

Model 6:

$$\begin{aligned}\mu = {} & \beta_1 H + \beta_2 O + \beta_3 F + \beta_4 V + \beta_5 T + \beta_6 (H \cdot O) + \beta_7 (H \cdot V) + \beta_8 (H \cdot T) + \beta_9 (H \cdot F) \\ & + \beta_{10} H^2 + \beta_{11} O^2 + \beta_{12} (O \cdot V) + \beta_{13} (O \cdot F) + \beta_{14} (O \cdot T) + \beta_{15} V^2 + \beta_{16} (V \cdot F) \\ & + \beta_{17} (V \cdot T) + \beta_{18} F^2 + \beta_{19} (F \cdot T) + \beta_{20} T^2 + \beta_{21} (A \cdot H) + \beta_{22} (A \cdot 0) + \beta_{23} (A \cdot F) \\ & + \beta_{24} (A \cdot V) + \beta_{25} (A \cdot T) + \beta_{26} (A^2 \cdot H) + \beta_{27} (A^2 \cdot O) + + \beta_{28} (A^2 \cdot F) + \beta_{29} (A^2 \cdot V) \\ & + \beta_{30} (A^2 \cdot T) + \beta_{31} (C \cdot H) + \beta_{32} (C \cdot O) + \beta_{33} (C \cdot F) + \beta_{34} (C \cdot V) + \beta_{35} (C \cdot T) + \beta_{36} (P \cdot H) \\ & + \beta_{37} (P \cdot O) + \beta_{38} (P \cdot F) + \beta_{39} (P \cdot V) + \beta_{40} (P \cdot T) + \beta_{41} (M \cdot H) + \beta_{42} (M \cdot O) + \beta_{43} (M \cdot F) \\ & + \beta_{44} (M \cdot V) + \beta_{45} (M \cdot T) + \beta_{46} (NK \cdot H) + \beta_{47} (NK \cdot O) + \beta_{48} (NK \cdot F) + \beta_{49} (NK \cdot V) + \beta_{50} (NK \cdot T) \\ & + \beta_{51} (I \cdot H) + \beta_{52} (I \cdot O) + \beta_{53} (I \cdot F) + \beta_{54} (I \cdot V) + \beta_{55} (I \cdot T)\end{aligned}$$

where H,O,F,V,T,I,C,A are as previously defined, and
P = dummy for post-secondary education
M = dummy for male
NK = dummy for no children

(continued...)

was still rejected using these models. Hence it was clear that the differences between the respondents' responses with respect to their trade-offs was not attributable to the variables that were significant in explaining the differences between "clean" and "non-clean" respondents. Since it was apparent that the choice sets included in the medium and large data sets were indeed structurally different than those included in the small data set, it was decided to concentrate on the small data set for all further analysis.

7.5. Exploiting the ranking information

The gains to Chapman and Staelin's (1982) procedure for using information from second and subsequent rankings were also investigated as part of this study. The second ranked choice is the most preferred of the three alternatives that remain after the first choice is eliminated from the choice set. Thus a new data matrix can be created from the original by deleting all the rows corresponding to first choices and flagging as the most preferred choices originally ranked second. Combining the new data set with the original set effectively doubles the sample size. Similarly, third choices can be considered by deleting the first and second choices and flagging the third choices as most preferred. This procedure was applied to the full quadratic plus demographics model (Model 4) using the small data set. Table 7.6 provides the main coefficients

---

[21](...continued)
Note that age and experience with cardio-respiratory disease were already included in Model 2. The education variable included in Model 4 would be highly correlated with the post-secondary variable included in models 5 and 6 and hence was not used in those models.

results, and Table 7.7 provides the WTP and HTO results from adding the second- and third-ranked choices.

Table 7.6 reveals that incorporating both the first and the second-ranked choices is indeed successful in increasing the apparent precision of the estimates. The asymptotic error on willingness-to-pay and the health trade-off estimates fall for all of the attributes. Extending the procedure to consider the first three choices, however, does not further improve the estimates of willingness-to-pay. Standard errors begin to rise and the point estimates of the tax coefficients fall, resulting in large and less precise estimates of willingness-to-pay. Note that the WTP estimates on the non-health attributes are comparable using either first or first and second rankings, while the WTP estimate for the health attribute rose approximately 25%. This property is also signified in the reduction of the HTO estimates. When including third ranked choices as well, all of the WTP estimates increased proportionally resulting in the HTO estimates remaining relatively unchanged.

7.6. Construct Validity

The validity of a measure is the degree to which the instrument measures the theoretical construct under investigation. If we already knew the true WTP (or WTA) of a good or service, there would be no need to conduct a survey of stated preference. The two main types of validity are criterion and construct validity (Mitchell and Carson, 1989). Criterion validity is concerned with whether the measure of the construct is related to another measure which is the criterion (or "gold" standard). Therefore, criterion validity can be assessed for stated preference surveys only if one has observable market data on what people will actually pay for the good. Construct validity involves the degree to which the measure is correlated with other measures as predicted by theory. One form of construct validity, theoretical validity, asks whether the measure is related to measures of other constructs as predicted by theory. For instance, economic theory predicts that, everything else being equal, WTP should be greater the larger is one's income[22]. It may also be presumed that those individuals who experienced one of the attributes may exhibit a higher trade-off for that attribute in comparison to the other attributes. Hence, using the full quadratic model[23] (Model 3), validity tests were undertaken in order to examine whether such predicted hypotheses did

---

[22]This is based on the assumption that the environmental quality (or specifically the attributes employed) is a normal good.

[23]The full quadratic plus demographics (model 4) could not be used since the validity tests being undertaken related to some of the demographics that appear in model 2. Hence, the demographic variables could not be included in the regressions in order to avoid collinearity problems.

When using first choices one of the alternatives within Choice Set 1 were dropped from the analysis. Alternative 3 from that choice set was "nearly dominated" by the other alternatives and was rarely chosen as either the first ranked or second ranked alternative. It was ranked first by only three respondents who also reported their income information. However, none of the respondents in the low income group chose that alternative. Hence, there resulted a situation in which a particular pattern of a group of dependent variables (O^2=1, OV=1, F^2=0, FV=0) was associated with the dependent variable always being 0. This resulted in those variables and their respective standard errors being improperly estimated. Thus, it was decided to not use this alternative in the analysis of high and low income using first-choices only. Furthermore, when dropping this alternative there turned out to be still be a case of multicollinearity resulting in the variable F^2 being dropped from the logit regression.

indeed occur[24]. Both first choices only and first and second choices were employed in these tests.

In order to examine whether those with higher incomes did indeed express higher willingness-to-pay values, the entire sample was stratified into two sub-samples - high income and low-income - defined as annual household income of $50,000 or greater, or less than $50,000. Separate regressions were then estimated for each income group and willingness-to-pay was derived. These results are presented in Table 7.8. According to likelihood ratio tests the sub-samples were different from one another.[25] Examining the data of first choices only, WTP values were slightly higher for those respondents with higher incomes for each of the four attributes. However, when using first and second choices, those respondents with higher incomes exhibited a lower WTP for the health attribute. A plausible explanation for not finding a larger WTP differential may be attributable to the fact that those individuals with higher incomes also live in areas that do not experience poor air quality. In other words, they have already "paid" in order to live in areas of better air quality[26].

---

[24]It can be hypothesized as well that those respondents that support environmental organizations may exhibit higher WTP values than those respondents who do not. However, there were too few respondents (38) that did support such organizations to use in a test Hence, this was not examined. As well too few respondents claimed to not have experienced either bad odours or poor visibility (31 and 13 respondents respectively) to include these attributes in the analysis.

[25]Results from these LR tests are presented in the Appendix D.

[26]It was indeed found that those individuals who live in the areas of poor (continued...)

A second validity check involved examining the relationship between experience with cardio-respiratory disease and willingness-to-pay. In the survey, respondents were asked to state whether they or another family member had been diagnosed with various heart or lung conditions. The a priori expectation was that those residents that had suffered from a cardio-respiratory disease, or had relatives that did, would be willing to pay more for an improvement in the health attribute. A total of 109 of the 188 respondents who did not make any choice set violations stated that they or another family member had one of those conditions. Table 7.9 compares the WTP and HTO values of those respondents with the respondents that had no experience with the health variable. It was found that the other attributes were indeed valued lower in comparison to the health attribute in those respondents that had experience with cardio-respiratory disease as defined in the questionnaire. The health trade-offs for each of odour, black fallout, and poor visibility were lower for those who experienced cardio-respiratory diseases In other words, health has been valued higher in comparison to the other attributes by that cohort of respondents.

A similar check was conducted using those who had experienced black fallout. Contrary to what was a-priori expected, those who experienced BFO exhibited a lower WTP and a lower HTO than those who did not when using first choices only. When using both first and second choices those with BFO experience had both a higher WTP and a higher HTO for the attribute see Table

---

[26](...continued)
air quality reported lower household incomes (section 6.2).

7.10). It is unclear how to analyze these mixed results. A plausible explanation for the finding using first choices only is that those who experienced the attribute have become accustomed to BFO and did not find it to be as bothersome as the other attributes. However, given that the a prior expectation was found when both first and second choices was employed was assuring since this was found using more data from the respondents.

While the tests of construct validity were not definitive, they were promising. The finding that WTP was not explicitly positively related to income may have been due to regional differences. Given the finding that there were significant differences in income between regions and the fact that there were air quality differences between the regions gives credence to this hypothesis. Construct validity tests of income will be explored within regions in section 7.8. The finding that those with experience with the health attribute valued that same attribute, in comparison to the other attributes, higher than those without experience was promising as was the finding that those with BFO experience valued that attribute higher than those without the experience (using both first and second choices). Overall the validity tests were quite encouraging.

7.7. Geographic Distributions

Given the differences found with respect to both the demographic characteristics and the attribute experiences and concerns of the cities and townships that comprise the entire Region, and given that the some areas are geographically distinct from one another, it was decided to examine whether these differences extended to the trade-offs amongst the attributes as estimated by willingness-to-pay and the health trade-off. Hence, separate regressions were undertaken for each of the seven previously defined areas. Due to small sample sizes in some of the regions, the linear model with quadratic taxes was employed.

Table 7.11 presents both the regression results and the estimated trade-offs for each of the seven separate regions. The first point to note is that all of the coefficients are statistically significant (at the 90% confidence level). When examining the WTP and the HTO estimates there appears to be no obvious pattern in the results. None of the regions are significantly different from one another. The largest HTO for any of the attributes is for black fallout in Lower Hamilton where the estimated trade-off was 52% that of the value of the health attribute.

Given the particular geography of the regions it was decided to create three larger grouping out of the seven regions. One purpose of this exercise was to create regions with a sufficient number of observations to allow the estimation of the full quadratic model. Based on both the physical location of the seven regions and the survey results (both with respect to experience with the attributes and with the trade-off estimates) the following groupings were undertaken. Region 1 was comprised of Lower Hamilton and Stoney Creek. These two areas are located adjacent to each other and are comprised of a similar urban setting. They are also located closest to the industrial areas of the Hamilton-Wentworth Region and those areas that are generally believed to suffer from the worst air quality. Region 2 was comprised of Upper Hamilton, Dundas, and Ancaster. Once again these areas are located next to each other and are comprised of similar urban settings. Glanbrook and Flamborough, being the most rural of all the regions, were grouped together as Region 3.

Tables 7.12 and 7.13 present both the regression and trade-off results for these regions using the linear, quadratic tax model, and the full quadratic, respectively. Examining the results of the quadratic model estimation, the most noticeable feature is that the health trade-offs are highest in Region 1 (Lower Hamilton and Stoney Creek). Given the earlier reported survey results with respect to the experience with the air quality attributes, this result is not surprising. Recall from Table 6.5 that respondents from those areas reported to have the highest percentage of experience with bad odour and black fallout of all the regions. As well, their level of concern for those attributes was amongst the highest of all the regions (see Table 6.7). Furthermore, those attributes were ranked as most important by a higher percentage of respondents in those regions. Thus, these results are consistent with those results found in other sections of the survey.

7.8. Validity re-examined

Given that respondents were now grouped in the above regions it was decided to re-investigate the issue of construct validity with respect to income. The idea was that if all of the respondents had similar experiences and trade-offs in general. then perhaps there would be a greater variation in WTP between the high and low income groups. Only Regions 1 and 2 were examined, as only 22 of the 25 respondents in Region 3 reported their income status, and of those 22 only

4 belonged to the low income group. The full quadratic model was used employing both first and second choices. The trade-off results are presented in Table 7.14. Of note is that in Region 1, those in the high income group had WTP values that were more than 25% greater than those in the low income group for each of the attributes. This finding is relevant, particularly given that the HTO values are quite similar between the two groups. The findings with respect to WTP in Region 2 are however, not similar. WTP values were higher for only two of the attributes in the high income group. Hence, it is difficult to establish the validity of the results and the reason for such a finding may be explained by other differences between the groups as discussed previously.

7.8. Conclusions

The results demonstrate the need to control the quality of data collected in this type of conjoint analysis survey. The responses indicate that the ranking task was difficult. A number of respondents did not attempt the exercise at all. Others failed to follow the instructions in ranking the choices or chose alternatives that were dominated by others in the choice set. Excluding the respondents who committed dominance violations substantially improved the precision of the willingness-to-pay estimates. Excluding these respondents also increased the size and significance of the tax coefficients and led to somewhat lower point estimates of willingness-to-pay. After trying to adjust for those variables that predicted whether or not respondents committed violations these observations persisted and indeed the stratified samples did yield statistically different estimates of the regression coefficients.

The conclusions about exploitation of multiple rankings were mixed. It is clear that the precision of the estimates is improved by considering first and second rankings together, however, the critically important tax terms tend towards zero and lose significance under this procedure, inflating the WTP estimates. Asking for a ranking of four choice sets may also increase the cognitive difficulty of answering the questionnaire, leading to a reduced and non-random sample which may lead to biassed results. Investigating the trade-off here would be useful. It may be that there are alternative methods for exploiting the ranking data.

The conclusions with respect to the health trade-offs were more encouraging. While the WTP estimates appeared to be sensitive to the model employed and the data employed, the HTO estimates were much more stable. The health trade-off was stable across (a) model specification, (b) data set employed, and (c) number of rankings employed. Hence, one can have much more confidence with the health trade-off estimates than with the WTP estimates.

The results of the income validity test in the stratified regions was a further positive sign that there was a degree of validity to the overall results obtained. It was also encouraging to find that the estimated willingness-to-pay values were broadly consistent with what the average taxpayer pays for regional services. Respondents were willing-to-pay about half the amount of the allocation of property taxes that are used for education for a decrease in health risk. The other estimated willingness-to-pay values fell in line with what was being spent on services such as police, roadways, and waste management.

The estimates of willingness-to-pay were quite substantial, but they are consistent with other studies. Loehman *et al.* (1994) estimated WTP for increases in either health or visibility to be about US$13 per month, apparently at 1980 prices[27]. The health effects were not as explicitly defined as they were in the present study, nor was the actual value of the benefits. In an earlier study, Loehman and De (1982) estimated the yearly WTP for a one-day per year reduction in severe cough, severe shortness of breath, and minor eye irritation to be between US$7 and US$46, depending on the income and other factors[28]. Hall *et al.*(1992) combined the results from several contingent valuation as well as other willingness-to-pay studies to obtain a willingness-to-pay of US$23 (at 1990 prices) per day for a one-day-per-year reduction in minor-restricted-activity-days. Note that the restriction in these studies, to include only personal benefits, excludes the public good aspect of air quality which may be captured in the estimates obtained in the current study (see Chapter 5). Note that this was a WTP for a one-day reduction, as opposed to a more substantial reduction, and therefore would be expected to be smaller.

Overall, it appears that adopting the conjoint analysis methodology to value the attributes of air quality clearly permits much more detailed valuation of the attributes of environmental services than was previously possible. It has been demonstrated that a complex conjoint analysis experiment can be administered by

---

[27]The survey was conducted in 1980. It is assume that the values reflect that year's price levels.

[28]It is unclear in which year the survey was conducted.

mail to a representative sample of an urban population, although it is suspected we are close to the boundary of what can reasonably be accomplished in this manner. More expensive personal interview techniques may yield gains in comprehension, but the sheer magnitude of the rating task will require well motivated respondents.

Figure 7.1. Breakdown of medium and large data sets

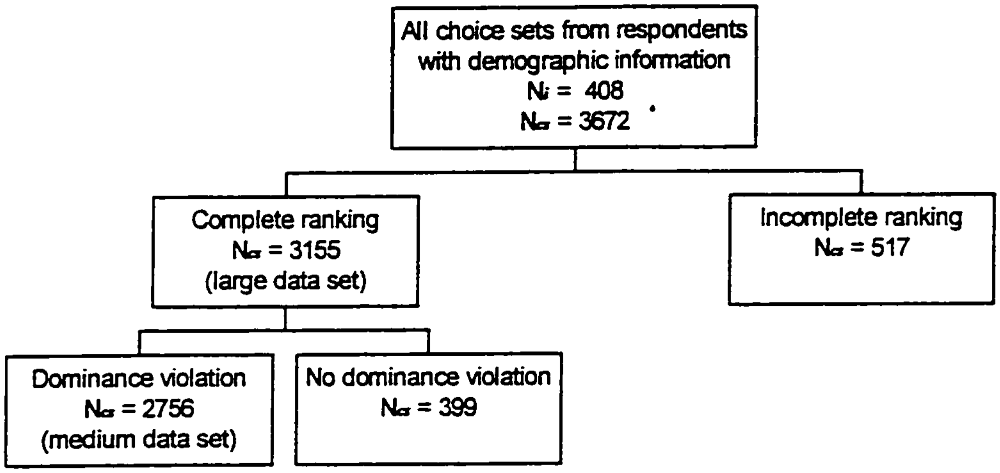


Table 7.1. Estimated coefficients by model
[first choices only, small data set]
(Asymptotic standard errors in parentheses)

	Model 1: Linear	Model 2: Linear with quadratic taxes	Model 3: Quadratic	Model 4: Quadratic with demographics
respondents	$N_i = 188$	$N_i = 188$	$N_i = 188$	$N_i = 148$
choice sets	$N_{CS} = 1692$	$N_{CS} = 1692$	$N_{CS} = 1692$	$N_{CS} = 1332$
pseudo R^2	.599	0.6041	0.6202	0.614
Log likelihood	-940.1	-928.7	-891.0	-712.6
H (health)	2.379 (.071)	2.278 (0.073)	1.927 (.136)	1.522 (.411)
O (odour)	.697 (.056)	.760 (.060)	.590 (.141)	1.113 (.387)
F (black fallout)	.886 (.067)	.894 (.070)	.638 (.148)	.593 (.448)
V (visibility)	.475 (.059)	.560 (.064)	.660 (.149)	.219 (.387)
T (tax)	-.032 (.004)	-.039 (005)	-.048 (.011)	-.058 (.029)
T^2		-.0005 (.0001)	-.0009 (.0003)	-.0009 (.0003)
H^2			1.214 (.231)	1.252 (.259)
H·O			.281 (.220)	.213 (.245)
H·F			.428 (.235)	.518 (.253)
H·V			.101 (.185)	.102 (.200)
H·T			-.014 (.012)	-.025 (.014)
O^2			1.229 (.454)	1.178 (.510)
O·F			.040 (.264)	-.012 (.292)
O·V			2.025 (.517)	2.002 (.576)
O·T			.008 (.010)	.006 (.011)
F^2			-1.84 (.496)	-1.771 (.555)
F·V			-1.081 (.497)	-1.022 (.553)
F·T			.024 (.014)	.016 (.014)
V^2			-.192 (.269)	-.181 (.295)
V·T			.016 (.013)	.021 (.014)
I·H				.002 (.003)
I·O				-.002 (.002)
I·F				.001 (.003)
I·V				.001 (.002)
I·T				-.0003 (.0002)
E·H				-.393 (.191)
E·O				-.208 (.166)
E·F				.033 (.209)
E·V				.003 (.176)
E·T				.005 (.013)
A·H				.006 (.007)
A·O				-.005 (.006)
A·F				-.003 (.007)
A·V				.008 (.006)
A·T				.0004 (.0005)
C·H				.319 (.148)
C·O				-.002 (.132)
C·F				.042 (.163)
C·V				-.0333 (.140)
C·T				.006 (.010)

Table 7.2. WTP for a one-third improvement in attributes, by model, in dollars per month
[first choices only, small data set]
(Asymptotic standard errors in parentheses)

	Model 1: Linear	Model 2: Linear with Quadratic taxes	Model 3: Full Quadratic	Model 4: Quadratic with demographics
health	74.5	58.1	40.2	39.7
	(8.0)	(6.8)	(9.2)	(10.3)
odour	21.8	19.3	12.3	13.1
	(3.1)	(2.6)	(4.7)	(5.5)
bfo	27.7	22.8	13.3	12.8
	(2.5)	(2.5)	(4.2)	(4.7)
visibility	14.9	14.2	13.8	14.4
	(2.1)	(1.9)	(4.7)	(5.2)

Table 7.3. Health trade-off, by model
[first choices only, small data set]
(asymptotic standard errors in parentheses)

	Model 1: Linear	Model 2: Linear with Quadratic taxes	Model 3: Full Quadratic	Model 4: Quadratic with demographics
odour	.29	.33	.31	.33
	(.024)	(.028)	(.067)	(.076)
bfo	.37	.39	.33	.32
	(.026)	(.029)	(.067)	(.074)
visibility	.37	.25	.34	.36
	(.026)	(.029)	(.076)	(.083)
taxes	-.013	-.017	-.025	-.025
	(.001)	(.002)	(.006)	(.007)

Table 7.4. Estimated coefficients by data set
[Model 4-quadratic model with demographics, first choices only]
(asymptotic standard errors in parentheses)

	Large Data Set		Medium Data Set		Small Data Set	
respondents	$N_i = 363$		$N_i = 361$		$N_i = 148$	
choice sets	$N_{CS} = 3155$		$N_{CS} = 2756$		$N_{CS} = 1332$	
pseudo R^2	.4377		.4947		.6141	
log Likelihood	-2459.3		-1930.4		-712.6	
H (health)	.685	(.166)	.742	(.1990)	1.522	(.411)
O (odour)	.614	(.168)	.423	(.188)	1.113	(.387)
F (black fallout)	.315	(.183)	.306	(.209)	.593	(.448)
V (visibility)	.159	(.166)	.292	(.197)	.219	(.387)
T (tax)	-.020	(.009)	-.038	(.012)	-.058	(.029)
T^2	-.00006	(.00007)	.0002	(.00009)	-.0009	(.0003)
H^2	.538	(.099)	1.130	(.138)	1.252	(.259)
H·O	-.075	(.087)	.419	(.124)	.213	(.245)
H·F	.115	(.091)	.530	(.128)	.518	(.253)
H·V·	-.042	(.076)	.224	(.100)	.102	(.200)
H·T	.004	(.005)	-.014	(.006)	-.025	(.014)
O^2	1.311	(.219)	1.118	(.292)	1.178	(.510)
O·F	-.072	(.122)	-.120	(.162)	-.012	(.292)
O·V	1.064	(.210)	1.540	(.312)	2.002	(.576)
O·T	.004	(.003)	-.006	(.004)	.006	(.011)
F^2	-1.08	(.214)	-1.201	(.295)	-1.771	(.555)
F·V	-1.212	(.228)	-.789	(.311)	-1.022	(.553)
F·T	.0007	(.004)	.001	(.005)	.016	(.014)
V^2	.029	(.108)	.097	(.162)	-.181	(.295)
V·T	-.006	(.003)	.009	(.005)	.021	(.014)
I·H	.010	(.00007)	.008	(.001)	.002	(.003)
I·O	.00004	(.001)	-.002	(.001)	-.002	(.002)
I·F	.004	(.001)	.003	(.001)	.001	(.003)
I·V	.003	(.001)	.002	(.001)	.001	(.002)
I·T	-.0001	(.00007)	-.00007	(.00008)	-.0003	(.0002)
E·H	-.134	(.081)	-.128	(.097)	-.393	(.191)
E·O	.034	(.081)	.084	(.088)	-.208	(.166)
E·F	-.075	(.088)	-.081	(.101)	.033	(.209)
E·V	-.069	(.080)	-.050	(.095)	.003	(.176)
E·T	.012	(.005)	.013	(.006)	.005	(.013)
A·H	.003	(.002)	.004	(.003)	.006	(.007)
A·O	-.0008	(.002)	-.0003	(.003)	-.005	(.006)
A·F	.0008	(.003)	.001	(.003)	-.003	(.007)
A·V	.004	(.002)	.005	(.003)	.008	(.006)
A·T	.0003	(0001)	.0003	(.0002)	.0004	(.0005)
C·H	.195	(.068)	.150	(.082)	.319	(.148)
C·O	.024	(.068)	.051	(.075)	-.002	(.132)
C·F	.017	(.074)	-.053	(.086)	.042	(.163)
C·V	-.016	(.068)	-.052	(.081)	-.0333	(.140)
C·T	-.009	(.004)	.003	(.005)	.006	(.010)

Table 7.5. WTP and HTO estimates by data set
[Model 4: full quadratic plus demographics, first choices only]
(asymptotic standard errors in parentheses)

	WTP			HTO		
	small data set	medium data set	large data set	small data set	medium data set	large data set
H	39.7 (10.3)	66.5 (13.5)	159.3 (50.9)	—		
O	13.1 (5.5)	18.5 (6.0)	67.0 (23.3)	.33 (.08)	.278 (.050)	.421 (.038)
F	12.8 (4.7)	22.5 (5.7)	57.7 (19.1)	.32 (.07)	.339 (.047)	.363 (.040)
V	14.4 (5.2)	27.9 (6.8)	52.1 (19.3)	.36 (.08)	.420 (.054)	.327 (.042)
T	—			0025 (.007)	.015 (.003)	.006 (.002)

Table 7.6. Estimated coefficients by number of choices employed
[Model 4: quadratic model with demographics, small data set]
(standard errors in parentheses)

	1st choices		1st and 2nd choices		1st, 2nd, and 3rd choices	
respondents	N_i =148		N_i =148		N_i =148	
choice sets	N_{CS} =1332		N_{CS} =2664*		N_{CS} =3996*	
pseudo R^2	.6141		.5063		.4455	
log Likelihood	-712.61		-1631.1		-2347.18	
H (health)	1.522	(.411)	1.971	(.324)	1.522	(.411)
O (odour)	1.113	(.387)	.853	(.261)	1.113	(.387)
F (black fallout)	.593	(.448)	795	(.310)	.593	(.448)
V (visibility)	.219	(.387)	.308	(.278)	.219	(.387)
T (tax)	-.058	(.029)	-.04	(.015)	-.058	(.029)
T^2	-.0009	(.0003)	-.0006	(.0001)	-.0009	(.0003)
H^2	1.252	(.259)	.249	(.121)	1.252	(.259)
H·O	.213	(.245)	.121	(.114)	.213	(.245)
H·F	.518	(.253)	.335	(.124)	.518	(.253)
H·V	.102	(.200)	-.216	(.107)	.102	(.200)
H·T	-.025	(.014)	-.013	(.007)	-.025	(.014)
O^2	1.178	(.510)	-.116	(.374)	1.178	(.510)
O·F	-.012	(.292)	.074	(.178)	-.012	(.292)
.O·V	2.002	(.576)	-.065	(.314)	2.002	(.576)
O·T	.006	(.011)	.006	(.004)	.006	(.011)
F^2	-1.771	(.555)	-.300	(.341)	-1.771	(.555)
F·V	-1.022	(.553)	-.076	(.399)	-1.022	(.553)
F·T	.016	(.014)	.012	(.006)	.016	(.014)
V^2	-.181	(.295)	-.486	(.158)	-.181	(.295)
V·T	.021	(.014)	.024	(.006)	.021	(.014)
I·H	.002	(.003)	.002	(.002)	.002	(.003)
I·O	-.002	(.002)	.002	(.002)	-.002	(.002)
I·F	.001	(.003)	.0004	(.002)	.001	(.003)
I·V	.001	(.002)	-.002	(.002)	.001	(.002)
I·T	-.0003	(.0002)	-.00001	(.00009)	-.0003	(.0002)
E·H	-.393	(.191)	-.255	(.151)	-.393	(.191)
E·O	-.208	(.166)	-.224	(.118)	-.208	(.166)
E·F	.033	(.209)	-.044	(.145)	.033	(.209)
E·V	.003	(.176)	.126	(.128)	.003	(.176)
E·T	.005	(.013)	-.004	(.007)	.005	(.013)
A·H	.006	(.007)	.001	(.005)	.006	(.007)
A·O	-.005	(.006)	-.006	(.004)	-.005	(.006)
A·F	-.003	(.007)	-.003	(.005)	-.003	(.007)
A·V	.008	(.006)	.009	(.005)	.008	(.006)
A·T	.0004	(.0005)	.00001	(.00002)	.0004	(.0005)
C·H	.319	(.148)	.382	(.12)	.319	(.148)
C·O	-.002	(.132)	.117	(.096)	-.002	(.132)
C·F	.042	(.163)	.088	(.116)	.042	(.163)
C·V	-.0333	(.140)	.098	(.103)	-.0333	(.140)
C·T	.006	(.010)	.002	(.005)	.006	(.010)

* Using 1st and 2nd choices there were 1332 four-alternative choice set and 1332 3-alternative choice sets. Using 1st, 2nd and 3rd choice sets, there were an additional 1332 two-alternative choice sets.

Table 7.7. WTP and HTO estimates by number of choices employed, [Model 4: full quadratic plus demographics, small data set] (asymptotic standard errors in parentheses)

	WTP			HTO		
	1st choices	1st and 2nd choices	1st, 2nd and 3rd choices	1st choices	1st and 2nd choices	1st, 2nd and 3rd choices
H	39.7 (10.3)	51.1 (5.7)	83.9 (9.4)	—	—	—
O	13.1 (5.5)	13.0 (2.7)	24.2 (3.4)	.33 (.08)	.25 (.03)	.29 (.02)
F	12.8 (4.7)	15.3 (2.4)	22.2 (2.9)	.32 (.07)	.30 (.04)	.26 (.03)
V	14.4 (5.2)	14.5 (2.6)	23.0 (3.3)	.36 (.08)	.28 (.04)	.27 (.03)
T	—	—	—	-0025 (.007)	-.02 (.002)	.01 (.001)

Table 7.8. Trade-offs by income level [Model 3: Full quadratic, small data set] (asymptotic standard errors in parentheses)

	1st choices[§]				1st and 2nd choices			
	WTP		HTO		WTP		HTO	
	high*	low*	high	low	high	low	high	low
N_i	92	63	92	63	92	63	92	63
log likelihood	-439	-282	-453	-282	-997	-698	-997	-698
pseudo -R^2	0.607	0.632	0.606	0.641	0.515	0.505	0.515	0.505
health	33.6 (9.5)	35.6 (14.5)	-	-	48.6 (6.3)	53.3 (10.2)	-	-
odour	7.6 (4.6)	15.4 (9.3)	.248 (.0960	.432 (.120)	12.1 (3.0)	13.7 (4.6)	.250 (.042)	.257 (.056)
bfo	11.7 (4.9)	12.4 (7.0)	.360 (.095)	.349 (.114)	13.9 (2.7)	16.5 (4.3)	.287 (.043)	.309 (.057)
visibility	13.2 (5.5)	6.8 (5.3)	.420 (.107)	.189 (.125)	17.0 (3.2)	17.9 (4.8)	.350 (.044)	.335 (.058)

*Low income: those with household income below $50,000
High income: those with household income of $50,000 or greater
[§]note that f2 was dropped - see section 7.6 for details
Note: only includes those respondents who stated their household income

Table 7.9. Trade-offs by cardio-respiratory experience
[Model 3: Full quadratic, small data set]
(asymptotic standard errors in parentheses)

	1st choices				1st and 2nd choices			
	WTP		HTO		WTP		HTO	
	yes	no	yes	no	yes	no	yes	no
N_i	109	79	109	79	109	79	109	79
log likelihood	-474	-408	-474	-408	-1124	-909	-1124	-909
pseudo -R^2	0.652	0.586	0.652	0.586	0.539	0.485	0.539	0.485
health	52.7 (20.0)	26.5 (7.0)	-	-	54.1 (7.0)	57.1 (9.3)	-	-
odour	14.6 (8.9)	8.1 (4.1)	.277 (.095)	.308 (.099)	13.8 (3.2)	16.4 (4.3)	.256 (.038)	.287 (.047)
bfo	14.9 (7.6)	9.9 (3.9)	.283 (.095)	.375 (.100)	16.0 (2.9)	16.9 (3.7)	.296 (.038)	.300 (.049)
visibility	17.2 (9.6)	10.2 (4.0)	.326 (.106)	.386 (.112)	17.2 (3.2)	21.5 (4.8)	.318 (.040)	.377 (.050)

Table 7.10. Trade-offs by black fallout experience
[Model 3: Full quadratic, small data set]
(asymptotic standard errors in parentheses)

	1st choices				1st and 2nd choices			
	WTP		HTO		WTP		HTO	
	yes*	no*	yes*	no*	yes*	no*	yes*	no*
N_i	125	60	125	60	125	60	125	60
log likelihood	-613.1	-258	-613	-258	-1391	-616	-1391	-616
pseudo -R^2	0.607	0.655	0.607	0.655	0.5023	0.541	0.5023	0.541
health	41.0 (12.0)	32.8 (10.1)	-	-	57.1 (7.3)	49.9 (8.3)	-	-
odour	14.6 (6.7)	6.7 (4.5)	.355 (.087)	.205 (.098)	17.7 (3.5)	9.3 (3.5)	.311 (.036)	.188 (.051)
bfo	12.5 (5.3)	14.3 (5.8)	.304 (.085)	.437 (.113)	18.8 (3.1)	11.26 (3.0)	.330 (.038)	.226 (.049)
visibility	12.9 (5.9)	12.8 (5.9)	.314 (.095)	.391 (.127)	17.6 (3.4)	22.3 (4.6)	.308 (.039)	.445 (.053)

*yes/no refer to experience with the attribute
Note: only includes those respondents who stated their bfo experience

Table 7.11. Estimated coefficients and trade-off values by 7 regions [Model 2: linear with quadratic tax term,, 1[st] and 2[nd] choices] (standard errors in parentheses)

	Hamilton Upper	Hamilton Lower	Stoney Creek	Ancaster	Dundas	Flam-borough	Glanbrook
respondents	N_i=65	N_i=56	N_i=17	N_i=4	N_i=11	N_i=19	N_i=6
choice sets	N_{cs}=1170	N_{cs}=1008	N_{cs}=306	N_{cs}=252	N_{cs}=198	N_{cs}=342	N_{cs}=108
pseudo R^2	0.478	0.486	0.56	0.506	0.587	0.493	0.63
log Likelihood	-759	-643.6	-167.4	-154.6	-101.6	-215.5	-49.5
H (health)	2.086 (.091)	2.040 (.098)	2.493 (.214)	2.344 (.221)	2.839 (.311)	1.889 (.158)	2.862 (.437)
O (odour)	.706 (.069)	.717 (.075)	.977 (.153)	.845 (.153)	1.014 (.206)	.862 (.131)	.822 (.280)
F (black fallout)	.778 (.079)	1.063 (.093)	1.119 (.188)	.711 (.172)	1.270 (.256)	.787 (.150)	1.444 (.361)
V (visibility)	.563 (.072)	.579 (.079)	.796 (.170)	.519 (.166)	.374 (.202)	.745 (.143)	.851 (.331)
T (tax)	-.031 (.004)	-.037 (.005)	-.048 (.009)	-.046 (.009)	-.046 (.011)	-.057 (.010)	-.072 (.022)
T^2	-.0006 (.0001)	-.0008 (.0001)	-.0007 (.0002)	-.0004 (.0002)	-.0006 (.0003)	-.001 (.0002)	-.001 (.0005)
WTP_H	67.5 (8.2)	54.7 (6.6)	51.9 (9.2)	50.8 (9.0)	62.3 (14.7)	33.4 (5.7)	39.6 (11.3)
WTP_O	22.8 (3.6)	19.2 (3.0)	20.3 (4.7)	18.3 (4.4)	22.2 (6.9)	15.2 (3.3)	11.4 (4.8)
WTP_F	25.2 (3.0)	28.5 (3.2)	23.3 (4.0)	15.4 (3.5)	27.7 (6.2)	13.9 (2.6)	20.0 (5.3)
WTP_V	18.2 (2.8)	15.5 (2.5)	16.6 (3.8)	11.2 (3.7)	8.2 (4.6)	13.1 (2.8)	11.8 (4.8)
HTO_O	.338 (.034)	.352 (.038)	.392 (.063)	.360 (.064)	.357 (.070)	.456 (.072)	.287 (.100)
HTO_F	.373 (.035)	.521 (.040)	.449 (.066)	.303 (.068)	.447 (.074)	.416 (.074)	.504 (.104)
HTO_V	.270 (.035)	.283 (.002)	.319 (.071)	.221 (.073)	.132 (.073)	.393 (.077)	.297 (.116)

Table 7.12. Estimated coefficients and trade-off values by 3 regions, [Model 2: linear model with quadratic tax term, 1st and 2nd choices] (standard errors in parentheses)

	Region 1*	Region 2*	Region 3*
respondents	N_i=73	N_i =90	N_i =25
choice sets	N_{cs}=1314	N_{cs} = 1620	N_{cs} = 450
pseudo R^2	0.501	0.492	0.519
log Likelihood	-815	-1023.5	-269
H (health)	2.128	2.181	2.044
	(.089)	(.081)	(.148)
O (odour)	.771	.748	.851
	(.068)	(.060)	(.117)
F (black fallout)	1.070	.803	.899
	(.083)	(.068)	(.138)
V (visibility)	.622	.543	.759
	(.072)	(.062)	(.130)
T (tax)	-.039	-.035	-.059
	(.004)	(.003)	(.009)
T^2	.0008	-.0005	-.001
	(.0001)	(.00008)	(.0002)
WTP_H	54.1	63.04	34.5
	(5.4)	(5.9)	(5.0)
WTP_O	19.6	21.6	14.4
	(2.6)	(2.7)	(2.7)
WTP_F	27.2	23.2	15.2
	(2.6)	(2.2)	(2.3)
WTP_V	15.8	15.7	12.8
	(2.1)	(2.1)	(2.4)
HTO_O	.362	.343	.417
	(.033)	(.028)	(.059)
HTO_F	.503	.368	.440
	(.034)	(.029)	(.062)
HTO_V	.292	.249	.371
	(.035)	(.029)	(.004)

*Region 1:Lower Hamilton and Stoney Creek
Region 2: Upper Hamilton, Dundas and Ancaster
Region 3: Flamborough and Glanbrook

Table 7.13. Estimated coefficients and trade-off values by 3 regions [Model 3: quadratic, 1[st] and 2[nd] choices] (standard errors in parentheses)

	Region 1*	Region 2*	Region 3*
respondents	N_i=73	N_i =90	N_i =25
choice sets	N_{cs}=1314	N_{cs} = 1620	N_{cs} = 450
pseudo R^2	.517	.515	.545
log Likelihood	-787.9	-976.6	-254.4
H (health)	2.238 (.135)	2.407 (.133)	2.453 (.262)
O (odour)	.690 (.113)	.527 (.108)	.727 (.221)
F (black fallout)	.856 (.122)	.506 (.111)	.666 (.230)
V (visibility)	.690 (.121)	.818 (.127)	1.155 (.258)
T (tax)	-.039 (.006)	-.038 (.006)	-.087 (.018)
T^2	-.0006 (.0002)	-.0005 (.0001)	-.001 (.0004)
H^2	.060 (.174)	.302 (.163)	.127 (.317)
H·O	-.124 (.165)	.291 (.153)	.026 (.293)
H·F	.233 (.178)	.457 (.170)	.176 (.324)
H·V	-.113 (.150)	-.374 (.151)	-.537 (.290)
H·T	-.001 (.010)	-.010 (.008)	-.017 (.020)
O^2	.350 (.580)	-.189 (.488)	-.172 (.966)
O·F	.561 (.282)	-.392 (.223)	-.125 (.431)
O·V	.006 (.489)	.193 (.418)	-.083 (.839)
O·T	.010 (.006)	.002 (.005)	.016 (0013)
F^2	-.321 (.522)	-.632 (.454)	-.585 (.902)
F·V	-.630 (.615)	.128 (.523)	.027 (1.026)
F·T	.011 (.008)	.014 (.007)	.007 (.017)
V^2	-.809 (.252)	-.217 (.196)	-.511 (.395)
V·T	.009 (.009)	.025 (.008)	.057 (.021)
WTP_H	57.4 (9.6)	63.9 (10.1)	28.3 (5.8)
WTP_O	17.7 (4.5)	14.0 (4.0)	8.4 (3.4)
WTP_F	22.0 (4.4)	13.4 (3.3)	7.7 (3.1)
WTP_V	17.7 (4.3)	21.7 (4.8)	13.3 (3.5)
HTO_O	.308 (.047)	.219 (.042)	.296 (.086)
HTO_F	.383 (.049)	.210 (..044)	.271 (.088)
HTO_V	.308 (.050)	.340 (.046)	.035 (.007)

*Region 1:Lower Hamilton and Stoney Creek
Region 2: Upper Hamilton, Dundas and Ancaster
Region 3: Flamborough and Glanbrook

Table 7.14. Income validity, by region employed
[Model 3: Full quadratic, small data set]
(asymptotic standard errors in parentheses)

	Region 1				Region 2			
	WTP		HTO		WTP		HTO	
Income Group	high*†	low*	high	low	high	low	high	low
N_i	25	35	25	35	48	24	48	24
log likelihood	-238	-412	-250	-412	-516	-238	-516	-238
pseudo -R^2	0.559	0.474	0.553	0.474	0.519	0.556	0.519	0.556
health	63.4 (16.0)	54.4 (14.3)	–	–	53.6 (10.7)	55.0 (16.4)	–	–
odour	19.3 (7.3)	17.5 (7.2)	.304 (.067)	.322 (.080)	9.4 (4.5)	7.7 (6.0)	.175 (.062)	.140 (.084)
bfo	23.3 (7.3)	19.9 (6.6)	.368 (.073)	.366 (.081)	11.4 (3.9)	10.1 (5.3)	.213 (.064)	.183 (.083)
visibility	21.8 (7.5)	17.2 (6.7)	.343 (.073)	.316 (.081)	16.6 (5.2)	22.1 (9.0)	.310 (.065)	.401 (.097)

*Low income: those with household income below $50,000
High income: those with household income of $50,000 or greater
†f2 was not included for Region 1 - see section 7.8 for details
Note: only includes those respondents who stated their household income

Chapter 8

CONCLUSIONS

The first objective of this thesis was to investigate the current state of research of non-market valuation of health care interventions. In the review of health care contingent valuation studies, one of the main findings was the lack of consistent and transparent reporting of methods. Contingent valuation studies must be more consistent in providing the methods employed and the theoretical implications of the methods. Indeed it was found that there is still much debate with respect to methods employed and it is thus imperative that authors of contingent valuation studies must provide these details in order for their work to be accurately judged and interpreted.

The NOAA Panel provided some useful guidelines, but it must be noted that the Panel was assessing a particular application of the contingent valuation methodology - natural resource damage assessment. As has already been noted, these guidelines are having an impact on the application of contingent valuation studies. However, it is important that the guidelines are critically evaluated before they are accepted as the standard in the application of the contingent valuation method in health and environment cost-benefit analysis studies. This has been done for health care cost-benefit analysis studies (Diener *et al.*, 1998).

The literature survey exemplified a shortcoming of the contingent valuation method - namely that it may not be the appropriate method to value interventions or programs with multiple outcomes. An alternative to contingent valuation that is particularly applicable in cases of multi-attribute outcomes is conjoint analysis. In a conjoint analysis study, attributes can be simultaneously valued and trade-offs between attributes can be obtained. This can be particularly useful when evaluating public programs in a collectively funded system with finite resources.

The second objective of the thesis was to investigate the potential for using conjoint analysis for valuing health and non-health benefits of environmental improvements. Hence, a case study was conducted in which these techniques were applied to the valuation of air quality changes in the Regional Municipality of Hamilton-Wentworth. The specific objectives of the case study were (i) to obtain as accurate information as possible about health trade-offs and willingness-to-pay for the attributes of air quality improvements in Hamilton-Wentworth, taking into account the limited budget for the study, and (ii) to contribute to the methodology of the conjoint analysis method by gathering evidence on the nature and effect of cognitive difficulties and of exploring the use of ranking information.

The attributes of air quality that were employed in the survey were health effects (measured as the number of hospital admissions and increased mortality due to air pollution), number of days of bad odour, number of days of black fallout and number of days of poor visibility. All of the attributes were measured on a monthly basis. Respondents were required to rank 9 choice sets, each consisting of 4 alternatives. Each alternative was defined using different combinations of the levels of each alternative. Obtaining rankings allowed more information to be garnered from each respondent than in a standard discrete choice study.

The overall results of the case study were encouraging. The valuations obtained did appear to be credible given that they were comparable with what was spent on local services with property tax dollars and that they were consistent with results from other studies. An important finding was that the willingness-to-pay estimates were very sensitive to which model and which data set was employed. The trade-offs among the other attributes were much more stable and showed that using taxes as an attribute may provide some uncertainties with respect to responses.

Using both the first and second choices of those respondents that provided a complete ranking and committed no dominance violations, the willingness-to-pay estimates were approximately $50 for a one-third improvement in the health attribute and between $13 and $15 for improvements in the other three attributes. Each of the other attributes was valued between 25% and 30% of the health attribute. However, given the small sample size, the willingness-to-pay estimates should be interpreted cautiously and should be considered as illustrative. The health trade-offs may be more reliable from a policy perspective as they did clearly indicate that the health attribute was the most important one.

The responses did indicate that the ranking task was difficult. Out of 515 returned surveys, only 115 respondents were able to rank all 9 choice sets completely without making a dominance violation. It is unclear whether this finding was attributable to the number of choice sets included or due to the fact that the task itself was too difficult. It would have been of interest to interview the respondents after they had completed the survey in order to enquire about these issues. Future research should include such an analysis. As well, it may be too difficult for respondents to complete a complicated conjoint analysis questionnaire by mail. In-person interviewing may have decreased the number of choice sets with incomplete rankings or dominance violations. Clearly these were weaknesses of the study and future research should address these concerns. It should be noted that other studies that do not include checks for dominance violations may include responses from individuals that did not understand their task. In the absence of checks for dominance violations it is difficult to assess the degree of cognitive difficulties.

With respect to the exploitation of multiple rankings, the study showed that it may be best to incorporate some but not all of the ranking information. While incorporating first and second rankings appeared to increase the precision of the willingness-to-pay and health trade-off estimates, there did not appear to be benefits from also incorporating the third ranked choices. Incorporating multiple rankings requires a minimum of three alternatives in each choice set which may be a more cognitively difficult task for respondents than employing paired comparisons. Although multiple rankings allow one to obtain more data from each respondent, overall the survey results may be less reliable if the consequences are a smaller sample size. Caution should be used in future studies when considering both the number of alternatives in each set as well as the number of choice sets included in a survey. In addition, careful consideration should be given to the survey methods employed (i.e. in-person, mail, etc.).

The findings of this thesis point towards areas of future research in the field if the appropriate funding was available. In order to explore the reasons for the cognitive difficulties it would be worthwhile to conduct two parallel surveys - one by mail and one in-person. The results could be compared in order to ascertain whether in-person interviews yields better results with respect to cognitive understanding of the task. Furthermore, it would be useful to conduct post survey interviews in order to determine from the respondents themselves what specifically may have caused the difficulties. It may also be worthwhile to present respondents with smaller choice sets, perhaps paired comparisons, to gather whether that would affect the results. It would also be of interest to conduct an analogous contingent valuation study which would allow one to assess how respondents may value the attributes when they are presented separately from one another.

In summary, it appears that conjoint analysis ranked data can provide useful estimates of the willingness-to-pay for and trade-offs between attributes of air quality. One of the limitations of this study was the small sample of respondents that provided a complete ranking and committed no dominance violations. This should lead to cautious interpretation of the willingness-to-pay estimates that were obtained. While more research is necessary before policy makers can be more confident with the results, it is clear that the results from a conjoint analysis study provide a valuable starting point in any economic evaluation of public goods, particularly health and the environment.

REFERENCES

Adamowicz, W., J. Louviere, and M. Williams. 1994. Combining revealed and stated preference methods for valuing environmental commodities. *Journal of Environmental Economics and Management* 26:271-92.

Bates, J. 1988. Econometric issues in stated preference analysis. *Journal of Transport Economics and Policy* 22:59-69.

Beggs, S., S. Cardell, and J. Hausman,. 1981. Assessing the potential demand for electric cars. *Journal of Econometrics* 16:1-19.

Bishop, R. C., Champ, and P. A., Mullarkey, D. J. 1995. Contingent Valuation. In *The Handbook of Environmental Economics,* ed. Bromley, D. W., Cambridge: Blackwell.

Bockstael, N. E. 1995 Travel Cost Models. In *The Handbook of Environmental Economics,* ed. Bromley, D. W., Cambridge: Blackwell. .

Burnett, R. T., R. Dales, D. Krewski, R. Vincent, T. Dann, and J. R. Brook.1995. Associations between ambient particulate sulfate and admissions to Ontario hospitals for cardiac and respiratory diseases. *American Journal of Epidemiology* 142(1):15-22.

Carson, R. T., J. Wright, N. Carson, A. Alberini, and N. Flores. 1995. *A bibliography of contingent valuation studies and papers.* National Resource Damage Assessment.

Cattin, P., and D. Wittink. 1982. Commercial use of conjoint analysis: A survey. *Journal of Marketing* 46:44-53.

Chapman, R. G. 1984. An approach to estimating logit models of a single decision makers choice behavior. *Advances in Consumer Research* 11:656-661

Chapman, R. G., and R. Staelin. 1982. Exploiting rank ordered choice set data within the stochastic utility model. *Journal of Marketing Research* XIX:288-301.

City of Hamilton. 1996. *Current Estimates.*

Desvouges, W. H., F. R. Johnson, S. P. Hudson, A. R. Gable, and M. C. Ruby. 1996. *Using Conjoint Analysis and Health-State Classifications to Estimate the Value of Health Effects of Air Pollution: Pilot Test Results and Implications. Final Report.* North Carolina: Triangle Economic Research.

Dickson, L. L., J.D. Heale, and L. W. Chambers. 1995. *Fact Book on Health in Hamilton-Wentworth.* McMaster University: Health Priorities Analysis Unit, Faculty of Health Sciences.

Diamond, P. D. and J. A. Hausman. 1994. Contingent valuation; is some number better than no number? *Journal of Economic Perspectives* 8(4):45-66.

Diener, A., C. Donaldson, B. O'Brien, and A. Gafni. 1999. *Contingent valuation in health care: a caution against the uncritical adoption of the NOAA guidelines.* Manuscript. University of Nebraska Medical Center.

Diener, A., B. O'Brien, and A. Gafni. 1998. Health care contingent valuation studies: a review and classification of the literature. *Health Economics* 7(4):313-26.

Dillman, D. A.. 1978. *Mail and Telephone Surveys: The Total design Method,* New York: John Wiley and Sons.

Dockery, D. W., and C. A. Pope. 1994. Acute respiratory effects of particulate air pollution. *Annual Review of Public Health* 15:107-32.

Drummond, M. F., B. O'Brien, G.L. Stoddart, and G. Torrance. 1997. *Methods for the Economic Evaluation of Health Care Programmes.* New York: Oxford University Press.

Elliott, S. J., P. Kreuger, D. Cole, R. Hall, N. Voorberg,, S. Thorne, and S. Wakefield. 1997. *Perceptions of Air Pollution: The North Hamilton Survey. Final Report,* Prepared for the Hamilton-Wentworth Air Quality Initiative..

Freeman, A. M. 1995 Hedonic Pricing Methods. In Bromley DW, editor. *The Handbook of Environmental Economics.* Cambridge: Blackwell.

Goldstein, H. 1995 *Multilevel Statistical Analysis.* London: Kluwer Academic Publishing.

Green, P. E., and V. Srinivasan. 1978. Conjoint analysis in consumer research: issues and outlook. *Journal of Consumer Research* 5:103-23.

Greene, W. H. 1993. *Econometric Analysis*. New York: Macmillan Publishing Company .

Hall, J., A. Winer, M. Kleinman, F. Lurmann, V. Brajer, and S. Colome.1992. Valuing the health benefits of clean air. *Science* 255:812-7.

HAQI. 1997. *Hamilton-Wentworth Air Quality Initiative Summary Report*, October 1997.

Hanemann, W. M. 1994. Valuing the environment through contingent valuation. *Journal of Economic Perspectives* 8(4):19-44.

Hensher, A., and J. J. Louviere. 1983. Identifying individual prefernces for international air fares. *Journal of Transport Economics and Policy* 17:225-245.

Hicks, J. R. 1943. The four consumer's surpluses. *Review of Economic Studies* 8:108-15.

Hicks, J. R. 1941. The rehabilitation of consumers' surplus. *Review of Economic Studies* 9:108-16.

Hsaoi, C. 1986. *Analysis of Panel data*. Cambridge: Cambridge University Press.

Johansson, P. O . 1993. *Evaluating Health Risks: An Economic Approach*, Great Britain: Cambridge University Press.

Johnson, F. R., W. H. Desvouges, M. C. Ruby, D. Stieb, and P. De Civita.. 1998. Eliciting stated health preferences: An application to willingness-to-pay for longevity. *Medical Decision Making* 18 suppl:S57-S67.

Jones-Lee, M. W., M. Hammerton, and P. R. Phillips. 1985. The value of safety: results of a national sample survey. *Economic Journal* 95(95):49-72.

Kahneman, D., and J. Knetch. 1992. Valuing public goods: the purchase of moral satisfaction. *Journal of Environmental Economics and Management* 22:57-70.

Keeney, R. L., and H. Raiffa 1993. *Decisions with Multiple Objectives. Preferences and Value Trade-offs. Great Britain*: Cambridge University Press.

Kroes, E.P., and R. R. Sheldon. 1988. Stated preference methods. *Journal of Transport Economics and Policy* 22:11-25.

Lareau, T. J., and Rae. 1989. Valuing WTP for diesel odour reductions: An application of contingent ranking technique. *Southern Economic Journal* 55(3):728-42.

Layton, D. F. 1995. *Specifying and Testing Econometric Models for Stated Preference Surveys*. Dissertation, University of Washington.

Loehman, E., and D. De. 1982 Application of stochastic choice modelling to policy analysis of public goods: a case study of air quality improvements. *The Review of Economics and Statistics* 64:474-80.

Loehman, E., S. Park,and D. Boldt. 1994. Willingness-to-pay for gains and losses in visibility and health. *Land Economics* 70(4):478-98.

Louviere, J. J. 1988A. *Analyzing Decision Making: Metric Conjoint Analysis*. Newbury Park, Ca: Sage Publications Inc.

Louviere, J. J. 1988B. Conjoint analysis modelling of stated preferences. *Journal of Transport Economics and Policy* 22:93-119.

Louviere, J. J., and G. Woodworth. 1983. Design and analysis of simulated consumer choice or allocation experiments: An approach based on aggregate data. *Journal of Marketing Research* 350-67.

Luce, R., and P. Suppes. 1965. Preference, utility, and subjective probability. In *Handbook of Mathematical Psychology - Volume 3*. ed. Luce, R., Bush, R., Galanter, E. New York: John Wiley and Sons.

Magat, A. W., W. K. Viscusi, and J. Huber. 1988. Paired comparison and contingent valuation approaches to morbidity risk valuation. Journal of Environmental Economics and Management 15:395-411.

Marin, A., and G. Psacharopoulis. 1982. The reward for risk in the labour market: evidence from the UK and a reconciliation with other studies. Journal of Political Economy 21:1-16.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, ed. Zarembka, P. 105-42. New York: Academic Press..

McFadden, D. 1986. The choice theory approach to market research. *Marketing Science* 5(4):275-97.

Mishan, E. J. 1988. *Cost-benefit analysis*. New York City: Routledge.

Mitchell, R.C., and R. T. Carson. 1989. Using Surveys to Value Public Goods: *The Contingent Valuation Method.* Washington, D.C.: Resources for The Future.

Opaluch, J., S. Swallow, and T. Weaver. 1993. Evaluating impacts from noxious facilities: including public preferences in current siting mechanisms. *Journal of Environmental Economics and Management* 24:41-59.

Petersen, R. G. 1985. *Design and Analysis of Experiments*. New York: Dekker.

Portney, P. R. 1994. The contingent valuation debate: why economists should care. *Journal of Economic Perspectives* 8(4):3-18.

Propper, C. 1991. Contingent valuation of time spent on NHS waiting lists. *Economic Journal* 100:193-199.

Propper, C 1995. The disutility of time spent on UK National Health Service Waiting lists. *Journal of Human Resources* XXX (4):677-699

Rae, D. E. 1983. The value to visitors of improving visibility at Mesa Verde and Great Smoky National Parks. In *Managing Air Quality and Scenic Resources at National Parks,* ed. Rowe RD, Chestnut LG. Boulder, Co.: Westview Press.

Ryan, M., A. Scott, S. Farrar P. Shackley, S. Vick, and E. McIntosh. 1996. *Using conjoint analysis in health care: unresolved methodological issues*. HERU Discussion Paper No 2/96. Aberdeen: University of Aberdeen.

Ryan, M., and J. Hughes. 1997. Using conjoint analysis to assess women's preferences for miscarriage management. *Health Economics* 6(3):261-74.

Schulze, W., G. McClelland, M. Doane, E. Balistreri, R. Boyce, B. Hurd, and R. Simenauer. 1995. *An Analysis of Stated Preferences for Superfund Site Cleanup*. Manuscript, Cornell University.

Smith, V. K., and W. H. Desvouges. 1986. *Measuring Water Quality Benefits*. Boston: Kluwer Nijhoff Publishing.

The Regional Municipality of Hamilton-Wentworth. 1996. *1996 Budget*.

Torrance, G. W., D. Feeny, W. Furlong, R. D. Barr, Y. Zhang, and Q. Wang. 1996. Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. *Medical Care*. 34(7):702-22

Viscusi, W. K., A. W. Magat, and J. Huber. 1991. Pricing environmental health risks: survey assessments of risk-risk and risk-dollar trade- offs for chronic bronchitis. *Journal of Environmental Economics and Management* 21:32-51.

Appendix A.

The Questionnaire

Attitudes Towards Air Quality in Hamilton-Wentworth

A Survey

R.A. Muller
A.A. Diener
Department of Economics
McMaster University

February, 1997

Attitudes Towards Air Quality in Hamilton-Wentworth

R. Andrew Muller and Alan A. Diener
Department of Economics,
McMaster University, Hamilton, Ont.
Tel: (905) 525-9140, x22765
Fax: (905) 521-8232

This survey is being conducted by McMaster University on behalf of the Hamilton Air Quality Initiative, which is a project being undertaken by the Regional, Provincial and Federal governments together with representatives from a variety of other organizations and citizens-at-large. Thank you for answering the survey.

We will be asking how the quality of the air in Hamilton-Wentworth affects you. Your participation is voluntary. If you would rather not answer a question please skip it. Your answers will be treated in strict confidence.

There are no right or wrong answers to this survey. Please answer each question the way you feel, not the way you think you ought to feel.

Section A

Q1. How would you describe the air quality in your neighbourhood compared to the rest of Southern Ontario? (Circle the number of your answer)

1	2	3	4	5	6
Much better than the rest of Southern Ontario	Somewhat better than the rest of Southern Ontario	The same as the rest of Southern Ontario	Somewhat worse than the rest of Southern Ontario	Much worse than the rest of Southern Ontario	Don't Know

Q2. How concerned are you about annoying odours due to air pollution in Hamilton-Wentworth? (Circle number)

1	2	3	4	5	6
Not at all concerned	A little concerned	Somewhat concerned	Very concerned	Extremely concerned	Don't Know

Q3. How concerned are you about black sooty material (black fallout) falling from the air in Hamilton-Wentworth? (Circle number)

1	2	3	4	5	6
Not at all concerned	A little concerned	Somewhat concerned	Very concerned	Extremely concerned	Don't Know

Q4. How concerned are you about poor visibility in Hamilton-Wentworth? (Circle number)

1	2	3	4	5	6
Not at all concerned	A little concerned	Somewhat concerned	Very concerned	Extremely concerned	Don't Know

Q5. How concerned are you about the health effects of air pollution? (Circle number)

1	2	3	4	5	6
Not at all concerned	A little concerned	Somewhat concerned	Very concerned	Extremely concerned	Don't Know

Many issues concern people these days. Some people are concerned about poor air quality. Other people are more concerned about other issues. Please indicate whether you think each of the following issues is a more serious issue than air quality, a less serious issue than air quality, or about as serious an issue as air quality.

Q6. The level of crime: (circle the number of your answer)
1. MORE SERIOUS ISSUE THAN AIR QUALITY
2. LESS SERIOUS ISSUE THAN AIR QUALITY
3. EQUALLY AS SERIOUS AS AIR QUALITY

Q7. The level of taxes: (circle number)
1. MORE SERIOUS ISSUE THAN AIR QUALITY
2. LESS SERIOUS ISSUE THAN AIR QUALITY
3. EQUALLY AS SERIOUS AS AIR QUALITY

Q8. Clearing snow from and improving the paving of Regional roads: (circle number)
1. MORE SERIOUS ISSUE THAN AIR QUALITY
2. LESS SERIOUS ISSUE THAN AIR QUALITY
3. EQUALLY AS SERIOUS AS AIR QUALITY

Q9. The quality of the educational system: (circle number)
1. MORE SERIOUS ISSUE THAN AIR QUALITY
2. LESS SERIOUS ISSUE THAN AIR QUALITY
3. EQUALLY AS SERIOUS AS AIR QUALITY

Q10. The level of unemployment: (circle number)
1. MORE SERIOUS ISSUE THAN AIR QUALITY
2. LESS SERIOUS ISSUE THAN AIR QUALITY
3. EQUALLY AS SERIOUS AS AIR QUALITY

Q11. What do you consider to be the signs of good air quality?

____

____

____

Sometimes the air in Hamilton-Wentworth has a bad odour. People describe odours differently. Some odours are like rotten eggs, others are like coal, others are like acid. There may be other kinds of odour. Sometimes people go indoors to avoid the odour. Some areas of Hamilton-Wentworth are more affected than others.

Q12. Have you ever experienced bad odour problems, such as those described above, in Hamilton-Wentworth? (circle number)

1. YES
2. NO ; GO TO QUESTION 14

Q13. If yes, how often, on average, over the past year? (circle number)

1. LESS THAN ONE DAY PER MONTH
2. 1-3 DAYS PER MONTH
3. 4-6 DAYS PER MONTH
4. 7-10 DAYS PER MONTH
5. MORE THAN 10 DAYS PER MONTH
6. DAILY

Sometimes black sooty material falls out of the air onto outdoor furniture, houses and cars. This is often called black fallout. When there is black fallout people may have to clean their outdoor furniture and fixtures before using them. Some areas of Hamilton-Wentworth are more affected than others.

Q14. Have you ever experienced this black fallout in Hamilton-Wentworth?

1. YES
2. NO ; GO TO QUESTION 16

Q15. If yes, how often (on average over the year)

1. LESS THAN ONE DAY PER MONTH
2. 1-3 DAYS PER MONTH
3. 4-6 DAYS PER MONTH
4. 7-10 DAYS PER MONTH
5. MORE THAN 10 DAYS PER MONTH
6. DAILY

Sometimes the air in Hamilton-Wentworth is clear, sometimes it is not. At times there is a haze over the city like the one in the picture below. This poor visibility can be caused by air pollutants being trapped near the ground or by a chemical reaction in the air. Most poor visibility days occur in the spring, summer, and fall.

A poor visibility day in Hamilton-Wentworth

Q16. Have you ever noticed such poor visibility days in Hamilton-Wentworth?
1. YES
2. NO ; GO TO QUESTION 18

Q17. If yes, how often on average? (Consider only the spring, summer and fall months)
1. LESS THAN ONE DAY PER MONTH
2. 1-3 DAYS PER MONTH
3. 4-6 DAYS PER MONTH
4. 7-10 DAYS PER MONTH
5. MORE THAN 10 DAYS PER MONTH
6. DAILY

Poor air quality can affect health. Some effects such as headaches or sore throats are minor. Others may be more serious. Some people in Hamilton-Wentworth suffer from breathing or heart conditions. Poor air quality can make these conditions worse. When air quality is poor, some people with these conditions must go to the hospital. Some heart and lung conditions made worse by air pollution are chronic bronchitis, emphysema, asthma, angina, and heart disease. Air pollution can also lead to premature deaths among people who are already sick. Children and seniors are more sensitive to these effects.

Q18. Do you believe the quality of air in Hamilton-Wentworth has affected your health or the health of another member of your household? (Circle answer)
1. Yes
2. No. Go to Question 20

Q19. If yes, how often do you notice the effects of air quality on your health or the health of another member of your household? (Circle answer)
1. LESS THAN ONE DAY PER MONTH
2. 1-3 DAYS PER MONTH
3. 4-6 DAYS PER MONTH
4. 7-10 DAYS PER MONTH
5. MORE THAN 10 DAYS PER MONTH
6. DAILY

Q20. Were you ever told by a doctor or any other health professional that you, or any of your family members, currently have any of the following lung or heart conditions? (check all that apply)

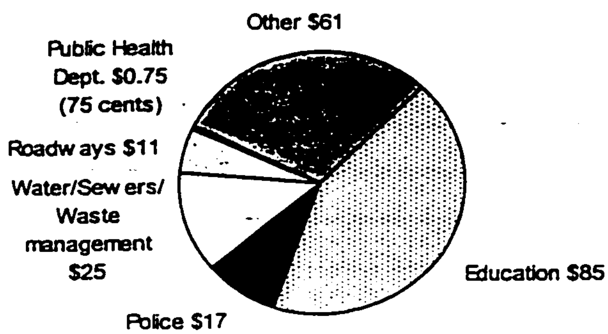
condition	**you**	**another family member**
asthma		
heart disease		
angina		
emphysema or chronic bronchitis		
any other breathing or heart disease		

Q21. We have discussed four attributes of air quality: bad odour, black fallout, poor visibility and health effects. Which is the most important to you? *(Put a 1 beside it).* Which is the least important to you? *(put a 4 beside it)* Of the remaining two attributes which concerns you the most? *(Put a 2 beside it. Put a 3 beside the remaining attribute.)*

- •BAD ODOUR ______
- •BLACK FALLOUT ______
- •POOR VISIBILITY ______
- •HEALTH EFFECTS ______

Property taxes collected in Hamilton-Wentworth are used to pay for many services. If you own your home you pay these taxes directly. If you rent your home your landlord uses some of your rent money to pay the taxes. In 1996, the typical household in Hamilton-Wentworth paid about $200 per month in property taxes. Of this total, about $85 was spent on education, about $25 was spent on water services, sewers and waste management, about $17 was spent on police services, about $11 was spent on roadways, and about $0.75 (75 cents) was spent on the Department of Public Health. (See chart below)

Shares of Regional Taxes
Total taxes paid = $200 per month



Q22. Som epeople believe property taxes are just about right, others believe that they are too high, others believe that they are too low. Do you think that property taxes are too high, too low, or about right? (Circle number)

1. TOO HIGH
2. TOO LOW
3. ABOUT RIGHT
4. DON'T KNOW

Section B

If we do nothing about pollution air quality may get worse. If we spend money to reduce pollution air quality may improve. The current level of each attribute that has been discussed in this survey is given in the table below. The following questions ask how you would balance changes in property taxes with changes in air quality. In each question, we will give you four choices. Choice A is always the current situation. Choices B, C and D are alternatives to the current situation. In each alternative, the attributes of air quality may be the same, about one-third better, or about one-third worse, than the current situation. The following chart shows what we mean.

Attribute	Current Situation	About One-third Better	About One-third Worse
Bad Odour	**4 days of bad odour per month**	**3 days of bad odour per month**	**5 days of bad odour per month**
Black Fallout (BFO)	**3 days of black fallout per month**	**2 days of black fallout per month**	**4 days of black fallout per month**
Visibility	**3 days of poor visibility per month**	**2 days of poor visibility per month**	**4 days of poor visibility per month**
Health Effects	**18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air**	**12 extra hospital admissions per month & 1 extra death per month, compared to perfectly clean air**	**24 extra hospital admissions per month and 3 extra deaths per month, compared to perfectly clean air**

Reducing pollution can be expensive. Taxes may be needed to offset these costs. Therefore each alternative will also include a change in property taxes or rental payments for your entire household. Please remember that any money spent on increased taxes or rent cannot be spent on anything else. You can refer to the chart on page 7 to recall what is presently spent on some regional services.

There are nine sets of choices. For each set, we will ask which choice you like the best and which you like the least. Then we will ask which, of the remaining two choices you like the best. All of these questions deal with matters of opinion. There are no right or wrong answers and you do not have to explain any of your answers.

THESE CHOICE SETS DO NOT DESCRIBE ANY ACTUAL CHOICES FACING THE REGION OF HAMILTON-WENTWORTH RIGHT NOW. THEY ARE CONSTRUCTED ONLY TO HELP YOU TO DESCRIBE HOW YOU WOULD BALANCE THE VARIOUS ATTRIBUTES OF AIR QUALITY AND PROPERTY TAXES OR RENTAL PAYMENTS.

Here is a sample question so you understand what we mean. Choice C has a better level of each air quality attribute, and a lower level of monthly taxes than choice A. Choice B, on the other hand, has a worse level of each attribute and a higher monthly tax or rent payment in comparison to choice A.

	State of the Environment			
Attribute	Choice A (Current situation)	Choice B	Choice C	Choice D
Bad Odour	4 days of bad odour per month	One-third Worse	One-third Better	Same
Black Fallout (BFO)	3 days of BFO per month	One-third Worse	One-third Better	Same
Poor Visibility	3 days of poor visibility per month	One-third Worse	One-third Better	Same
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	One-third Worse	One-third Better	Same
Monthly Property Taxes or Rent	$200 per month for typical household	$5 more taxes/rent	$5 less taxes/rent	$5 more taxes/rent
Rank	*2*	*4*	*1*	*3*

Please put a 1 under the choice you like BEST.
Please put a 4 under the choice you like LEAST.
Consider the remaining choices. Put a 2 under the one you like best. Put a 3 under the other one.

A Sample Question

Suppose that you like choice C the best. You would put a 1 under choice C as shown.
Suppose that you like choice B the least. You would put a 4 under choice B as shown.
Choice A and D remain. Suppose you like Choice A better than Choice D.
You would put a 2 under that Choice A and a 3 under Choice D, as shown.

Now we will give you nine choice sets in which the choices are not as straightforward as in the sample question. Please rank the choices in each in the same way as you did for the sample question

Q23. **Choice Set 1**

	State of the Environment			
Attribute	Choice A (Current Situation)	Choice B	Choice C	Choice D
Bad Odour	4 days of bad odour per month	Same as current situation	One-third Worse	One-third Better
Black Fallout (BFO)	3 days of BFO per month	Same as current situation	Same as current situation	One-third Better
Poor Visibility	3 days of poor visibility per month	Same as current situation	One-third Worse	One-third Better
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	Same as current situation	One-third Worse	One-third Better
Monthly Property Taxes or Rent	$200 per month for typical household	$5 more taxes/rent	$5 less taxes/rent	Same as current situation
Rank				

Please put a 1 under the choice you like BEST.
Please put a 4 under the choice you like LEAST.
Consider the remaining choices. Put a 2 under the one you like best. Put a 3 under the other one.

Q24. **Choice Set 2**

	State of the Environment			
Attribute	Choice A (Current Situation)	Choice B	Choice C	Choice D
Bad Odour	4 days of bad odour per month	One-third Better	One-third Better	Same as current situation
Black Fallout (BFO)	3 days of BFO per month	Same as current situation	One-third Better	One-third Worse
Poor Visibility	3 days of poor visibility per month	One-third Worse	Same as current situation	One-third Better
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	Same as current situation	One-third Better	One-third Better
Monthly Property Taxes or Rent	$200 per month for typical household	$20 more taxes/rent	Same as current situation	$20 less taxes/rent
Rank				

Please put a 1 under the choice you like BEST.
Please put a 4 under the choice you like LEAST.
Consider the remaining choices. Put a 2 under the one you like best. Put a 3 under the other one.

Q25. **Choice Set 3**

	State of the Environment			
Attribute	Choice A (Current Situation)	Choice B	Choice C	Choice D
Bad Odour	4 days of bad odour per month	One-third Better	Same as current situation	One-third Worse
Black Fallout (BFO)	3 days of BFO per month	One-third Worse	One-third Better	Same as current situation
Poor Visibility	3 days of poor visibility per month	Same as current situation	One-third Worse	One-third Better
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	One-third Worse	One-third Better	Same as current situation
Monthly Property Taxes or Rent	$200 per month for typical household	$10 less taxes/rent	Same as current situation	$10 more taxes/rent
Rank				

Please put a 1 under the choice you like BEST.
Please put a 4 under the choice you like LEAST.
Consider the remaining choices. Put a 2 under the one you like best. Put a 3 under the other one.

Q26 **Choice Set 4**

	State of the Environment			
Attribute	Choice A (Current Situation)	Choice B	Choice C	Choice D
Bad Odour	4 days of bad odour per month	Same as current situation	One-third Better	One-third Worse
Black Fallout (BFO)	3 days of BFO per month	Same as current situation	One-third Worse	Same as current situation
Poor Visibility	3 days of poor visibility per month	One-third Worse	Same as current situation	One-third Better
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	Same as current situation	One-third Better	One-third Worse
Monthly Property Taxes or Rent	$200 per month for typical household	$25 more taxes/rent	Same as current situation	$25 less taxes/rent
Rank				

Please put a 1 under the choice you like BEST.
Please put a 4 under the choice you like LEAST.
Consider the remaining choices. Put a 2 under the one you like best. Put a under the other one.

Q27. **Choice Set 5**

	State of the Environment			
Attribute	Choice A (Current Situation)	Choice B	Choice C	Choice D
Bad Odour	4 days of bad odour per month	One-third Better	One-third Worse	Same as current situation
Black Fallout (BFO)	3 days of BFO per month	One-third Better	One-third Worse	Same as current situation
Poor Visibility	3 days of poor visibility per month	One-third Better	One-third Worse	Same as current situation
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	Same as current situation	One-third Better	One-third Worse
Monthly Property Taxes or Rent	$200 per month for typical household	$10 more taxes/rent	Same as current situation	$10 less taxes/rent
Rank				

Please put a 1 under the choice you like BEST.
Please put a 4 under the choice you like LEAST.
Consider the remaining choices. Put a 2 under the one you like best. Put a 3 under the other one.

Q28. **Choice Set 6**

	State of the Environment			
Attribute	Choice A (Current Situation)	Choice B	Choice C	Choice D
Bad Odour	4 days of bad odour per month	One-third Better	One-third Worse	Same as current situation
Black Fallout (BFO)	3 days of BFO per month	Same as current situation	One-third Better	One-third Worse
Poor Visibility	3 days of poor visibility per month	One-third Worse	Same as current situation	One-third Better
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	One-third Worse	Same as current situation	One-third Better
Monthly Property Taxes or Rent	$200 per month for typical household	$50 less taxes/rent	$50 more taxes/rent	Same as current situation
Rank				

Please put a 1 under the choice you like BEST.
Please put a 4 under the choice you like LEAST.
Consider the remaining choices. Put a 2 under the one you like best. Put a 3 under the other one.

Q29. **Choice Set 7**

	State of the Environment			
Attribute	Choice A (Current Situation)	Choice B	Choice C	Choice D
Bad Odour	4 days of bad odour per month	One-third Better	One-third Worse	Same as current situation
Black Fallout (BFO)	3 days of BFO per month	Same as current situation	One-third Better	One-third Worse
Poor Visibility	3 days of poor visibility per month	One-third Worse	Same as current situation	One-third Better
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	One-third Better	One-third Worse	Same as current situation
Monthly Property Taxes or Rent	$200 per month for typical household	Same as current situation	$15 less taxes/rent	$15 more taxes/rent
Rank				

Please put a 1 under the choice you like BEST.
Please put a 4 under the choice you like LEAST.
Consider the remaining choices. Put a 2 under the one you like best. Put a 3 under the other one.

Q30. **Choice Set 8**

	State of the Environment			
Attribute	Choice A (Current Situation)	Choice B	Choice C	Choice D
Bad Odour	4 days of bad odour per month	Same as current situation	One-third Worse	One-third Better
Black Fallout (BFO)	3 days of BFO per month	One-third Better	Same as current situation	One-third Worse
Poor Visibility	3 days of poor visibility per month	One-third Worse	One-third Better	Same as current situation
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	One-third Worse	One-third Better	Same as current situation
Monthly Property Taxes or Rent	$200 per month for typical household	$15 more taxes/rent	Same as current situation	$15 less taxes/rent
Rank				

Please put a 1 under the choice you like BEST.
Please put a 4 under the choice you like LEAST.
Consider the remaining choices. Put a 2 under the one you like best. Put a 3 under the other one.

Q31. **Choice Set 9**

	State of the Environment			
Attribute	Choice A (Current Situation)	Choice B	Choice C	Choice D
Bad Odour	4 days of bad odour per month	One-third Better	Same as current situation	One-third Worse
Black Fallout (BFO)	3 days of BFO per month	One-third Better	Same as current situation	One-third Worse
Poor Visibility	3 days of poor visibility per month	One-third Better	Same as current situation	One-third Worse
Health Effects	18 extra hospital admissions per month and 2 extra deaths per month, compared to perfectly clean air	One-third Worse	One-third Better	*Same as current situation*
Monthly Property Taxes or Rent	$200 per month for typical household	$20 more taxes/rent	*Same as current situation*	$20 more taxes/rent
Rank				

Please put a 1 under the choice you like BEST.
Please put a 4 under the choice you like LEAST.
Consider the remaining choices. Put a 2 under the one you like best. Put a 3 under the other one.

Section C:

Now we would like to ask you a few questions about yourself.

Q32. In what year were you born?
________YEAR

Q33. What is the highest level of education you have completed? (Circle number)
1. LESS THAN GRADE 8
2. ELEMENTARY SCHOOL, GRADE 8
3. HIGH SCHOOL (GRADE 12 OR 13)
4. POST SECONDARY (CERTIFICATE OR DIPLOMA)
5. UNIVERSITY DEGREE (B.A., B.Sc.)
6. POST GRADUATE (M.A., M.Sc., Ph.D.)

Q34. What is your current employment status? (Circle number)
1. FULL TIME HOUSEWORK
2. FULL-TIME PAID EMPLOYMENT
3. PART-TIME PAID EMPLOYMENT
4. SELF-EMPLOYED
5. UNEMPLOYED
6. RETIRED
7. STUDENT

Q35. Which number best represents your household type? (Circle number)
1. COUPLE WITH CHILDREN
2. COUPLE WITHOUT CHILDREN
3. LONE PARENT WITH CHILDREN
4. NON-FAMILY, ONE PERSON
5. NON-FAMILY, MORE THAN ONE PERSON

Q36. How many people live in your household (including yourself)?
__________NUMBER

Q37. How many are adults 18 years or older?
________NUMBER

Q38. How many of the adults are aged 65 or older?
_______NUMBER

Q39. Are you currently employed, or have you ever worked in the health care field? (For example, medical doctor, nurse, health researcher, etc) (Circle number)
1. YES
2. NO

Q40. Do you support any environmental organizations (for example, Greenpeace, Earthroots, Pollution Probe) with time or money? (Circle number)
1. YES
2. NO

Q41. What is your sex? (Circle number)
1. MALE
2. FEMALE

Q42. Which range best represents the total income of your household in 1995, before taxes. (Circle number)

1. under $20,000	6. $60,000-$69,999
2. $20,000-$29,999	7 $70,000-$79,999
3. $30,000-$39,999	8. $80,000-$89,999
4. $40,000-$49,999	9. $90,000-$99,999
5. $50,000-$59,999	10. $100,000 or over

Q43. How long have you lived in your present home? (Circle number)
1. LESS THAN ONE YEAR
2. 1-5 YRS
3. MORE THAN 5 YEARS

Q44. What type of dwelling do you live in? (Circle number)
1. DETACHED HOUSE
2. SEMI-DETACHED HOUSE
3. HIGH-RISE APARTMENT BUILDING
4. LOW-RISE APARTMENT BUILDING (3 FLOORS OR LESS)

Q45. Do you own your home or do you rent? (Circle number)
1. OWN
2. RENT

Q46. How many hours per week do you spend outdoors engaging in physical activity such as exercise gardening, or walking? (Circle number)
1. LESS THAN 5 HOURS
2. 5-10 HOURS
3. MORE THAN 10 HOURS

Q47. Please tell us about any part of this questionnaire that was hard for you to understand

Q48. Is there anything else you would like to say about air quality in Hamilton-Wentworth or about this questionnaire.

Appendix B.

Creating the Factorial Design for the Questionnaire

Following Petersen (1985) a 1/9 partial factorial design was created from the 3 levels of the 5 attributes. The partial factorial design was divided into 9 blocks of 3 so as to permit estimation of all first and second order interactions of the factors. Specifically, the following steps were undertaken.

1. A general 3^5 factorial design was created with levels 1, 2 and 3 in each of the 5 factors A, B, C, D and E. Denote the level of factor i by $x_i, i=A,B,C,D,E$.

2. A 1/9 partial fractional design was made with defining effects ABCDE and A^2BCDE^2 (Petersen, 1985, 210ff) by selecting that subset of the original 243 combinations which satisfied the equations

$$x_A + x_B + x_C + x_D + x_E = 1$$
$$2x_A + x_B + x_C + x_D + 2x_E = 1$$

Note that under the first defining effect, all first order effects are aliased with fourth order or fifth order interactions. For example, the aliases of the first order effect of A are

$$
\begin{aligned}
A(ABCDE) &= A^2BCDE \\
&= (A^2BCDE)^2 \\
&= AB^2C^2D^2E^2
\end{aligned}
$$

and

$$
\begin{aligned}
A(A^2B^2C^2D^2E^2) &= A^3B^2C^2D^2E^2 \\
&= B^2C^2D^2E^2
\end{aligned}
$$

where the second equality in the first set of equations follows from the rules for reducing symbolic expressions in Petersen (201). Similarly, all second order interactions are aliased to higher order interactions. For example, AB is aliased to

$$
\begin{aligned}
AB(ABCDE) &= A^2B^2CDE \\
&= (A^2B^2CDE)^2 \\
&= ABC^2D^2E^2
\end{aligned}
$$

The second defining effect was derived by confounding AE with BCD. All first and second order interactions are aliased to third order and higher interactions. For example, A is aliased to

$$A(A^2BCDE^2) = BCDE^2$$

and B is aliased to

$$B(A^2BCDE^2)=A^2B^2CDE^2$$
$$=(A^2B^2CDE^2)^2$$
$$=ABC^2D^2E$$

3. The 1/9 factorial design was blocked into 9 blocks of three questions each by confounding further second-order interactions with higher order interactions using the confounding effects ABC^2DE^2 and AB^2CDE^2 (see Petersen, 213ff). To do this two orthogonal three-way partitions were generated using the equations

$$b_1=x_A+2x_B+x_C+x_D+2x_E\ (\text{modulo}3)$$
$$b_2=x_A+x_B+2x_C+x_D+2x_E\ (\text{modulo}3)$$

Crossing the three values of b_1 and b_2 yielded 9 blocks of three questions each. The result of this procedure is that the defining effects ABC^2DE^2 andAB^2CDE^2 cannot be distinguished from the effects of blocking.

4 The order of the questions was randomized within blocks.

Appendix C.

FURTHER REGRESSION RESULTS (NOT PRESENTED IN THE MAIN TEXT)

Table C1. Estimated coefficients by model
[first & second choices, small data set]
(Asymptotic standard errors in parentheses)

	Model 1: Linear	Model 2: Linear with quadratic taxes	Model 3: Quadratic	Model 4: Quadratic with demographics
respondents	N_i =188	N_i =188	N_i =188	N_i =148
choice sets	N_{CS} =3384	N_{CS} = 3384	N_{CS} = 3384	N_{CS} =2664*
pseudo R^2	.477	.496	.514	.506
Log likelihood	-2199	-2120	-2044	-1631
H (health)	2.328 (.055)	2.130 (.055)	2.305 (.087)	1.971 (.324)
O (odour)	.718 (.039)	.763 (.042)	.617 (.072)	.853 (.261)
F (black fallout)	.925 (.047)	.906 (.049)	.673 (.076)	795 (.310)
V (visibility)	.489 (.041)	.595 (.044)	.785 (.081)	.308 (.278)
T (tax)	-.032 (.002)	-.038 (.002)	-.041 (.004)	-.04 (.015)
T^2		-.0007 (.0001)	-.0006 (.0001)	-.0006 (.0001)
H^2			.186 (.110)	.249 (.121)
H·O			.084 (.103)	.121 (.114)
H·F			.309 (.113)	.335 (.124)
H·V			-.269 (.098)	-.216 (.107)
H·T			-.007 (.006)	-.013 (.007)
O^2			.019 (.340)	-.116 (.374)
O·F			-.013 (.159)	.074 (.178)
O·V			.115 (.290)	-.065 (.314)
O·T			.006 (.004)	.006 (.004)
F^2			-.519 (.313)	-.300 (.341)
F·V			-.181 (.364)	-.076 (.399)
F·T			.011 (.005)	.012 (.006)
V^2			-.442 (.140)	-.486 (.158)
V·T			.020 (.006)	.024 (.006)
I·H				.002 (.002)
I·O				.002 (.002)
I·F				.0004 (.002)
I·V				-.002 (.002)
I·T				-.00001 (.00009)
E·H				-.255 (.151)
E·O				-.224 (.118)
E·F				-.044 (.145)
E·V				.126 (.128)
E·T				-.004 (.007)
A·H				.001 (.005)
A·O				-.006 (.004)
A·F				-.003 (.005)
A·V				.009 (.005)
A·T				.00001 (.00002)
C·H				.382 (.12)
C·O				.117 (.096)
C·F				.088 (.116)
C·V				.098 (.103)
C·T				.002 (.005)

Table C2. Estimated coefficients by model
[first and second and third choices, small data set]
(Asymptotic standard errors in parentheses)

	Model 1: Linear	Model 2: Linear with quadratic taxes	Model 3: Quadratic	Model 4: Quadratic with demographics
respondents	N_i =188	N_i =188	N_i =188	N_i =148
choice sets	N_{CS} =5076	N_{CS} =5076	N_{CS} =5076	N_{CS} =3996*
pseudo R^2	.400	.414	.452	.4455
Log likelihood	-3229	-3148	-2950	-2347
H (health)	2.080 (.044)	1.914 (.045)	2.068 (.059)	1.522 (.411)
O (odour)	.590 (.030)	.611 (.031)	.598 (.043)	1.113 (.387)
F (black fallout)	.735 (.034)	.684 (.035)	.531 (.049)	.593 (.448)
V (visibility)	.449 (.030)	.491 (.031)	.651 (.049)	.219 (.387)
T (tax)	-.026 (.001)	-.025 (.001)	-.025 (.002)	-.058 (.029)
T^2		-.0006 (.0001)	-.0003 (.0001)	-.0009 (.0003)
H^2			.409 (.078)	1.252 (.259)
H·O			.414 (.071)	.213 (.245)
H·F			.673 (.081)	.518 (.253)
H·V			.156 (.061)	.102 (.200)
H·T			-.016 (.005)	-.025 (.014)
O^2			.077 (.204)	1.178 (.510)
O·F			-.122 (.133)	-.012 (.292)
O·V			.453 (.137)	2.002 (.576)
O·T			-.004 (.003)	.006 (.011)
F^2			-.741 (.160)	-1.771 (.555)
F·V			-.005 (.228)	-1.022 (.553)
F·T			-.004 (.004)	.016 (.014)
V^2			-.292 (.124)	-.181 (.295)
V·T			-.0002 (.003)	.021 (.014)
I·H				.002 (.003)
I·O				-.002 (.002)
I·F				.001 (.003)
I·V				.001 (.002)
I·T				-.0003 (.0002)
E·H				-.393 (.191)
E·O				-.208 (.166)
E·F				.033 (.209)
E·V				.003 (.176)
E·T				.005 (.013)
A·H				.006 (.007)
A·O				-.005 (.006)
A·F				-.003 (.007)
A·V				.008 (.006)
A·T				.0004 (.0005)
C·H				.319 (.148)
C·O				-.002 (.132)
C·F				.042 (.163)
C·V				-.0333 (.140)0
C·T				.006 (.010)

Table C3. Estimated coefficients by income level *
[Model 3-full quadratic, small data set]
(asymptotic standard errors in parentheses)

	First choices only				First and second choices			
	low income		high income		low income		high income	
respondents	$N_i = 63$		$N_i = 92$		$N_i = 63$		$N_i = 92$	
choice sets	$N_{CS} = 567$		$N_{CS} = 825$§		$N_{CS} = 1134$		$N_{CS} = 1656$	
pseudo R^2	.633		.606		.505		.515	
Log likelihood	-282.1		-439.1		-697.9		-997.1	
H (health)	2.345	(.330)	1.778	(.178)	2.002	(.132)	2.404	(.131)
O (odour)	1.013	(.345)	.402	(.177)	.515	(.118)	.600	(.108)
F (black fallout)	.818	(.319)	.618	(.191)	.617	(.125)	.688	(.112)
V (visibility)	.444	(.289)	.700	(.201)	.671	(.128)	.842	(.121)
T (tax)	-.066	(.027)	-.053	(.015)	-.038	(.007)	-.049	(.006)
T^2	-.001	(.0006)	-.001	(.003)	.0007	(.0002)	-.0006	(.0001)
H^2	.723	(.446)	1.362	(.316)	.242	(.184)	.200	(.160)
H·O	-.270	(.476)	.303	(.295)	.294	(.170)	-.013	(.154)
H·F	.121	(.450)	.574	(.320)	.300	(.192)	.390	(.162)
H·V	.251	(.340)	.014	(.253)	-.047	(.157)	-.389	(.146)
H·T	.0007	(.027)	-.029	(.016)	-.003	(.010)	-.018	(.009)
O^2	-329	(.609)	-.473	(.460)	-.696	(.709)	.260	(.469)
O·F	.466	(.504)	-.216	(.370)	.603	(.314)	-.315	(.219)
O·V	.196	(.294	.243	(.232)	-.866	(.614)	.472	(.395)
O·T	.038	(.025)	.001	(.013)	.010	(.007)	.001	(.006)
F^2 †	-	-	-	-	.696	(.643)	-.973	(.436)
F·V	.410	(.684)	.646	(.556)	.457	(.746)	-.356	(.503)
F·T	.041	(.031)	.011	(.016)	.019	(.009)	.003	(.008)
V^2	-.877	(.532)	.050	(.364)	-.848	(.288)	-.216	(.191)
V·T	.024	(.031)	.027	(.017)	.025	(.009)	.025	(.008)

*These regression results correspond to the estimated trade-offs reported in table 7.8.
† Note that f2 was dropped when using first choices only due to collinearity
§ The three choice sets in which the respondent ranked the dropped alternative first were dropped

Table C4. Estimated coefficients by cardio-respiratory experience [Model 3-full quadratic, small data set] (asymptotic standard errors in parentheses)*

	First choices only				First and second choices			
	yes		no		yes		no	
respondents	N_i =109		N_i =79		N_i = 109		N_i =79	
choice sets	N_{CS} =981		N_{CS} =711		N_{CS} =1962		N_{CS} =1422	
pseudo R^{2}	.652		.586		.534		.485	
Log likelihood	-473.9		-407.7		-1124.3		-909.4	
H (health)	1.991	(.189)	1.873	(.207)	2.446	(.121)	2.151	(.128)
O (odour)	.552	(.205)	.576	(.200)	.626	(.099)	.617	(.108)
F (black fallout)	.563	(.213)	.703	(.211)	.724	(.104)	.636	(.113)
V (visibility)	.650	(.213)	.722	(.224)	.779	(.109)	.812	(.124)
T (tax)	-.038	(.014)	-.071	(.019)	-.045	(.006)	-.038	(.006)
T^2	-.0008	(.0003)	-.001	(.0004)	-.0007	(.0001)	-.0005	(.0001)
H^2	1.563	(.343)	.856	(.321)	.084	(.153)	.280	(.160)
H·O	.529	(.327)	.098	(.309)	.142	(.140)	.010	(.154)
H·F	.560	(.347)	.400	(.328)	.137	(.157)	.484	(.163)
H·V	.247	(.274)	-.048	(.267)	-.273	(.133)	-.286	(.147)
H·T	-.005	(.017)	-.033	(.020)	.001	(.008)	-.016	(.009)
O^2	1.482	(.601)	.807	(.765)	-.085	(.484)	.223	(.485)
O·F	.202	(.384)	-.189	(.370)	.267	(.225)	-.305	(.228)
O·V	2.673	(.696)	1.266	(.843)	-.204	(.414)	.519	(.412)
O·T	.003	(.014)	.014	(.015)	.007	(.005)	.004	(.006)
F^2	-2.387	(.662)	-1.146	(.816)	-.223	(.445)	-.899	(.449)
F·V	-1.350	(.686)	-.572	(.802)	-.171	(.515)	-.287	(.522)
F·T	.034	(.019)	.012	(.019)	.016	(.007)	.004	(.008)
V^2	-.196	(.397)	-.124	(.371)	-.600	(.199)	-.299	(.200)
V·T	.008	(.017)	.033	(.021)	.027	(.008)	.011	(.008)

Table C5. Estimated coefficients by black fallout experience
[Model 3-full quadratic, small data set]
(asymptotic standard errors in parentheses)[§]

	First choices only				First and second choices			
	yes**		no		yes		no	
respondents*	N_i =125		N_i =60		N_i =125		N_i =60	
choice sets	N_{CS} =1125		N_{CS} =720		N_{CS} =2250		N_{CS} =1440	
pseudo R^2	.607		.655		.502		.541	
Log likelihood	-613.1		-258.1		-1391.4		-616.3	
H (health)	1.782	(.158)	2.453	(.303)	2.179	(.102)	2.648	(.177)
O (odour)	.632	(.175)	.502	(.245)	.677	(.086)	.497	(.140)
F (black fallout)	.541	(.172)	1.073	(.327)	.718	(.092)	.598	(.138)
V (visibility)	.559	(.171)	.959	(.340)	.670	(.093)	1.186	(.175)
T (tax)	-.043	(.012)	-.075	(.025)	-.038	(.005)	-.053	(.009)
T^2	-.0006	(.0003)	-.002	(.0006)	-.0005	(.0001)	-.0008	(.0002)
H^2	1.393	(.281)	.657	(.429)	.062	(.131)	.534	(.209)
H·O	.290	(.278)	.213	(.365)	.057	(.122)	.111	(.201)
H·F	.720	(.281)	-.428	(.474)	.279	(.134)	.443	(.215)
H·V	.216	(.217)	-.223	(.396)	-.173	(.114)	-.569	(.199)
H·T	-.027	(.015)	.025	(.024)	-.009	(.007)	-.007	(.011)
O^2	1.359	(.539)	1.228	(.920)	.346	(.403)	-.554	(.666)
O·F	-.018	(.329)	.455	(.496)	.235	(.197)	-.542	(.284)
O·V	2.330	(.308)	1.414	(1.034)	.208	(.338)	.048	(.576)
O·T	.014	(.012)	-.004	(.018)	.006	(.005)	.002	(.007)
F^2	-1.969	(.578)	-1.635	(1.014)	-.530	(.364)	-.638	(.626)
F·V	-1.104	(.594)	-1.417	(.973)	-.540	(.430)	.479	(.717)
F·T	.016	(.016)	.052	(.028)	.009	(.006)	.012	(.010)
V^2	-.236	(.334)	-.391	(.461)	-.549	(.175)	-.358	(.255)
V·T	.010	(.014)	.030	(.026)	.010	(.006)	.044	(.012)

[§]These regressions correspond tp the trade-off values reported in table 7.10
*Note that not all of the respondents provided this data, hence the total N is 185
**yes/no refers to experience with the attribute

Table C6. Estimated coefficients by income level, region 1[*]
[Model 3-full quadratic, small data set, first and second choices]
(asymptotic standard errors in parentheses)[*]

	Region 1[‡]				Region 2			
	low income		high income		low income		high income	
respondents	$N_i = 35$		$N_i = 25$		$N_i = 48$		$N_i = 24$	
choice sets	$N_{CS} = 630$		$N_{CS} = 448$[†]		$N_{CS} = 864$		$N_{CS} = 432$	
pseudo R^2	.457		.559		.519		.556	
Log likelihood	-411		-238		-516		-238	
H (health)	1.742	(.150)	3.07	(.328)	2.326	(.181)	2.622	(.296)
O (odour)	.560	(.148)	.933	(.233)	.406	(.154)	.368	(.227)
F (black fallout)	.637	(.157)	1.128	(.248)	.495	(.157)	.481	(.227)
V (visibility)	.550	(.147)	1.053	(.262)	.721	(.172)	1.051	(.313)
T (tax)	-.032	(.008)	-.048	(.012)	-.043	(.009)	-.048	(.014)
T^2	-.0005	(.0002)	-.0004	(.0002)	-.0007	(.0002)	-.0008	(.0003)
H^2	.208	(.223)	-.308	(.355)	.486	(.230)	.413	(.370)
H·O	.186	(.211)	-.652	(.360)	.297	(.221)	.544	(.332)
H·F	.221	(.235)	.383	(.343)	.609	(.238)	.651	(.383)
H·V	.117	(.186)	-.434	(.315)	-.327	(.212)	-.496	(.354)
H·T	-.002	(.013)	.015	(.019)	-.033	(.012)	-.004	(.018)
O^2	.591	(.460)	-.601	(.520)	-.345	(.683)	-.981	(1.128)
O·F	1.148	(.429)	-.061	(.487)	-.545	(.311)	-.262	(.561)
O·V	-.155	(.194)	-.532	(.286)	.103	(.566)	-.182	(.934)
O·T	.009	(.009)	.010	(.012)	-.005	(.008)	.008	(.011)
F^2	-	-	-	-	-.593	(.629)	.138	(.998)
F·V	-.822	(.506)	.310	(.611)	.309	(.737)	.860	(1.224)
F·T	.019	(.011)	-.009	(.014)	.018	(.012)	.015	(.014)
V^2	-1.120	(.401)	.434	(.434)	.020	(.266)	-.649	(.515)
V·T	.007	(.011)	.016	(.016)	.034	(.012)	.054	(.020)

[*]These regressions correspond to the trade-offs reported in table 7.14
[‡]Choice set 1, alternative 3 was dropped from all respondents as discussed in the text (section 7.8).
[†]Two choice sets in which the respondent chose the dropped alternative were dropped (see section 7.8 for details)
[‡]f2 was dropped due to collinearity

Appendix D
Likelihood Ratio Test Results

This appendix contains information about the likelihood ratio (LR) tests that were used throughout Chapter 7. They are organized according to the section in which they appeared in that chapter.

The likelihood ratio (LR) statistic is calculated using the following formula:

$$LR = 2[\mathcal{L}(\theta_{UNR}) - \mathcal{L}(\theta_{RES})] \sim \chi^2$$

Where $\mathcal{L}(\theta_{UNR})$ is the log likelihood of the unrestricted model and $\mathcal{L}(\theta_{RES})$ is the log likelihood of the restricted model. (Note that rejection of the null hypothesis is based on the 5% significance level.)

<u>**Section 7.2**</u>

Model 3 (quadratic model) vs Model 4 (quadratic model plus demographics)
H_0 : coefficients on all the demographic variables are zero (i.e. the quadratic model is correct)

$\mathcal{L}(\theta_{UNR}) = -713$
$\mathcal{L}(\theta_{RES}) = -891$

$$LR = 2[-713+891] = 356$$

# of restrictions: 40-20=20

$p < .001$

Hence, the null hypothesis, H_0, is rejected - i.e. the quadratic model plus demographics is the correct one.

Model 1 (linear model) vs Model 2 (linear model with quadratic tax term)
H_0 : coefficients on all "extra" variables are zero (i.e. linear-quadratic taxes-model is correct)

$\mathcal{L}(\theta_{UNR}) = -928.7$
$\mathcal{L}(\theta_{RES}) = -940.1$

LR = 2 [-928.7 +940.1]
= 22.6
# of restrictions = 6-5=1

χ^2 (1, .99) = 6.63

p< .001

Hence, the null hypothesis, H_0, is rejected - i.e. the linear model is the correct one.

Model 2 (linear model with quadratic tax term) vs Model 3 (fullquadratic model):
H_0 : coefficients on all "extra" variables are zero (i.e. linear-quadratic taxes-model is correct)

LR = 2 [-890.98+ 928.74]
= 73.54

# of restrictions = 20-6=14

p<.001

Hence, the null hypothesis, H_0, is rejected - i.e. the quadratic model is the correct one.

Section 7.4.

Using Model 5[1]:

small vs medium data set

H_0 : The incremental data used in the medium data set that is not included in the small data provides different estimates than the data employed in the small data set only

$\mathcal{L}(\theta_{UNR}) = -858.5-1293.7 = -2152.2$
$\mathcal{L}(\theta_{RES}) = -2248.7$

LR = 2[-2152.2 + 2248.7] = 193

\# of restrictions =50

p =<.001

Hence the null hypothesis is rejected / not-rejected

small vs large data set

H_0 : The incremental data used in the large data set that is not included in the small data provides different estimates than the data employed in the small data set only

$\mathcal{L}(\theta_{UNR}) = -858.5 - 1829.3 = -2687.8$
$L(\theta_{RES}) = -2839.9$

LR = 2[-2687.8 +2839.9] = 304.2

\# of restrictions = 50

p < .001

Hence the null hypothesis is rejected

[1]See section 7.4. for details on Model 5 and Model 6.

Using Model 6:

small vs medium data set

H_0 : The incremental data used in the medium data set that is not included in the small data provides different estimates than the data employed in the small data set only

$\mathcal{L}(\theta_{UNR}) = -857.9 - 1277.9 = -2135.8$
$\mathcal{L}(\theta_{RES}) = -2241.2$

LR = 2[-2135.8+2241.2] = 210.8

LR = 210.8
# of restrictions = 55

$p < .001$

Hence the null hypothesis is rejected

small vs large data set

H_0 : The incremental data used in the large data set that is not included in the small data provides different estimates than the data employed in the small data set only

$\mathcal{L}(\theta_{UNR}) = -857.9 - 1806.2 = -2664.1$
$\mathcal{L}(\theta_{RES}) = -2828.2$

LR = 2[-2664.1+2828.2]=328.2
# of restrictions = 55

$p < .001$

Hence the null hypothesis is rejected / not-rejected

Section 7.6.

low-income vs high-income

- Model 3 (quadratic) was employed

first choices

H_0 : Low-income and high income respondents had the same estimated coefficients

$\mathcal{L}(\theta_{UNR}) = -282.1-439.1= -721.2$

$\mathcal{L}(\theta_{RES}) = -737.2$

LR = 32

# of restrictions = 19

p = .011

Hence the null hypothesis is rejected

first and second choices

H_0 : Low-income and high income respondents had the same estimated coefficients

$\mathcal{L}(\theta_{UNR}) = -697.9-997.1= -1695$

$\mathcal{L}(\theta_{RES}) = -1712.4$

LR = 34.8

# of restrictions = 20

p = .021

Hence the null hypothesis is rejected

cardio-respiratory experience

first choices
H_0 : those who experienced cardio-respiratory disease and those who did not had the same estimated coefficients.

$\mathcal{L}(\theta_{UNR})$ = -473.9-407.7 =-881.6
$\mathcal{L}(\theta_{RES})$ = -891.0

LR = 18.8

# of restrictions = 20

p = .535

Hence the null hypothesis is rejected

first and second choices
H_0 : those who experienced cardio-respiratory disease and those who did not had the same estimated coefficients.

$\mathcal{L}(\theta_{UNR})$ = -1124.3-909.4 = -2033.7
$\mathcal{L}(\theta_{RES})$ = -2044.3

LR = 21.2
# of restrictions = 20

p = .385

Hence the null hypothesis is rejected

bfo experience

first choices

H_0 : those who experienced black fallout and those who did not had the same estimated coefficients.

$\mathcal{L}(\theta_{UNR})$ = -613.1-258.1= 871.2

$\mathcal{L}(\theta_{RES})$ = -883.0

LR = 23.6

# of restrictions = 20

p = .260

Hence the null hypothesis is rejected

first and second choices

H_0 : those who experienced black fallout and those who did not had the same estimated coefficients.

$\mathcal{L}(\theta_{UNR})$ = -1391.4-616.3 = -2007.7

$\mathcal{L}(\theta_{RES})$ = -2021.7

LR = 28

# of restrictions = 20

p = .109

Hence the null hypothesis is rejected

Section 7.7.

seven regions
H_0 :All seven regions produce the same estimates

$\mathcal{L}(\theta_{UNR})$ = -759-643-167-155-102-216-50=-2093
$\mathcal{L}(\theta_{RES})$ = -2119.9

LR = 53.8

# of restrictions = 6

p <.001

Hence the null hypothesis is rejected

three regions (Model 2)
H_0 : All three regions produce the same estimates
H_1 :

$\mathcal{L}(\theta_{UNR})$ = -815-1024-269 = -2108
$\mathcal{L}(\theta_{RES})$ = -2119.9

LR = 23.8

# of restrictions = 6

p < .001

Hence the null hypothesis is rejected

three regions (Model 3)
H_0 : All three regions produce the same estimates.

$\mathcal{L}(\theta_{UNR})$ = -788-977-254=-2019
$\mathcal{L}(\theta_{RES})$ = -2044.3

LR = 50.6

# of restrictions = 20

p <.001

Hence the null hypothesis is rejected

Section 7.8.

Region 1: low income vs high income
H_0 : Low-income and high income respondents within Region 1 had the same estimated coefficients

$\mathcal{L}(\theta_{UNR})$ = -411-238 = -649
$\mathcal{L}(\theta_{RES})$ = -663.5

LR = 29

# of restrictions = 20

p = .088

Hence the null hypothesis is rejected

Region 2: low income vs high income
H_0 : Low-income and high income respondents within Region 2 had the same estimated coefficients

$\mathcal{L}(\theta_{UNR}) = -516-238 = -754$
$\mathcal{L}(\theta_{RES}) = 976.6$

LR = 445

# of restrictions = 20

$p < .001$

Hence the null hypothesis is rejected