EVALUATING THE BENEFITS AND EFFECTIVENESS OF PUBLIC POLICY

F. Mikael Sandström

AKADEMISK AVHANDLING

som för avläggande av ekonomie doktorsexamen vid Handelshögskolan i Stockholm framlägges för offentlig granskning fredagen den 10 september 1999, kl 13.15 i sal 750, Handelshögskolan, Sveavägen 65, Stockholm

STOCKHOLM SCHOOL
OF ECONOMICS

EVALUATING THE BENEFITS AND EFFECTIVENESS OF PUBLIC POLICY

STOCKHOLM SCHOOL OF ECONOMICS
EFI, THE ECONOMIC RESEARCH INSTITUTE

EFI Mission
EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution which works independently of economic, political and sectional interests. It conducts theoretical and empirical research in management and economic sciences, including selected related disciplines. The Institute encourages and assists in the publication and distribution of its research findings and is also involved in the doctoral education at the Stockholm School of Economics.
EFI selects its projects based on the need for theoretical or practical development of a research domain, on methodological interests, and on the generality of a problem.

Research Organization
The research activities are organized in nineteen Research Centers within eight Research Areas. Center Directors are professors at the Stockholm School of Economics.

ORGANIZATION AND MANAGEMENT	
Management and Organisation; (A)	Prof Sven-Erik Sjöstrand
Center for Ethics and Economics; (CEE)	Adj Prof Hans de Geer
Public Management; (F)	Prof Nils Brunsson
Information Management; (I)	Prof Mats Lundeberg
Man and Organisation; (PMO)	Acting Prof Jan Löwstedt
Industrial Production; (T)	Prof Christer Karlsson
MARKETING	
Center for Information and Communication Research; (CIC)	Adj Prof Bertil Thorngren
Center for Consumer Marketing; (CCM)	Associate Prof Magnus Söderlund
Marketing, Distribution and Industrial Dynamics; (D)	Prof Lars-Gunnar Mattsson
ECONOMIC PSYCHOLOGY	
Center for Risk Research; (CFR)	Prof Lennart Sjöberg
Economic Psychology; (P)	Prof Lennart Sjöberg
ACCOUNTING, CONTROL AND CORPORATE FINANCE	
Accounting and Managerial Finance; (B)	Prof Lars Östman
Managerial Economics; (C)	Prof Peter Jennergren
FINANCE	
Finance; (FI)	Prof Clas Bergström
ECONOMICS	
Center for Health Economics; (CHE)	Prof Bengt Jönsson
International Economics and Geography; (IEG)	Prof Mats Lundahl
Economics; (S)	Prof Lars Bergman
ECONOMICS STATISTICS	
Economic Statistics; (ES)	Prof Anders Westlund
LAW	
Law; (RV)	Prof Erik Nerep

Chairman of the Board: Prof Sven-Erik Sjöstrand. *Director:* Associate Prof Bo Sellstedt,

Adress
EFI, Box 6501, S-113 83 Stockholm, Sweden • Internet: www.hhs.se/efi/
Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail efi@hhs.se

EVALUATING THE BENEFITS AND EFFECTIVENESS OF PUBLIC POLICY

F. Mikael Sandström

STOCKHOLM SCHOOL OF ECONOMICS
EFI, THE ECONOMIC RESEARCH INSTITUTE

A Dissertation for the degree of
Doctor of Philosophy, Ph.D. (Economics)
Stockholm School of Economics

ISBN NR 91-7258-518-8

Keywords: random utility maximization, travel cost analysis, conditional logit, multinomial logit, logit, panel data, instrumental variables, Baltic Sea, eutrophication, water quality, mammography, taxes, eco-labeling, automobiles, accelerated vehicle retirement programs, scrappage

Distributed by:
EFI, Stockholm School of Economics, Box 6501, SE-133 83 Stockholm, Sweden

Printed by:
Elanders Gotab, Stockholm 1999

Then God said, "And now we will make human beings; they will be like us and resemble us. They will have powers over the fish, the birds, and all animals, domestic and wild, large and small."

Genesis 1:26

Contents

Preface

It is now a beautiful evening of the first day of July, and I have missed another of the deadlines I put down for myself. However, until quite recently I used to push the deadline by half a year at a time. Lately, I have moved it by one month at a time, and now, I am actually down to moving it by only days each time. Today, I *will* finish.

It is wonderful to have reached the goal. Finally, I will be able to put the thesis behind me, and move on. At the same time, I realize that what has made the travel (and the travail) worthwhile has been the journey more than the goal. It is often inspiring to do research. However, being a Ph.D. student also involves unheard of amounts of privation, frustration, desperation, exasperation, aberration, and dilapidation. The consolation, however, is the great friendship shown by so many people. I cannot thank everyone who has in some way or another assisted me with my work, since I have been told that the preface must not be longer than the rest of the thesis. If somebody feels forgotten, then be sure that you are not.

The list of "Thank-Yous" has to begin with my academic advisor, Professor Per-Olov Johansson. There is no doubt that all the four essays have gained immensely from his advice. He has also had the ability to push me along at crucial times during the work on the thesis. I seriously doubt that I would ever have reached the goal had it not been for those pushes. In addition, I have enjoyed having him as my advisor. Thank you P-O!

Professor Lars Bergman has been the second half of my advisory committee and has given valuable advice, especially on the first of the essays. Further, some of the ideas on eco-labeling originates with his note on the "green firm" and with the discussions we have had on why companies often care more about the environment than one would expect.

The person who first got me interested in environmental economics and travel cost analysis was Professor Bengt Kriström. Whether that is to his credit or discredit, I will let others judge. He has always been available as an extra advisor also after he left the Stockholm School of Economics for the Swedish University of Agricultural Science in Umeå. The discussant on my licentiate thesis was Professor Lars Hultkrantz, CTS, Dalarna University. He gave many useful comments at that occasion, and he has continued to do so since. He also supported me and cheered me up at one particular occasion when life as a Ph.D. student seemed unusually arduous. That meant a lot to me. The assistance of Tore Söderqvist has been vital in getting this thesis together, especially since he helped me buy a powerful enough computer. Tore has helped me many other times as well and has been a fun conference companion.

I have to extend a special thanks to Carin Blanksvärd and Pirjo Furtenbach. There is no doubt that my life during my years at the department would have been much more difficult without their help with administrative matters – large and small.

My fellow Ph.D. students. I will not even try to mention all names, but will only say that I owe you all for making my years at the Stockholm School of Economics so enjoyable and for creating a stimulating atmosphere at the Department of Economics.

I do have to mention Marcus Asplund, however. He has given me important insights into the field of industrial organization and has served as an econometrics consultant. Also, without knowing it, he solved a puzzle which might have taken me months to sort out. Just when I thought everything was almost finished, and that I only had to run a few more regressions – only slightly different than the old ones – everything went wrong. The results were simply incomprehensible. Marcus gave me the clue to the solution! Had he not, I might have had to go back to my initial half-year postponements of the dissertation. We also shared office during one year – Marcus' last year before his dissertation –, which was fun, if somewhat odd, considering his working habits. It was a little peculiar to find him sleeping in the office in the mornings.

My roommate during the rest of the time has been Richard Friberg. He has read all my papers, usually several versions, and has given excellent comments. He has served as a discussion partner. He has encouraged me, and on occasions he has discouraged me from undertaking new tasks instead of getting done with the thesis. Not least, he has put up with my views on politics, the Olympic Games, Göran Persson, the Swedish judicial system, Bill Clinton, Göran Kropp... To count Richard among ones friends is a privilege.

Numerous people have helped me to obtain data on sight depth, on the average life length of scrapped cars, on consumer interest rates, on the population of Swedish parishes, on the number of alcohol serving licenses per municipality, on gasoline prices, on.... I cannot mention them all, but would like to mention two. Dennis Nordberg at the Consumer Authority gave me access to their data free of charge. Bengt Lundgren, MD, President of the Swedish Society of Breast Imaging, helped me distribute and collect a survey to all the medical districts of Sweden. Essay 2 would not have been possible without his help.

As an economist, one is sometimes accused of having too narrow a view of how society works. There are a couple of people I would like to mention, not for assisting with the thesis, but rather for making me a better economist by extending my views beyond the utility-maximizing consumer and profit

maximizing firm. Conversations with Professor Erik Dahmén, at the lunch table, have hopefully prevented me from getting trapped in the toolbox. At the same lunch table, discussions with the late Professor Johan Myhrman provided me with a better understanding of economics. The team at the SSE Riga also broadened my views, not only by giving me a taste for Irish Whiskey. Continued association with Johan Stein over lunch now and then broadened them further. And of course, I must mention Lars Bergkvist, who has never let me get away with a shaky argument since we first met as undergraduate students at the Stockholm School of Economics twelve years ago.

I have received financial support from Stockholm School of Economics, the Swedish Environmental Protection Board, the Government Committee on Automobile Taxation (trafikbeskattningsutredningen) and Apoteksbolagets Fond för forskning i hälsoekonomi och socialfarmaci. This support is gratefully acknowledged.

Where would I be without my family? Since I can hardly see the difference between a BMW and a Lada, I ran into some trouble when I was going to do research about cars. My brother Henrik's close to encyclopedic knowledge in this area was indispensable. His assistance as a babysitter has also been extremely valuable when both my wife Ylva and I have had our schedules over-full. Ylva is my fiercest critic and staunchest supporter. She is my best friend and the center of gravity of my life. She also helped me with the mammography article since she now works at the breast cancer unit at one of Stockholm's hospitals. After a day of frustration at work, picking up my daughters Tove and Torun at the kindergarten has made me realize that I really have nothing to complain about in this world.

And now, as I write these final words of my thesis, I realize that the time is passed midnight, which means I missed another deadline.

Nacka, 2 July 1999
Mikael Sandström

Introduction and summary

This dissertation deals with two different aspects of public policy. A central question to ask when policy is formed is if the welfare of the individuals in society will be improved by the policy. To be able to answer that question, we must have methods to value the effects of public policy. This is the issue of the two first essays. These two essays each describe an application of the travel cost method, a method which is used to value commodities for which no markets exist.

A second important issue is how a given policy objective should best be met. To determine this, we must be able to evaluate the effects of different policy instruments. This is the issue of the third and fourth essays. In these essays, policy measures designed to improve the environmental characteristics of the automobile fleet are analyzed.

The four essays can be read independently. The notation may differ between the essay.

In the first section below, a brief introduction is given to the travel cost method and the two first essays are summarized. The next section is devoted to automobiles and the environment, and in particular to measures that may reduce the environmental harm of road traffic. The main features of the Swedish system for taxation of automobiles are described, as is the attempt to use "eco-labeling" to induce consumers to buy less polluting automobiles. That section also contains a summary of the third and fourth essays.

1. Two applications of the travel cost method

The first and second essays are application of the travel cost method, which is a method for valuing goods that are not traded on a market, i.e. that do not have market prices. Both of the essays use models that are based on the random utility maximization (RUM) model. Before going deeper into the theory behind this method, an attempt is made to place the essays in a broader context.

1.1 A background to the travel cost method

A main goal of economics is to develop methods that may be used to ascertain whether or not a given policy measure will improve the welfare of society. While acknowledging that this approach in fact implies strong value judgements, this usually involves the use of some form of social cost-benefit analysis. In other words, we attempt to translate changes in the welfare of each individual in society into a monetary measure, and then sum these over all individuals. Possibly, different weights are put on the welfare change of different individuals, to reflect for example equity considerations. However, the basic approach remains the same.

If the policy measure under analysis mainly affects goods that are traded on a market, we are often able to use market prices in our cost-benefit analysis. However, policy measures may also affect goods for which no market exists. Such non-market goods may be public goods, i.e. goods characterized by non-rivalry in consumption. By the term "non-rivalry" we mean that one individual's consumption of the good does not reduce another individual's consumption possibilities for the same good. An example of non-rivalry in consumption is air quality. If A "consumes" a given level of air quality, this does not reduce B's possibility to "consume" air quality. When our policy affects environmental amenities we are often forced to consider changes in the provision of public goods and are thus unable to use market prices. A non-market good may, on the other hand, be a private good that for some reason is not traded on a market. This is often the case when dealing with health economic issues. Since health care is greatly subsidized in most western countries, markets do not exist for many medical services. Alternatively, if markets do exist, the public involvement in the health care sector may mean that prices are severely distorted. Then prices from these markets cannot be used in a cost-benefit analysis, since they will not reflect the scarcity value of the resources used to produce these services.

When market prices cannot be used to evaluate the effects of a policy we must revert to some other valuation method. Attempts to fashion such methods have followed two main paths. Researchers following the first of these have tried to construct hypothetical markets for the non-market good that they want to value. The most common such method, the contingent valuation method, well illustrates this approach. While there are many variants of this method, the main features are the same. The subjects of the study are presented with a hypothetical situation constructed by the researcher. They are then posed with a question that is designed to elicit their valuation of this scenario. (See Carson, 1991, for a treatment of valuation methods using constructed markets and for references to studies using this approach.)

The travel cost method is the result of research following the other path. Instead of relying on hypothetical markets, this line of research has focused on how data on observed behavior may be used to make inference on individuals' valuation of non-market goods. Such valuation methods are usually termed either indirect valuation methods, since values are indirectly traced from individuals' behavior, or market based valuation methods, since they rely on data on market behavior.

The genesis of market based valuation methods may be traced to Adam Smith, who recognized that the wage of a laborer will in part reflect the disutility from his particular occupation.[1] Thus the wage in a dirty and hazardous occupation should be higher than the wage in another, otherwise comparable, but cleaner and safer occupation. In other words, an employer needs to pay a premium to induce workers to take an unsafe job. The wage, which is the price of a good which is traded on a market, namely labor, will thus in part reflect the value of goods that are not directly traded on any markets, namely health risks. In so-called "hedonic" wage models, the wages in occupations involving different levels of risk are analyzed to obtain a measure of the risk premium contained in the wage. This premium is then used as a measure of the workers' valuation of health risks. These values can then be used to evaluate policies that reduce such risks. (See Viscusi, 1992 for a summary of a number of such studies.)

The central insight that forms the foundation of hedonic wage models, and also of all other indirect valuation methods, is that a non-market good can often be interpreted as a characteristic of a market commodity (Bockstael and McConnell, 1993). In the case of occupational hazards, health risks may be viewed as an aspect of a market commodity – labor. A related valuation method focuses on the price of housing. Some public good, e.g. air quality, is then seen as a characteristic of housing, and price differences between housing in different areas, with different air quality, is used to infer the value house-owners put on this public good.

In a travel cost model, we view a trip as a commodity which is valued by the consumer. Non-market goods, such as air or water quality, that affect the site of destination of the trip, may then be viewed as characteristics of the trip. Thus, these non-market goods will influence the value to the consumer from undertaking

---

[1] "The wages of labour vary with the ease or hardship, the cleanliness or dirtiness, the honourableness or dishonourableness of the employment. Thus in most places, take the year round, a journeyman taylor earns less than a journeyman weaver. His work is much easier. A journeyman weaver earns less than a journeyman smith. His work is not always easier, but it is much cleanlier. A journeyman blacksmith, though an artificer, seldom earns so much in twelve hours as a collier, who is only a labourer, does in eight. His work is not quite so dirty, is less dangerous, and is carried out in day-light, and above ground. Honour makes a great part of the reward of all honourable professions. In point of pecuniary gain, all things considered, they are generally under-recompensed, as I shall endeavour to show by and by. Disgrace has the contrary effect. The trade of a butcher is a brutal and odious business; but it is in most places more profitable than the greater part of common trades. The most detestable of all employments, that of public executioner, is in proportion to the quantity of work done, better paid than any common trade whatever." (Smith, 1776, Book I, Chapter X, Part I.)

a trip. The keystone of the travel cost method is that it is possible to infer something about the value of such goods from studying travel behavior, or perhaps more precisely, from the costs incurred by consumers to visit a given site.

The earliest applications of the travel cost method usually aimed to value a specific site. The non-market good which was to be valued could be the availability of a national park. By relating the number of visitors from "zones" at an increasing distance from the site to the cost of travel to the site, a "demand curve" was constructed, and this demand curve was used to form welfare measures. This so-called zonal travel cost method relied on aggregate data on the number of visitors. In later years, however, the focus of the travel cost literature has shifted almost entirely to studies using micro data (Smith, 1989). In particular, theoretical developments and not least the increase in computer speed has shifted attention to models that explicitly take account of the discreteness of travel demand. In other words, modern travel cost models take account of the fact that a travel decision is better characterized as a choice between discrete alternatives than as a matter of choosing the optimal amount of a perfectly divisible good, as is assumed in the "text-book model" of consumer choice.

The theoretical underpinning of discrete choice travel cost models is the random utility maximization (RUM) model. In a RUM framework, we essentially regard a trip as a differentiated good. On a given choice occasion, the consumer chooses the alternative that gives him the highest level of utility. The basic feature of a RUM model is that we assume that the utility to an individual from a given choice may be separated into two parts. One part is assumed to depend on observable characteristics of the choice and/or of the individual, and one part is assumed to be non-observable. The latter part will appear random to an outside observer. There are different interpretations of the non-observable component of utility in a RUM model, but in the applications in this dissertation, we will assume that this part of utility is deterministic to the individual, and appears random only to the outside observer. (See Anderson, de Palma and Thisse, 1992, for an extensive treatment of the RUM model.)

Assuming a RUM model implies that we can obtain estimates of the probability that a randomly selected individual will choose a given alternative, in the travel cost setting, that she will undertake a certain trip. If the characteristics of an alternative – a trip – changes, then this probability will also change. From such changes in probabilities we can obtain measures of the corresponding change in welfare. That is the basic methodology employed in the first and second essays.

1.2 Summary of Essay 1 – Valuing the quality of water

The first essay employs a RUM travel cost model to estimate the benefits from improved water quality in the seas around Sweden. The model is estimated using the nested multinomial logit (NMNL) and conditional logit (CL) specifications. The background to this valuation study is that eutrophication, and the consequent decrease in water quality, has received increasing attention and is considered to be a serious environmental problem. Eutrophication is caused by an oversupply of nutrients, i.e. of nitrogen and phosphorus, to the water. Such emissions have increased manyfold during the 20th century, especially during its latter half. Among the negative effects of eutrophication is that the frequency of algal blooms increases. Also, the transparency of the water is reduced since primary production increases. This may lower the quality of seaside recreation.

The database from which data on travel behavior are obtained comprises of data from interviews with 2 000 – 4 000 Swedes per month during five years. We select from this database only travel to the coast during the three summer months, June – August. Since we use such a comprehensive data set, it is reasonable to think that it well covers this class of recreational travel. If we want to make policy simulations, it is important that all close substitutes to a site where water quality changes are included in the analysis. Since for Swedish consumers travel to the different sites along the Swedish coast are likely not to be close substitutes to any other good, that condition can reasonably be assumed to be satisfied.

To obtain any welfare estimates we must be able to link travel behavior to environmental quality in some meaningful way. In Essay 1, this link is sight depth. This seemingly simple measure of such a complex feature as water quality appears to serve well. In the travel cost model, this variable significantly affects travel behavior. To be able to evaluate policy measures to reduce eutrophication, we need to establish that sight depth is indeed linked to the supply of nutrients to the water. That such a link exists is well known. However, to perform policy simulations, we must obtain a quantitative measure of how the nutrients affect the sight depth. To accomplish this, a regression of sight depth on the concentration of phosphorus and nitrogen was run.

Any welfare measure from a travel cost study is sensitive to the specification of the travel cost variable. The standard method to measure travel cost is to multiply the travel distance by some calculated kilometer cost of travel, and then to add some estimate of the cost of travel time. Randall (1994) has criticized this approach. He points out that the travel cost is likely to consist of large subjective components. In the study described in Essay 1, a separate cost equation was used to obtain a measure of the out-of-pocket cost of travel per kilometer. To this cost, we added the cost of time as estimated by Johansson and Mortazawi (1995). They used the same database that is used for the travel cost study. Thus the data are allowed to determine how the cost of travel to a destination is calculated. Hopefully, this goes some way towards meeting Randall's criticism.

Two sets of policy simulations were undertaken. One set assumes a uniform change of the nutrient load along the entire Swedish coastline. The consumer surplus from a reduction of the nutrient load by 50 percent is estimated to be around 240 mSEK if the NMNL model is used and 540 mSEK if the CL model is used.[2] These values are rather low compared to the values obtained from contingent valuation studies of the same problem (Söderqvist, 1996), and correspond to but a small fraction of the estimated cost for such a clean-up (Gren, Elofsson and Jannke, 1995). Thus, the study presented in Essay 1 cannot be used as an argument for expensive measures to reduce the eutrophication of the seas around Sweden.

The other set of policy simulations assumes a change in the nutrient load in the Laholm Bay in southwest Sweden. The consumer surplus for a 50 percent reduction in the nutrient load in the bay is estimated to be 12 mSEK if the NMNL model is used and 32 mSEK if the CL model is used. These estimates are of the same order of magnitude as the estimated cost for such a reduction (Gren and Zylicz, 1993), but they are still smaller than the cost.

1.3 Summary of Essay 2 – Valuing a public health program

In Essay 2, the travel cost method is applied to a rather different problem – mammographic examinations. In many western countries, it is common to screen asymptomatic women for breast cancer, using mammography. The reason for this is that early detection of breast cancer significantly improves the chances of survival, and also makes it possible to use less invasive treatment, such as breast preserving surgery. In Sweden, as in several other countries, large publicly financed programs have been implemented to screen a large proportion of women in the age groups with the highest risk of breast cancer. While the potential benefits of such programs are substantial, so are the costs. Thus, valuation of mammography is an important task.

Previous to the study presented in this dissertation, only one application has been made of the travel cost method to a health economic problem, by Clarke (1996) who also studies mammography. Unlike Clarke (1996), who uses a traditional, ordinal model, the RUM model in the second essay is extended to an expected utility setting. In a traditional RUM model, it is assumed that the consumer is certain about the characteristics of each alternative. In many cases, this is a reasonable simplification. However, that is hardly the case when we are dealing with a health economic problem where the reduction of the risk of an adverse health outcome is pertinent.

---

[2] As of June 1999, SEK 1 ≈ USD 0.12 ≈ EURO 0.11.

A reason why few market based valuation studies have been performed in the health economic area may be that some doubts have existed about the applicability of such methods to issues where the good we want to value is of a truly essential nature. (See Johansson, 1997, for a discussion.) The theoretical conditions under which benefit estimates from a market based valuation study are valid are stronger than the conditions under which the Hicksian welfare measures are well defined. In particular, the market good that is used to infer the value of the non-market good must be non-essential. By "non-essential" we mean that it must be possible to fully compensate the individual if he completely refrains from using this market good. Thus, we must be careful when applying the travel cost method, not to national parks or other amenities that the individual may live without, but to essential goods such as health and the probability of death. However, in Essay 3 it is shown that the non-essentiality condition need not imply a severe limitation on the applicability of market based valuation methods to health economic problems.

In this essay, a somewhat novel approach is taken in formulating the travel cost variable. Instead of computing the travel cost, an auxiliary cost function is formulated. The cost of a trip is assumed to be a function of several variables, rather then just of the distance traveled, and this function is estimated simultaneously with the travel cost model.

The results from the estimation of the model are somewhat disappointing. In particular, the cost variable is not significantly different from zero at any ordinary level of significance. However, using the admittedly shaky estimated model to obtain benefit estimates gives quite reasonable results. The estimated benefits per mammography is between SEK 600 and approximately SEK 2 000. The cost per mammography has been estimated to be around SEK 1 000. While it would not be reasonable to draw any policy conclusions from the results, it is somewhat comforting to note that the costs and benefits are at least of the same order of magnitude.

One explanation of the poor showing of the model may be that the data on the women's travel to the mammographic examination is not detailed enough and that the variation in travel cost is too small. These problems may be remedied in future studies. Thus there is no reason why market based methods should not find health economic applications.

2. Automobiles, public policy and the environment

While the two first essays form attempts to measure the benefits of public policy, the third and fourth essays focus on what policy measures should be used to reach given objectives. They deal with different aspects of automobile taxation, and also with the Swedish system for eco-labeling of automobiles. (We discuss this term below.)

2.1 Policy measures to reduce pollution

Road transport is one of the most serious threats to our environment. For this reason, policies to reduce vehicle emissions have been introduced. While taxation of automobiles was originally introduced for fiscal reasons, attempts have been made to use taxation to reduce road traffic, and to alter the composition of the vehicle fleet in a "green" direction. In Sweden, both the sale and the ownership of automobiles have been taxed. In addition, automobile fuel is taxed. All of these taxes have been altered at different times in order to induce consumers to purchase less environmentally harmful vehicles. Naturally, many other forms of regulation have also been used, not least emission standards that regulate the maximum amount that an automobile may emit of certain substances. Here, we will however, only deal with two specific policies other than taxes, namely eco-labeling, and scrappage premiums.

We begin by considering eco-labeling schemes. In the traditional "tragedy-of-the-commons" story, individuals in society will not take into account the externalities caused by their behavior when they make their consumption decisions. Thus, we must use either direct regulations or economic incentives, e.g. Pigouvian taxes, to induce the consumers to internalize the environmental effects of their behavior. However, the increasing demand for products that are deemed to be less harmful to the environment bears witness to that there is more to this tale. One may speculate as to the reason for this demand for "green" products – if it is a consequence of altruism, of social pressure or some other factor – but it appears to be a fact that consumers care about the environmental characteristics of the products they buy.

Eco-labeling schemes have been implemented in a number of countries to make it easier for consumers to know which products are environmentally friendly. Under such a scheme, products that satisfy certain criteria are allowed to carry a label that is meant to indicate to the consumers that this product is better for the environment than products not carrying the label.[3] The authorizing body may be a government agency or a private organization, such as an environmental group. In Sweden, a government sponsored eco-labeling scheme for automobiles has been in operation since 1991. This scheme has also been associated with tax rebates for

[3] Whether eco-labeled products really are superior to other products with regard to environmental characteristics is an open question. However, that is a question of how well designed are the eco-labeling schemes and not an issue of how they affect consumer behavior, which is the focus of this discussion.

eco-labeled cars. The third essay deals with the effects on the composition of the car park from tax incentives and eco-labeling.

The observation that old cars pollute more than new cars, at least per kilometer driven, has been used as an argument to introduce policies that will reduce the life-length of automobiles. In particular, "accelerated vehicle retirement program" have been advocated. Under such a scheme, a scrappage premium is paid to car owners who scrap their cars. Usually, an accelerated vehicle retirement program is only in operation during a brief period of time. In Sweden, however, a scrappage premium program has been in operation since 1976. Scrappage premiums and other determinants of the life-length of automobiles are in the focus of the fourth essay.

2.2 Summary of Essay 3 – Automobile sales

In Essay 3, we focus on the composition of the car park, i.e. on what determines which car models consumers buy. Thus, we study the sale of new cars of different makes and models. The analysis is based on a model taken from the industrial organization literature. This implies that both the supply and demand sides of the market are modeled, which contrasts with most previous research on taxation. The model is adopted from Berry (1994), who utilizes a RUM model from which aggregate demand functions are constructed for each car model on the market. The producers of automobiles are modeled as price-setting producers of differentiated goods. In Essay 3, Berry's model is extended with an explicit model of how the conditional utility from choosing a given alternative is determined. Since an automobile is a consumer durable, we must consider the benefit stream from the car during its entire life. In particular, we must take into account that periodic taxes, such as a road tax, and "once off" taxes, such as a registration tax, might affect demand differently.

The model is estimated using panel data estimation techniques. A complication in models with imperfect competition is that price and quantities are determined simultaneously. Econometric methods that do not take account of this simultaneity may produce inconsistent estimates. Thus, the model is estimated using instrumental variables. To construct instruments, we use the predictions from the demand model on how different variables should interact.

The system by which eco-labeled cars have received preferential tax treatment has changed several times. This makes it possible to separate the "eco-labeling effect" from the effects of tax incentives. In other words, we are able to determine if consumers only respond to the tax incentives or if "green" considerations in themselves affect the consumption decision. The somewhat surprising conclusion is that the eco-labeling effect dominates. In fact, the tax credit to eco-labeled cars does not even have a significant effect on the demand for such cars. Naturally, it would be premature to conclude that we could abolish tax incentives and rely purely on eco-labeling to promote the "greening" of the vehicle fleet. One reason why the tax incentives do not have a significant effect may be that the incentives have been rather small. However, the result illustrates that "green preferences" among consumers may have a non-negligible effect on demand.

2.3 Summary of Essay 4 – Automobile scrappage

While the third essay dealt with data on the sale of different car models, Essay 4 focuses on the size and age of the whole car park. We use a version of the chain-of-replacement model (see e.g. Massé, 1962) to model the determinants of the life-length of automobiles, and of the size of the automobile park. This model is then used to study the effects of a temporary or a permanent scrappage scheme on the life-length of cars and the size of the car park. Since different taxes may affect the life-length differently, we also consider the effects of changes in taxes. An empirical application is then made to the scrappage rate of cars in Sweden.

The theoretical model predicts that an increase in the periodic tax would not affect the life-length of cars but would reduce the size of the automobile stock, while an increase in the registration tax would increase the life-length of cars and reduce the size of the car stock. A permanent scrappage premium would be exactly equivalent to a negative registration tax with regard to the life-length of automobiles. With regard to the size of the car stock, a permanent scrappage premium would not be equivalent to a negative sales tax, but would have the same qualitative effect. A temporary scrappage premium would have no permanent effects on the life-length of automobiles or on the size of the car stock. Rather, cars older than a certain age, which would depend on the size of the premium, would be scrapped and immediately replaced by new cars. The empirical application largely supports the predictions from the theoretical model.

It is concluded that the effects on emissions from the introduction of temporary or permanent scrappage premiums will be ambiguous. Since a permanent scrappage premium will increase the size of the car stock, the sign and size of the change in emissions will depend on how quickly the environmental characteristics of cars improve. A temporary scrappage premium will cause emissions initially to drop if new cars are cleaner than old ones. However, when the initial effect has worn off, the car park will be older than it would have been in the absence of the program. Thus, there may be an offsetting increase in emissions. Naturally, the increased production and scrappage of cars will also have negative environmental effects.

References

[1] Anderson, Simon P, André de Palma and Jacques-François Thisse (1992) *Discrete choice theory of product differentiation.* Cambridge, Mass. and London: MIT Press.

[2] Berry, Steven T. (1994) "Estimating discrete-choice models of product differentiation." *Journal of Economics.* 25(2):242-262.

[3] Bockstael, N. E. and K. E. McConnell (1993) "Public goods as characteristics of non-market commodities." *Economic Journal.* 103(September), 1244-1257.

[4] Carson, Richard T. (1991) "Constructed markets." in J.B. Braden and C.D. Kolstad, eds. *Measuring the demand for environmental quality.* Amsterdam and New York: North-Holland.

[5] Clarke, Philip M. (1996) "Cost-benefit analysis and mammographic screening: a travel cost approach." *Journal of Health Economics.* 17:767-787.

[6] Gren, Ing-Marie, Katarina Elofsson and Paul Jannke (1995) "Cost of nutrient reductions to the Baltic Sea." Beijer Discussion Paper Series, No. 70. Stockholm: Beijer International Institute of Ecological Economics, Royal Swedish Academy of Sciences.

[7] Gren, Ing-Marie, and Tomaz Zylicz (1993) "Cost-effectiveness of the Baltic Sea clean-up: Will wetlands reconcile efficiency with biodiversity?" Beijer Discussion Paper Series, No. 24. Stockholm: Beijer International Institute of Ecological Economics, Royal Swedish Academy of Sciences.

[8] Johansson, Per-Olov (1997) "On the use of market prices to evaluate medical treatments." *Journal of Health Economics.* 16:609-615.

[9] Massé, Pierre (1962) *Optimal investment decisions.* Englewood Cliffs, N.J.: Prentice-Hall, Inc.

[10] Randall, Alan (1994) "A difficulty with the travel cost method." *Land Economics.* 70:88-96.

[11] Smith, Adam (1776, 1993) *An inquiry into the nature and causes of the wealth of nations. A Selected Edition.* Oxford and New York: Oxford University Press.

[12] Smith, V. Kerry (1989) "Taking stock of progress with the travel cost recreation demand methods: Theory and implementation." *Marine resource economics.* 6:279-310.

[13] Söderqvist, Tore (1996) "Contingent valuation of a less eutrophicated Baltic Sea." Beijer Discussion Paper Series, No. 88. Stockholm: Beijer International Institute of Ecological Economics, Royal Swedish Academy of Sciences.

[14] Viscusi, W.K. (1992) *Fatal tradeoffs. Public and private responsibilities for risk.* New York and Oxford: Oxford University Press.

Included essays

ESSAY 1 – RECREATIONAL BENEFITS FROM IMPROVED WATER QUALITY

ESSAY 2 – CAN WE USE MARKET METHODS TO EVALUATE HEALTH RISKS?

ESSAY 3 – GREEN TAXES, ECO-LABELING AND THE AUTO MARKET

ESSAY 4 – THE ECONOMICS OF AUTOMOBILE SCRAPPAGE

Essay 1

Recreational benefits from improved water quality

Recreational benefits from improved water quality

F. Mikael Sandström*
Stockholm School of Economics

Abstract

In this paper, a random utility maximization (RUM) model of Swedish seaside recreation is used to estimate the benefits from reduced eutrophication of the seas around Sweden. Sight depth data from around the Swedish coast are used as a quality index related to eutrophication. The model is estimated using the nested multinomial logit (NMNL) and conditional logit (CL) specifications.

In order to test the relationship between this quality variable and the nutrient concentration in the water, a regression of sight depth on the concentration of phosphorus and nitrogen has been run. The results are used to make policy simulations.

Two sets of such simulations have been undertaken. One set assumes a uniform change of the nutrient load along the entire Swedish coastline. The consumer surplus from a reduction of the nutrient load by 50 percent is estimated to be around 240 mSEK if the NMNL model is used and 540 mSEK if the CL model is used.

The other set of policy simulations assumes a change in the nutrient load in the Laholm Bay in southwest Sweden. The consumer surplus for a 50 percent reduction in the nutrient load in the bay is estimated to be 12 mSEK if the NMNL model is used and 32 mSEK if the CL model is used.

---

* Assistance and valuable comments have been given by Lars Bergman, Per-Olov Johansson, Bengt Kriström and seminar participants at the Stockholm School of Economics, CTS, Dalarna University, the Department of Forest Economics of the Swedish University of Agricultural Sciences, Umeå and at the First World Congress of Environmental and Resource Economists, Venice. The analysis of the Tourism and Travel Database has been carried out in cooperation with, and with assistance from, Professor Lars Hultkrantz and Reza Mortazavi, CTS. The work would not have been possible without the scores of people who have kindly assisted me with data on water quality and other data. Financial support from the National Swedish Environmental Protection Board is gratefully acknowledged.

Introduction

Seas, lakes and rivers provide a large number of services. This paper focuses on one of these, namely their recreational function. The travel cost method is applied to estimate recreational benefits from water quality improvements.

We would suspect water quality changes caused by environmental degradation to have important effects on the recreationist's benefit from a trip to for example a lake or a seaside resort. Chemical substances or pathogenic bacteria in the water may have health consequences. High concentration of algae or other biological substances can be unpleasant. Oil spills may make bathing impossible. As a consequence, improved recreational quality may account for a substantial part of the benefits from water clean-up programs. Bockstael, Hanemann and Strand (1987) cite Freeman's (1979) assertion that over half the value from improved water quality is usually due to recreational values.

In some instances, water pollution will have only local effects. More often, however, entire aquatic systems are affected. Also, demand for recreation at a certain site cannot be studied in isolation. How the consumers' decisions are affected by quality changes at one or several sites will intimately depend on the availability of substitutes. For these reasons, a study of recreational costs or benefits from policies that may change water quality should ideally include 1) all sites that are affected by the proposed policy and 2) all sites that are possible substitutes to any site that is affected by the proposed policy.

Violating the first of these conditions will simply mean that we measure only part of the total recreational benefits, which may be what we are interested in, anyway. However, violations of the second conditions may have more important consequences. Interpreting the condition in the light of Hanemann and Morey (1992) gives some insight into why that is. They analyze the case where data is available only on a subset of the commodities that enter the individual's utility function. If the group of data commodities is, using the term of Blackorby, Primont and Russel (1978), recursively separable[1] from all other commodities, a set of demand functions can be constructed based only on the prices and quality aspects of the separable group of commodities. Terming such a demand system a partial demand system, a compensated (equivalent) variation measure based on this demand system is consequently called a partial compensated (equivalent) variation.

Hanemann and Morey (1992) show that this partial compensated variation is a lower bound on compensated variation calculated from the complete demand system. If the data commodities have zero income effects, then the two measures

[1] The separability assumed by Hanemann and Morey (1992) requires that if x is a vector of data commodities, with which a vector b of non-market goods is associated, and z is a vector of non-data commodities, with which a vector c of non-market goods is associated, then we should be able to represent a consumer's preferences with a utility function on the form: $u=u(u_d(x,b),z,c)$, where $u_d(x,b)$ is thus a "sub-utility" function.

will be identical. In other cases the size of the difference will roughly depend on the marginal rate of substitution between the data commodities and the other commodities – the smaller the degree of substitutability, the smaller the difference. If, on the other hand, some of the non-data commodities are close substitutes to the data commodities the divergence can be substantial.

In a recreational demand model, we would thus like to isolate a category of recreation for which we can reasonably assume that the individual would not easily substitute for example with other forms of recreational activities. If we think that seaside recreation on the French Mediterranean coast is quite similar to such recreation in Spain or Italy, then we would make a mistake by only studying sites on the French side of the boarders.

The present study makes use of the Tourism and Travel Database, the TDB, which is based on telephone interviews with a random sample of 2000-4000 Swedes each month. Information is assembled on recreational and business travel during each month (Swedeline, 1995). From the TDB, all data on summer recreation during the years 1990-1994 to the Swedish coast is selected. It is reasonable to expect that this would constitute a thorough coverage of this form of recreational travel.

It is also fair to believe that for Swedes, seaside recreation to sites at the Swedish coast is a group of commodities that has no close substitutes. Recreation by a lake is quite different in character. Traveling abroad can also reasonably be thought to be very different from domestic recreation. On the other hand, different Swedish coastal sites may be close substitutes to each other.

The seas around Sweden are the Skagerrak, the Kattegatt, the Öresund, and the Baltic Sea. Especially the last three of these have been heavily affected by eutrophication, which is in the focus of this study. Eutrophication is caused by an oversupply of nutrients to the water and has been suspected of causing changes in the macroalgal flora (Wennberg, 1987) and more frequent algal blooms (Granéli et alii, 1989). Also, a higher nutrient load decreases the transparency of the water by increasing primary production (Rosenberg, 1990). Eutrophication can thus be expected to have a negative impact on the quality of seaside recreation. A random utility maximization (RUM) model is used to evaluate hypothetical changes in the emissions of nutrients. A simple quality measure related to eutrophication, sight depth, is employed in the estimations. Even though this particular class of travel cost models has been widely used as an evaluation tool in the USA (Kaoru, 1995 and Hausman, Leonard and McFadden, 1995 are two examples) this is, to the knowledge of the author, the first European application of discrete choice models to recreational demand.

From the discussion above, it follows that to asses the recreational benefits from policy measures that have only a local effect we would still need to take account of recreation to most of the Swedish coast. To evaluate any policy measures that have recreational effects, similar problems are likely to occur. The method used in this study should thus have many applications.

Section 2 is a presentation of the model and the data. In any travel cost study, the definition of the travel cost variable will be of principal significance. Section 3 addresses this issue. To be able to use a recreational demand model to evaluate environmental change, another factor will also be crucially important. If the variable linking pollution to recreational behavior is not well chosen, the valuation exercise will be bootless. Firstly, we must be sure that the quality variable does in fact relate to the perceived quality of the site. Secondly, we need to establish the link between the quality variable and environmental variables that may be affected by policy, in this case, the link between nutrients in the water and sight depth. If behavior is affected by the quality variable, this is an indication that the first link does hold. In Section 4, a model is presented which tests and describes the second link in quantitative terms. Sections 5 and 6 present estimation results and the results from two policy simulations. The last section offers a few concluding remarks.

The model and the data[2]

In a RUM model, the individual is seen as choosing between a finite number, say N, of mutually exclusive goods (see, e.g., Smith, 1989, Bockstael, McConnell and Strand, 1991, Small and Rosen, 1981, Maddala, 1983 or McFadden, 1976). The alternatives are such that the individual either consumes a fixed quantity of a good, or does not consume it at all. In the travel cost framework, each alternative corresponds to a recreational site. The utility of an individual, conditional on choosing site j, is determined by a row vector of characteristics of the site, $\mathbf{b}_j$, the individual's income, y, the cost of visiting the site and of a random component, ε_j. Thus, his conditional utility can be written:

$$(1) \qquad \tilde{v}_j = \tilde{v}(y - p_j, \mathbf{b}_j, \varepsilon_j)$$

where use is made of Hanemann's (1982) result that in a pure discrete choice model, price and income must enter the utility function as y-p.

The individual will visit the site from which he derives the highest utility. In recreational demand models, it is usually assumed that $\tilde{v}_j$ is composed of a deterministic part, $V(y-p_j,\mathbf{b}_j)$, and a random term that is additive, and independently and identically distributed. The function V is commonly taken to be linear in the row vector $(y-p_j,\mathbf{b}_j)$.

Suppose the N alternatives can be divided into S subgroups, such that each alternative belongs to exactly one subgroup. The individual is then seen as choosing first among the subgroups, and then between the alternatives belonging to the chosen subgroup. The decisions are assumed to be independent over decision levels. In other words, different factors affect the decisions at the two

2 For a thorough discussion of the model used in this paper, and the underlying theoretical considerations, see Sandström (1996), accessible from: http://swopec.hhs.se/hastef/abs/hastef0121.htm.

levels. In the travel cost framework, the subgroups would be groups of similar sites, e.g., sites within the same region.

Suppose further that the vector of site characteristics can be divided into regional characteristics that are constant over sites belonging to the same region and site specific characteristics that vary also within the regions. Denote the set of choices j belonging to subgroup s by Σ_s, for s=(1,2,...,S) and write the deterministic part of the conditional utility of site $j \in \Sigma_s$, $V_{j,s}$, as:

(2) $$V_{j,s}=\mathbf{Y}_s\beta_a+\mathbf{Z}_{j,s}\beta_b,$$

where Y_s is the vector of regional characteristics of region s, $\mathbf{Z}_{j,s}$ are the characteristics of site j, and β_a and β_b are the associated column vectors of parameters. It can then be shown that the probability that the individual will choose subgroup s can be written:

(3) $$\pi_s = \frac{e^{\mathbf{Y}_s\beta_a + I_s(1-\sigma)}}{\sum_{t=1}^{S} e^{\mathbf{Y}_t\beta_a + I_t(1-\sigma)}}$$

where

$$I_t = \ln\left(\sum_{j\in\Sigma_t} e^{\mathbf{Z}_{j,t}\beta_b/(1-\sigma)}\right), \text{ for t=1,2,...,S.}$$

(see, e.g., Anderson, de Palma and Thisse, 1992 or McFadden, 1976)

Given that the individual has chosen subgroup s, the probability that he will choose alternative $j \in \Sigma_s$ can be written:

(5) $$\pi_{j|s} = \frac{e^{\mathbf{Z}_{j,s}\beta_b/(1-\sigma)}}{e^{I_s}} = \frac{e^{\mathbf{Z}_{j,s}\beta_b/(1-\sigma)}}{\sum_{k\in\Sigma_s} e^{\mathbf{Z}_{k,s}\beta_b/(1-\sigma)}}.$$

I_t is termed an inclusive value. It can be seen as a measure of the attractiveness of region t. The coefficient on the inclusive value, 1-σ, is often called the dissimilarity parameter, as it can be seen as a measure of the degree of similarity of alternatives belonging to each group. McFadden (1981) has shown that this coefficient must lie in the unit interval for the model to be consistent with stochastic utility maximization. A value close to zero implies great similarity between the alternatives in the subgroup, and a value close to one denotes little similarity. Restricting 1-σ to one reduces the NMNL model to the conditional logit model of McFadden (1973). In this paper, results are presented both from NMNL and conditional logit models.

The NMNL model can be estimated "from the bottom up". Thus, we first estimate (5) to obtain the parameters of the lowest level. Using these, the inclusive values of equation (4) are calculated. Treating the inclusive values as an independent variable, we can then estimate the top level using equation (3).

The function V is assumed to have the following form:

$$V_j=\beta_1 TTC_{j,s}+\beta_2 LNSIGHT_{j,s}+\beta_3 BEACH_{j,s}+\beta_4 LICENCE_{j,s}+\beta_5 SUN_s \tag{6}$$

where $TTC_{j,s}$ is the total cost of visiting site j in region s and $LNSIGHT_{j,s}$ is the natural logarithm of an index of sight depth at the site. $BEACH_{j,s}$ is the number of beaches, $LICENCE_{j,s}$ is the number of alcohol serving licenses per thousand inhabitants and SUN_s is the average hours of sunlight per month. The variable SUN_s is defined as a regional variable, while all the others are site specific.

The reason for taking the natural logarithm of sight depth is that it is reasonable to assume the marginal benefit of increased sight depth to be decreasing. The difference between one and two meters of sight depth is certainly going to be more noticeable than the difference between eleven and twelve meters.

To use the estimated model to evaluate policy, we need some measure of the change in welfare from a change in the characteristics of one or several sites. Assume that the marginal utility of income is approximately independent of prices and characteristics of the sites and that income effects are negligible. Further, assume that the recreational trip is non-essential and that the characteristics of the sites satisfy weak complementarity in the sense of Mäler (1974). (See also Bockstael and McConnell, 1993.) Under these conditions, Small and Rosen (1981) show that the following expression gives a valid measure of the compensating variation for a change in the characteristics of one or several sites:

$$CV(\mathbf{V}^1,\mathbf{V}^0)=\frac{1}{\lambda}\left[\ln\left(\sum_{j=1}^{N}e^{V_j}\right)\right]_{\mathbf{V}^0}^{\mathbf{V}^1} \tag{7}$$

where λ is the (constant) marginal utility of income and the vector $\mathbf{V}^i$, $i\in\{0,1\}$, is the vector of values of the deterministic part of utility of all alternatives, evaluated at the initial (i=0) and final (i=1) values of the site characteristics. Thus, $\mathbf{V}^i=\{V_1{}^i,V_2{}^i,\ldots V_N{}^i\}$, where $V_j{}^0$ ($V_j{}^1$) is deterministic utility from visiting site j with initial (final) values of site characteristics. Silberberg (1972) has shown that this line integral is path independent.

As can be seen from (3) and (5), the parameters of any characteristic that is constant over alternatives will cancel out of the expression of the probabilities and hence will not be possible to estimate. The individual's income will, naturally, be the same regardless of which site he chooses. However, marginal utility of income can still be obtained as the negative of the parameter on price, due to the Hanemann (1982) result referred to above.

There is no obvious way to decide what should constitute a "site" in a travel cost model. In the present study, the data dictates the definition. The lowest useful distinction of the destination of a trip in the TDB is municipality (kommun). Thus, each coastal municipality in Sweden is treated as a separate site.

The island of Gotland is excluded from the study, as are the four northern coastal counties of Sweden, that is, the coast of Norrland. The reason for not including Gotland is, firstly that it is an island, with the consequence that the travel patterns for visitors to Gotland are very different from those of visitors to the mainland coast, and secondly, that data on sight depth is not available. Norrland is excluded because, due to climate and demography, a very small fraction of seaside recreation takes place in this part of the country. In addition, water discharged from the large rivers of northern Sweden often makes the water in the northern Baltic Sea turbid. As a consequence, sight depth may be less functional as a quality measure in this area. In total, 66 municipalities are included in the study. In the NMNL model, the recreationist is assumed to first choose among four coastal regions, and then between the municipalities in the chosen region.

The characteristic of the sites in which we are most interested is the sight depth. The use of this variable as a measure of water quality, and how it is constructed, will be discussed in section 0. The number of beaches is included as an explanatory variable to control for differences in the general attractiveness of different sites, and for their varying size. Data was obtained from an ordinary road map (K.A.K. Bilatlas Sverige, 1992). The number of alcohol serving licenses per thousand inhabitants is included as a proxy for the nightlife at the site, to control for other possible attractions than just the sea. Data on the number of licenses were obtained from the Swedish Alcohol Inspection Board (Alkoholinspektionen, 1996), and population statistics were obtained from the Swedish Federation of Municipal Councils (Svenska Kommunförbundet, 1994). Naturally, the number of hours of sunlight will have an impact on site attractiveness. Data on this variable, which is calculated as an average for each month of the year over the years 1961-1990, were obtained from the Swedish Institute for Meteorology and Hydrology, SMHI (1994).

A key issue in travel cost research has been how travel cost should be measured. In particular, much attention has been given to the question of how the cost of travel time should be included in the analysis. (See, e.g., Randall, 1994, Englin and Shonkwiler, 1995 and Bockstael, Hanemann and Strand, 1987.) The next section describes how the travel cost variable is constructed.

Travel cost

The standard method of calculating travel cost in travel cost studies is to multiply the distance to the different sites with a hypothetical kilometer price, usually calculated on the basis of vehicle operating cost, gas price, etc. To this cost, a cost of travel time is added. In this study, a cost function is instead estimated based on stated cost (This approach is basically the same as that of Boonstra, 1993). A regression was run with stated cost as the dependent variable, and mode of travel, distance traveled, duration of the trip, choice of accommodation and number of household members participating in the trip as explanatory variables. Setting the model in the household production function framework provides a justification for this approach. (Sandström, 1996)

All data used in this regression are from the TDB, except the distance variable, which was obtained from a distance matrix from the Swedish Road Authority (Vägverket). The regression equation and estimation results are presented in appendices A and B, respectively.

The coefficients on the distance traveled for the different modes of transport are interpreted as the kilometer price of travel and were multiplied with the distance to each site to obtain the out of pocket cost of traveling to each site. To account for the cost of time, the results from Johansson and Mortazavi (1995) were used. They use the TDB to estimate the time cost of recreational travel for different modes of transport, using a RUM model. Their assumptions on the time it takes to travel a certain distance, and their estimates of the cost per hour of recreational travel were used to calculate the time cost to each destination. This cost was added to the out of pocket cost of travel to obtain total travel cost. The measure of travel time is thus obtained from the same data material that is used for the travel cost study and is based on revealed preferences.

A number of specifications of the model for the cost of a trip were tested. In some of the regressions, the cost of a trip was allowed to be non-linear in distance traveled. However, the results from these estimations were hard to interpret. With the assumption that the kilometer price of travel is independent of distance traveled, other alterations of the model do not appear to have any noteworthy effect on this kilometer price. The model thus appears reasonably robust to specification.

Hopefully, this approach leads us some way towards meeting the criticism expressed by Randall (1994). He claims that a fundamental problem with the travel cost method is that travel cost is unobservable. Using the stated cost and the detailed data available from the TDB may hopefully help us to come closer to a "true" measure of travel cost.

Water quality

The crucial factor in the travel cost model is the sight depth variable, as it provides the link with the nutrient load. In other words, it is the link between the environmental variables that can be influenced by policy, the concentration of phosphorus and nitrogen in the water, and recreational behavior. If this link cannot be modeled, the whole exercise will be meaningless. Data on sight depth had to be acquired from a number of sources, but turned out to be available for most stretches of the coast. (Details are available in Sandström, 1996.)

There is little doubt that a link does exist between nutrient load and sight depth. An increase in inflow of nutrients increases primary production, i.e., the content of organic material in the water, which reduces the transparency of the water (Rosenberg, Larsson and Edler, 1986). In an attempt to quantify this connection, a simple regression was run. Data on sight depth, water temperature and on the concentration of phosphorus and nitrogen from nine observation points in three municipalities in Östergötlands län (county) for the years 1975-1993 were obtained from Motala Ströms Vattenvårdsförbund, and the following model was estimated using OLS:

$$\text{(8)}\quad \underset{}{\text{LNSIGHT}} = \underset{(18.8)}{5.62} - \underset{(-4.47)}{0.0156} * \text{WTEMP} - \underset{(-12.7)}{0.625} * \ln(\text{TN}) - \underset{(-3.37)}{0.177} * \ln(\text{TP})$$

LNSIGHT is the natural logarithm of sight depth, WTEMP is water temperature, TN is total nitrogen content and TP is total phosphorus content. Water temperature was included to control for seasonal variations. The figures within parentheses are t-values calculated from White's consistent covariance matrix (See e.g. Greene, 1993).

The parameters on TN and TP have the expected negative signs and are significant at more than the 1 percent level of significance. Thus, the link between the quality index chosen for this study and the physical entities that can be affected by environmental policy measures seems to be validated.

A number of specifications of the sight depth model were estimated. The model with the highest adjusted R^2 was chosen. That turned out to be the log-log model. Such a relationship between sight depth and nitrogen and phosphorus also accords well with the results from marine biological research (Larsson).[3]

Estimation results

Separate models were estimated for, on the one hand, those traveling by car or public means of transportation and, on the other hand, for those traveling by private boat. For the first group, both a conditional logit and an NMNL model were estimated, while the small number of observations on boat recreationists allowed only the estimation of a conditional logit. The total size of the sample is 2 425 trips, of which 217 were made by private boat. The estimation results are presented in Table 1 below. Figures within parentheses are the asymptotic standard errors of the coefficients. In the top level of the NMNL model, the inclusive value is a random variable. The standard errors are adjusted to take account of this. Naturally, no inclusive value coefficient was estimated in the conditional logit models.

3 A Box-Cox regression (not reported) supports the choice of functional form. It gives a transformation parameter of 0.1.

Table 1 – Estimation results for the travel cost model

Variable	NMNL	Conditional logit	Conditional logit (boat)
Top decision level (X_S):			
IV	0.997	-	-
	(0.0632)		
SUN	0.283	0.299	0.612
	(0.0420)	(0.0444)	(0.424)
Lower decision level (Y_{jS}):			
TTC	-0.00108	-0.00101	-0.00213
	(0.0000639)	(0.0000309)	(0.000201)
LNSIGHT	0.269	0.575	-0.172
	(0.0702)	(0.0627)	(0.163)
BEACH	0.0164	0.0207	0.0420
	(0.00229)	(0.00199)	(0.0116)
LICENCE	0.583	0.544	0.478
	(0.0205)	(0.0186)	(0.0704)

All coefficients in the models for non-boat travel are significant at least at the 1 percent level of significance, and have expected signs. The inclusive value coefficient is almost exactly one, so that we would expect the conditional logit and NMNL models to be identical. All coefficients are indeed quite close, except for the sight depth coefficient. The results seem to suggest that the sight depth affects recreational behavior. However, the magnitude of this effect is rather uncertain.

In the boat model, the coefficients for TTC, BEACH and LICENCE are significant at least at the 1 percent level of significance, while LNSIGHT and SUN are not even significant at the 10 percent level. The coefficient for LNSIGHT does not have the expected sign. This may be due in part to the small sample size. However, it is also reasonable to believe that boat recreationists are less affected by water quality. In addition, boat recreation often implies travel through several different municipalities. It is therefore doubtful whether the data on this category of recreationists are reliable in this respect.

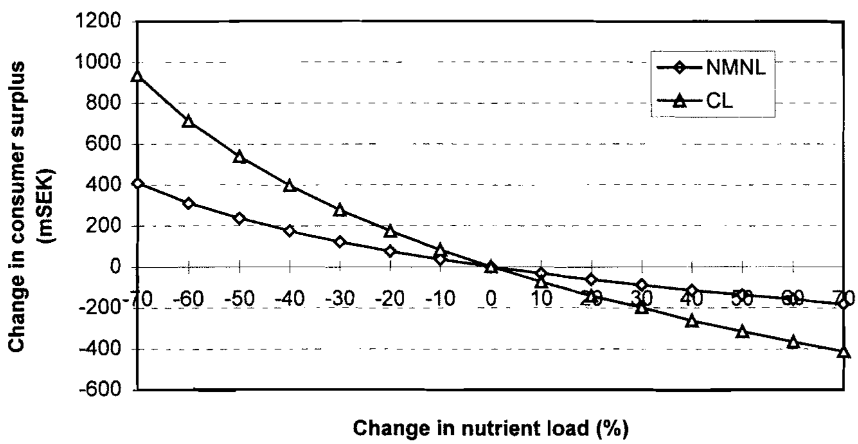
Policy simulations

Our models of the households' recreational behavior, and the sight depth model, make it possible to estimate the benefits from policy measures that change the nutrient levels in waters around the Swedish coast. Below, two experiments are presented. The first is an attempt to calculate the results of a uniform change in quality along the entire Swedish coast. The second deals with a change in quality in just one small region, the Laholm Bay.

The Laholm Bay has been seriously affected by eutrophication (See Rosenberg, Larsson and Edler, 1986 and Wennberg, 1987). The area is also one of the most popular seaside recreation areas in Sweden. Of the 2 208 trips in the sample made to all sites (boat recreationists excluded), 238 were made to the three municipalities around the Laholm Bay, i.e., over 10 percent.

The estimated sight depth equation (8) was used to construct a quality index for hypothetical changes in the nutrient load. The welfare change for each individual in the sample is then easily calculated using equation (7) and the parameter estimates of Table 1. The sum over all individuals in the sample was then multiplied by the inverse of the sampling ratio to obtain an estimate of the national welfare change. The resulting changes in consumers' surplus are displayed in Diagrams 1 and 2. The differences between the conditional logit (CL) and NMNL models are due to the different estimates of the sight depth coefficient.

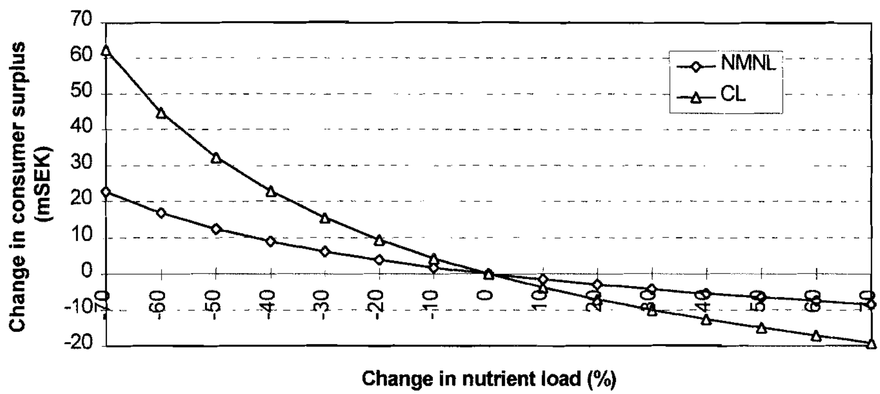
Diagram 1 – Policy experiment, entire Swedish coast



At a meeting in 1990 of the prime ministers of the states around the Baltic sea, a reduction target of 50 percent was set up for the nutrient load to the Baltic sea (Wulff and Niemi, 1992). The estimated recreational benefit from such a reduction is 240 mSEK and 540 mSEK in, respectively, the NMNL and conditional logit models. Even the highest of these figures correspond to less than ten percent of the estimated cost of such a reduction (Gren, Elofsson and Jannke, 1995). Considering this study alone, a drastic reduction of emissions of nutrients to the Baltic Sea can thus not be justified.

It should be noted that this model only captures part of the possible values from reduced eutrophication. Other forms of use values, e.g., for commercial fishing, are not taken into account. Non-use values are certainly not included in the benefit estimates. Also, day trips of less than 100 km are not included in the TDB. Further, the data precludes the estimation of a model that takes account of changes in the total number of trips (See Morey, 1991 for a discussion of this issue). However, since it is usually thought that recreational benefits constitute a major part of total benefits from improved water quality, the very low estimates are still striking. (See, e.g., Bockstael, Hanemann and Strand, 1987.)

Diagram 2 – Policy experiment, Laholm Bay



The benefit from a 50 percent reduction of the nutrient load to the Laholm Bay was estimated to be around 12 mSEK per year from the NMNL model, and 32 mSEK per year from the conditional logit model. The highest of these figures, 32 mSEK, comes close to the cost of a 50 percent reduction of the nutrient load in the Laholm Bay. Gren and Zylicz (1993) estimate that the cost of an efficient reduction by this proportion would be around 45 mSEK per year.

Concluding remarks

To be able to use models of recreational behavior to evaluate a policy measure, we must attempt to include not only recreation to the sites that are directly affected by the policy, but also to all sites that are likely to be close substitutes. The present study covers almost all domestic Swedish seaside recreation, and should thus meet up to this criterion.

To the knowledge of the author, the study is the first RUM travel cost model applied to European data. Previous European travel cost studies have either been single-site models (e.g., Bojö, 1985 and Strand, 1981), or have treated visits to all sites as a single good (e.g., Boonstra, 1993). Even though the issue dealt with is quite specific – eutrophication of the seas around Sweden – the method should have numerous other applications.

Perhaps the most important result in this paper is that the sight depth variable performs so well as a quality index. Also, instead of just assuming a relation between the quality variable and pollution, it has been shown that the link between this quality index and nutrient concentration can be established with simple econometric methods. Naturally, a more elaborate model could be developed.

In most travel cost studies, the travel cost variable is constructed by using some assessment of vehicle operating cost. In this study, a cost function is instead estimated, based on the stated cost for the trip. This approach goes some way towards solving the problem of defining the "true" cost of traveling to a site.

The benefit estimates from the trans-Baltic policy experiments are surprisingly low. Contingent valuation studies on reduced eutrophication of the Baltic sea yield willingness to pay estimates around ten times higher (Söderqvist, 1996). Inherent differences between the travel cost and the contingent valuation methods certainly account for some of the disparity. Still, more research is called for to explain the large divergence between the results.

References

[1] Alkoholinspektionen. (1996) Statistics over alcohol serving licenses by municipality. Stockholm.

[2] Amemiya, Takeshi (1973) "Regression analysis when the variance of the dependent variable is proportional to the square of its expectation." *Journal of the American Statistical Association*. 68:344:928-934.

[3] Anderson, Simon P., André de Palma and Jacques-François Thisse. (1992) *Discrete Choice Theory of Product Differentiation*. Cambridge, Mass. and London, England: MIT Press.

[4] Blackorby, Charles, Daniel Primont and R. Robert Russel (1978) *Duality, separability and functional structure: theory and economic applications*. New York: North Holland.

[5] Bockstael, Nancy E., W. Michael Hanemann and Ivar E. Strand, Jr. (1987) *Measuring the benefits of water quality improvements using recreational demand models*. Vol. II. Report to US EPA. Department of Agricultural and Resource Economics, University of Maryland.

[6] Bockstael, Nancy E. and Kenneth E. McConnell (1993) "Public goods as characteristics of non-market commodities." *The Economic Journal*. 103 (September) 1244-1257.

[7] Bockstael, Nancy E., Kenneth E. McConnell and Ivar E. Strand (1991) "Recreation." in J.B. Braden and C.D. Kolstad, eds. *Measuring the demand for environmental quality*. Amsterdam, New York, Oxford and Tokyo: North-Holland.

[8] Bojö, Jan (1985) with Lars Hultkrantz. *Kostnadsnyttoanalys av fjällnära skogar: fallet Vålådalen*. Research report, Stockholm School of Economics, Department of Economics. (English summary.)

[9] Boonstra, Froukje (1993) *Valuation of ski recreation in Sweden: A travel cost analysis*. Umeå Economic Studies No. 316, University of Umeå.

[10] Englin, Jeffrey and J.S. Shonkwiler. (1995) "Modeling recreational demand in the presence of unobservable travel cost: towards a travel price model." *Journal of Environmental Economics and Management*. 29:368-377.

[11] Freeman, A. M. III. (1979) *The benefits of air and water pollution control: A review and synthesis of recent estimates*. Technical report prepared for the Council on Environmental Quality, Washington, D.C.

[12] Granéli, Edna, Per Carlsson, Per Olsson, Bo Sundström, Wilhelm Granéli and Odd Lindahl. (1989) "From anoxia to fish poisoning: the last ten years of phytoplankton blooms in Swedish marine waters." in E.M. Cosper, V.M. Bricelj and E.J. Carpenter (eds.) *Novel phytoplankton blooms*. Coastal and estuarine studies, Springer-Verlag.

[13] Greene, William H. (1993) *Econometric analysis*. 2nd edition. New York: Macmillan.

[14] Gren, Ing-Marie, Katarina Elofsson and Paul Jannke. (1995) "Cost of nutrient reductions to the Baltic sea." Beijer Discussion Paper Series, No 70. Stockholm: Beijer International Institute of Ecological Economics, Royal Swedish Academy of Sciences.

[15] Gren, Ing-Marie and Tomaz Zylicz (1993) "Cost-effectiveness of the Baltic sea clean-up: Will wetlands reconcile efficiency with biodiversity?" ." Beijer Discussion Paper Series, No 24. Stockholm: Beijer International Institute of Ecological Economics, Royal Swedish Academy of Sciences.

[16] Hanemann, W. Michael (1982) "Applied welfare analysis with qualitative response models." Working paper, no 241, California Agricultural Experiment Station, Division of Agricultural Sciences, University of California, October 1982.

[17] Hanemann, W. Michael and Edward R. Morey. (1992) "Separability, partial demand systems, and consumer's surplus measures." *Journal of environmental economcs and management*. 22:3.

[18] Hausman, Jerry A., Gregory K. Leonard and Daniel McFadden. (1995) "A utility-consistent, combined discrete choice and count data model. Assessing

recreational use losses due to natural resource damage." *Journal of Public Economics*. 56:1-30.
[19] Johansson, Per and Reza Mortazavi. (1995) "The valuation of travel time: An empirical study using repeated samples of non-business trips." Umeå: Umeå University.
[20] Johansson, Per-Olov (1987) *The economic theory and measurement of environmental benefits*. Cambridge, UK: Cambridge University Press.
[21] *K.A.K. Bilatlas Sverige*. (1992) Stockholm: Nordstedts Förlag AB.
[22] Kaoru, Yoshiaki (1995) "Measuring marine recreation benefits of water quality improvements by the nested random utility model." *Resource and Energy Economics*. 17:119-136.
[23] Larsson, Ulf. "Fria vattenmassan." in "Miljökvalitetsbeskrivning av Norrtälje kommuns kustområde 1988-1992." Technical report 14. Stockholm: Department of system ecology, University of Stockholm.
[24] Maddala, G.S. (1983) Limited-dependent and qualitative variables in econometrics. Cambridge University Press.
[25] Mäler, Karl-Göran (1974) *Environmental economics: a theoretical inquiry*. Baltimore: John Hopkins University Press.
[26] McFadden, Daniel (1973) "Conditional logit analysis of qualitative choice behavior." in P. Zarembka, ed. *Frontiers in econometrics*. New York: Academic Press.
[27] McFadden, Daniel (1976) "Quantal choice analysis: a survey." *Annals of economic and social measurement*. 5:6:363-390.
[28] McFadden, Daniel (1981) "Econometric models of probabilistic choice." in C. Manski and D. McFadden eds., *Structural analysis of discrete data with econometric applications*. Cambridge, Mass.: MIT Press.
[29] Morey, Edward R. (1991) "What are per-trip consumer's surplus measures? And what, if anything, do they tell us about consumer's surplus?" Discussion papers in Economics, no. 91-6, Department of Economics, University of Colorado, Boulder.
[30] Randall, Alan (1994) "A difficulty with the travel cost method." *Land Economics*. 70:88-96.
[31] Rosenberg, Rutger (1990) "Negative oxygene trends in Swedish coastal bottom waters." *Marine Pollution Bulletine*. 21:7, pp 335-339.
[32] Rosenberg, Rutger (ed.), Ulf Larsson and Lars Edler (1986) *Eutrophication in marine waters surrounding Sweden*. Solna: National Swedish Environmental Protection Board, Report 3054.
[33] Sandström, Mikael (1996) "Recreational benefits from improved water quality: A random utility model of Swedish seaside recreation." Working

Paper no 121. Working paper series in economics and finance. Stockholm: Stockholm School of Economics.
Downloadable from: http://swopec.hhs.se/SWoPEc/hastef/abs/hastef0121.htm.

[34] Silberberg, E. (1972) "Duality and the many consumer's surpluses." *American Economic Review.* 62:942-952.

[35] Small, Kenneth A. and Harvey S. Rosen (1981) "Applied welfare economics with discrete choice models." *Econometrica.* 49:1:105-130.

[36] SMHI. (1994) "Sunshine Duration." Norrköping, Sweden: SMHI.

[37] Smith, V. Kerry (1989) "Taking stock of progress with the travel cost recreation demand methods: Theory and implementation." *Marine resource economics.* 6:279-310.

[38] Söderqvist, Tore. (1996) "Contingent valuation of a less eutrophic Baltic sea." Beijer Discussion Paper Series, No 88. Stockholm: Beijer International Institute of Ecological Economics, Royal Swedish Academy of Sciences.

[39] Strand, Jon (1981) "Beregning a samfunsøkonomisk verdi av fisket i Gaula-Vassdraget." Memorandum from Department of Economics, University of Oslo.

[40] Swedeline. (1995) *Turist och Resedatabasen.* (Tourism and Travel Database), Swedeline AB, Östersund, Sweden.

[41] Svenska Kommunförbundet (1994) *Folkmängd 1993-12-31.* Stockholm, Sweden.

[42] Wennberg, Tore. (1987) "Long-term changes in the composition and distribution of the macroalgal vegetation in the southern part of Laholm Bay, south-western Sweden, during the last thirty years." National Swedish Environmental Protection Board, report 3290, Solna. In Swedish with an English summary.

[43] Wulff and Niemi (1992) "Priorities for the restoration of the Baltic sea - A scientific perspective." *Ambio.* 21:2:193-195.

A. Sources of sight depth data:

[1] Göteborg och Bohus län. SMHI, Oceanografiska laboratoriet och Göteborgs och Bohus läns vattenvårdsförbund. Jan Szaron.

[2] Laholmsbukten. SMHI. Lars Edler.

[3] Skåne Danuta Lindsjö ed. (1993) Rapport från ett cyklopöga. Öresundsfonden. Mimeo.

[4] Haböbukten. SMHI, Oceanografiska laboratoriet. Jan Szaron.

[5] Blekinge län. Blekingekustens vattenvårdsförbund. Kalmarsundslaboratoriet, Högskolan i Kalmar, Inst. för naturvetenskap. Roland Enkvist.

[6] Kalmar län. Samordnad kustvattenkontroll i Kalmar län. Kalmarsundslaboratoriet, Högskolan i Kalmar, Inst. för naturvetenskap. Roland Enkvist.

[7] Östergötlands län. Motala ströms vattenvårdsförbund. Håkan Olsson, Länstyrelsen i Östergötlands län.

[8] Östergötlands län. Nyköpings kommun, Teknik, vattenlaboratoriet.

[9] Stockholms län och Södermanlands län. Ulf Larsson, Anders Sjösten and Katarina Skärlund. (1994) Himmerfjärdsundersökningen. Institutionen för systemekologi, Stockholms Universitet. Technical report no 15.

[10] Stockholms län. Stockholm Vatten. Christer Lännergren.

[11] Tierps kommun. ÅF-IPK. Karlit AB, recipientkontroll 1994. Barbro Grönberg, Länstyrelsen i Uppsala län.

[12] Älvkarleby kommun. Samordnad vattendragskontroll, 1993. Dalälvens vattenvårdsförening. Barbro Grönberg, Länstyrelsen i Uppsala län.

[13] Miscellaneous observations along the south and east coast. SMHI, Forskning och utveckling. Per Sandén.

B. Regression equation for the cost function

The regression equation used to obtain the kilometer cost of travel can be described as follows:

$$
\begin{aligned}
COST = {} & \gamma_0 + \gamma_{1,0} PERS + \sum_{k\in K} \gamma_{1,k} D_k PERS + \\
& + \gamma_{2,0} DIST + \sum_{m\in M} \gamma_{2,m} T_m DIST + \\
& + \gamma_{3,0} TNIGHTS + \sum_{n\in N} \gamma_{3,n} NIGHTS_n + \gamma_{4,0,0} DAYTRIP + \\
& + \sum_{k\in K} \left(\gamma_{4,k,0} D_k TNIGHTS + \sum_{n\in N} \gamma_{4,k,n} D_k NIGHTS_n \right) + \eta
\end{aligned}
\tag{A1}
$$

where COST is the total stated monetary cost of the trip, PERS is the number of persons participating, DIST is distance traveled and TNIGHTS the total number of trips away from home. $NIGHTS_n$ is the number of nights spent in accommodation category $n \in N$, N={hotel, family and friends, caravan, rented house}, and DAYTRIP is a dummy variable equal to one if the trip is a day-trip, zero otherwise. The D_k:s are dummy variables equal to one if the length of the stay away from home falls in the range $k \in K$, K={3-28,8-28,15-28}, and zero otherwise. The T_m:s are dummy variables equal to one if the mode of transport $m \in M$, M={public transport, private boat}, is chosen, and zero otherwise. The γ:s are parameters and η is a random term.

Two regressions were run, an ordinary least square regression, and a regression assuming the dependent variable to be lognormally distributed. (See Amemiya, 1973 for a discussion of the log-normal distribution.) The second of these was run to take account of the non-negativity constraint on the dependent variable. As can be seen below, the coefficient estimates hardly differ between the two regressions, while the t-values do. To take account of heteroskedasticity in the OLS regression, White's consistent covariance matrix was used to calculate the t-values. (See any econometrics textbook, e.g., Greene, 1993.) The sample used for estimation of the cost function comprised of 1 770 observations. Estimation was carried out in Limdep.

C. Regression results for the cost function

	OLS		Lognormal	
Variable	Coefficient	t-value	Coefficient	t-value
Constant	-795.65***	(-4.097)	-795.65***	(-27.336)
PERS	115.29***	(6.049)	115.29***	(27.219)
D_{3-28}*PERS	294.63***	(5.411)	294.63***	(3.805)
D_{8-28}*PERS	258.13*	(1.759)	258.13	(1.365)
D_{15-28}*PERS	-119.40	(-0.477)	-119.40	(-0.162)
$DIST_j$	2.0031***	(9.082)	2.0031***	(27.341)
T_P* $DIST_j$	-1.6639***	(-3.293)	-1.6640***	(-27.414)
T_B* $DIST_j$	0.74100	(0.944)	0.74101**	(2.489)
TNIGHTS	495.44***	(4.986)	495.44***	(23.991)
$NIGHTS_H$	516.94***	(2.879)	516.94**	(2.016)
$NIGHTS_F$	-135.26***	(-2.814)	-135.26***	(-25.256)
$NIGHTS_C$	-56.255	(-1.129)	-56.255***	(-20.996)
$NIGHTS_R$	412.04***	(2.899)	412.04***	(4.320)
DAYTRIP	530.88***	(2.960)	530.88***	(24.540)
D_{3-28}*TNIGHTS	-260.21***	(-3.215)	-260.21***	(-7.492)
D_{3-28}*$NIGHTS_H$	41.902	(0.205)	41.902	(0.117)
D_{3-28}*$NIGHTS_F$	85.885	(1.434)	85.885**	(2.253)
D_{3-28}*$NIGHTS_C$	158.95**	(2.410)	158.95***	(2.616)
D_{3-28}*$NIGHTS_R$	-98.684	(-0.676)	-98.684	(-0.843)
D_{8-28}*TNIGHTS	-100.10*	(-1.922)	-100.10**	(-2.019)
D_{8-28}*$NIGHTS_H$	28.027	(0.117)	28.027	(0.042)
D_{8-28}*$NIGHTS_F$	-39.016	(-0.742)	-39.016	(-0.775)
D_{8-28}*$NIGHTS_C$	0.35732	(0.006)	0.35732	(0.003)
D_{8-28}*$NIGHTS_R$	-41.421	(-0.696)	-41.421	(-0.342)
D_{15-28}*TNIGHTS	57.405	(1.026)	57.405	(0.543)
D_{15-28}*$NIGHTS_H$	-765.43***	(-2.702)	-765.43	(-0.614)
D_{15-28}*$NIGHTS_F$	93.500	(0.922)	93.500	(0.976)
D_{15-28}*$NIGHTS_C$	-56.559	(-0.976)	-56.559	(-0.274)
D_{15-28}*$NIGHTS_R$	-175.26**	(-2.188)	-175.26	(-0.709)
σ	-	-	1.3377***	(82.135)

Sub-indices H, F, C and R denote hotel, family and friends, caravan and rented house, respectively. The sub-indices on the D variables indicate the range of duration of the trips. P and B indicate travel by public means of transportation and private boat, respectively. The asterisks indicate significance at the 1 percent, 5 percent and 10 percent levels of significance, in a two tail t-test.

Essay 2

Can we use market methods to evaluate health risks?

Can we use market methods to evaluate health risks?

Mikael Sandström*
Stockholm School of Economics

Abstract

In recent years, a growing number of studies have used the contingent valuation method to value health care in monetary terms. Only very few studies employ methods based on market data. One reason for this may be that some doubts have existed about the applicability of such models in the field of health economics.

A theoretical model for the use of market prices to evaluate health care is discussed. The random utility approach is used to formulate a travel cost model. An auxiliary cost function is introduced, thus allowing the data to determine the cost of time, and the kilometer cost of transport.

An empirical application is made to mammography. The results leave many questions unanswered. However, the estimated welfare changes and the measures of the implicit values of a statistical life appear reasonable. A conclusion from the study is that market based valuation methods can fruitfully be applied to health economic problems.

*I am grateful for financial support from Apoteksbolagets Fond för forskning i hälsoekonomi och socialfarmaci. Assistance and valuable comments have been given by Per-Olov Johansson, Bengt Lundgren, Richard Friberg, Ylva Sandström and seminar participants at the Stockholm School of Economics.

1 Introduction

Methods to evaluate health care in monetary terms have become widely used. Most valuation studies in this field have relied on varieties of the contingent valuation method (CVM).[1] A problem with CVM studies is that they rely on hypothetical data, which has led some researchers to question the validity of this entire approach. (For an overview of the controversy surrounding the CVM, see the Journal of Economic Perspectives [13].) Even if we take a less extreme standpoint, it is clear that knowledge may be gained by applying other valuation methods. Not least, such knowledge may allow us to test the validity of CVM studies.

Valuation methods based on observed behavior have been extensively used to value environmental amenities. Perhaps the most widely used such method is the travel cost method (TCM). The fundament of the TCM is perhaps best understood in the light of its originally envisaged application: the valuation of national parks. In the United States, visitors pay only a relatively small fee to enter a national park. However, to enjoy the benefits of a national park, one must travel there. The cost of this journey can be substantial. In a famous letter to the US National Park Service, Harold Hotelling [10] proposed that this cost could be used to estimate a "demand curve" for a national park. This curve could then be used to infer what value the visitors derive from the park.

It is easy to see the analogy to valuation problems involving medical care. In many industrial countries, health care is provided at zero or nominal cost to the consumer. Still, to undergo medical treatment, the consumer will often have to incur some associated costs, such as travel to a hospital, loss of work income, etc. However, market based methods have not been widely used in the field of health economics. A few valuation studies of health risks have been undertaken (see e.g. Rosen [21] and Viscusi [27]), but to my knowledge, Clarke [5] is the only travel cost study of a health care program, previous to the present study. However, although Clarke analyzes mammographic screening, he does so in a deterministic setting, i.e. without analyzing the implications of risk and uncertainty for the assumptions of the model.

Health economic valuation problems involve commodities that affect the health of the individual, or the probability of death. This introduces two complications. Firstly, we must acknowledge that we may be dealing with truly essential goods, and not national parks and other amenities that the individual can live without. (See Johansson [11] for a discussion of how this issue may be addressed.) Secondly, the nature of many health economic problems forces us to deal explicitly with risk and uncertainty, and to carefully analyze what the implications are for the assumptions behind our model.

In Section 2 below, I will review the theory, and will argue that most restrictions that follow from the introduction of risk, uncertainty, and extreme choice situations apply regardless of which valuation method is used. The additional assumptions underlying market based valuation methods are not overly restrictive. In other words, market based

[1] See e.g. Johansson [12] for a summary of a few CVM studies of health economic problems.

valuation methods may be useful in health economic applications. However, additional care is sometimes warranted in the interpretation of the underlying assumptions.

The main difficulties lie instead in, firstly, translating the theoretical model into a model which can be used in applied work, and secondly, the empirical difficulties one is likely to encounter. The presence of risk and uncertainty may aggravate such difficulties. A large number of issues need to be dealt with in specifying and estimating a market based model and in using it to obtain welfare measures. As pointed out by Bockstael and McConnell [4], how the researcher tackles these issues will have important consequences for the estimated benefits. Of principal importance is, naturally, that some market good, or goods, is available, which is in some empirically traceable way associated with the good we wish to value.

These empirical complications are well illustrated by the application of a travel cost model described in Sections 3 and 4. I employ the model to evaluate mammographic screening, i.e. the same problem studied by Clarke [5]. Mammographic screening is undertaken to detect breast cancer in asymptomatic women. Breast cancer is the most frequent malignant tumour among women in many industrialized countries, including Sweden. It is a deadly disease. In fact, for Swedish women aged between 35 and 54 years, it is the most common cause of death. Early detection significantly increases the chance of survival. In addition, less invasive treatment is possible if the tumour is treated at an early stage (Lidbrink [16]). Thus, programs to screen large groups of women for breast cancer have large potential benefits, and have been implemented in many countries. However, the associated costs are also large. There is an ongoing discussion about whether or not the benefits of screening with mammography justify the costs, and also about how large those benefits are. (See e.g. Wright and Mueller [30]). It is thus important to evaluate screening programs.

One advantage of the method employed in the study presented in this paper, is that it takes its departure in the women's own valuation of the perceived benefits, and in the behavior resulting from this valuation. Some of the benefits from mammography can be objectively measured, such as the reductions in the risk of death etc. However, other benefits will be of a more subjective nature. Undergoing mammography may for example reduce anxiety. The model discussed below will potentially also capture such "subjective" values.

The paper is organized as follows. A theoretical model is formulated and analyzed in Section 2. In the following section, 3, the model is operationalized by specifying a probability distribution for unobserved individual characteristics, and by parameterizing the individual's indirect utility function. The data used are described and the estimation results are presented in Section 4. Also, estimates of the compensating variation for an elimination of mammography are calculated. These and the implicit values of a statistical life are analyzed, and compared to other studies, and to the cost of screening programs. In the final section, 5, the main conclusions are summarized.

2 A theoretical model

In this section, we will first consider the claim that market prices can be used to value non-market goods (Subsection 2.1). The implications for the applicability of the theory to health economic problems will be examined. In the following subsection (2.2), some additional complications following from the introduction of risk and uncertainty will be explored. (The implications for measures of the value of a statistical life are discussed in Appendix B). After that, in Subsection (2.3), the model will be extended to discrete choice situations. Finally (Subsection 2.4), the main conclusions on when it is appropriate to apply market based valuation methods are summarized.

2.1 Using market prices to value non-market goods

The theoretical foundations underlying the use of market prices to value changes in the provision of non-market commodities have been thoroughly explored. (See e.g. Bockstael and McConnell [3].) Thus, we will only briefly review the most important assumptions. We will also consider what limitations these assumptions place on the applicability of market based valuation methods to health economic problems.

Suppose the non-market commodity b is complementary with some market good, say x_1. Then, if b is increased[2], the compensated demand curve for x_1 will shift to the right. It is well known that, for a price change, the area to the left of the compensated demand curve, for the initial (final) utility level, between initial and final prices measures the compensating (equivalent) variation. Mäler [18] has suggested that, under two conditions, this result may be extended to changes affecting a non-market commodity, say a change from b_0 to b_1, by evaluating the following integrals:

$$\Delta W = \int_{p_1^0}^{\infty} x_1(\mathbf{p}^0, b^1, v^r)dp_1 - \int_{p_1^0}^{\infty} x_k(\mathbf{p}^0, b^0, v^r)dp_1 \tag{1}$$

where p_1^0 is the price of x_1, $x_1(\mathbf{p}^0, b^i, v^r)$ is the compensated demand for x_1, given $b^i, i = 0, 1$, and $\mathbf{p}^0$ is a vector of prices of x_1 and other market goods in the economy. The choice of a reference level of utility, v^r, determines if (1) will measure the compensating variation, CV, or the equivalent variation, EV. If we have $v^r = v^0$ it will measure the CV and if we have $v^r = v^1$ it will measure the EV. The welfare change from a change in b would thus be the change in the area to the left of the compensated demand for x_1.

The two conditions under which (1) will measure the CV or EV are perhaps best understood if we apply Shephard's lemma to the integrals in the expression. If we define the expenditure function $m(\mathbf{p}, b, v)$ as the lowest income necessary to obtain utility v given prices $\mathbf{p}$ and a level of the non-market commodity equal to b, then by Shephard's

[2] We will, throughout this paper, assume that b is a good, i.e. that utility is always non-decreasing in b.

lemma, $\partial m/\partial p_1 = x_1$. Thus:

$$\int_{p_1^0}^{\infty} x_1(\mathbf{p}^0, b^1, v^r)dp_1 - \int_{p_1^0}^{\infty} x_k(\mathbf{p}^0, b^0, v^r)dp_1 =$$

$$= \lim_{p_1 \to \infty} m\left(p_1, \widehat{\mathbf{p}}^0, b^1, v^r\right) - m\left(\mathbf{p}^0, b^1, v^r\right) -$$

$$- \lim_{p_1 \to \infty} m\left(p_1, \widehat{\mathbf{p}}^0, b^0, v^r\right) + m\left(\mathbf{p}^0, b^0, v^r\right) \quad (2)$$

where $\hat{\mathbf{p}}$ is the vector of all prices except p_1. It should be quite obvious that if the two limits are equal, expression (2) will define CV (for $v^r = v^0$) or EV (for $v^r = v^1$), since it will be equal to the change in income necessary to keep utility at the reference level when b changes.[3] For the limits to be equal, we need firstly for the limits to actually exist. In other words, the expenditure function must be bounded for $b \in \{b^0, b^1\}$, at both the initial and final levels of utility. This condition is usually termed "non-essentiality". Secondly, we need the limits to be equal. This will be the case if b is "weakly complementary" with x_1. This condition implies that if the consumer does not consume x_1, then he will not care about b.

If the weak complementarity and non-essentiality conditions are satisfied, we can be sure that CV and EV are well defined.[4] However, the reverse is not true in general. Thus, we may have situations when CV and/or EV are well defined, but when market based methods may not be used to measure these. In other words, our valuation problem must satisfy more stringent conditions for us to be able to use market based methods than when using the contingent valuation method. As we shall see in the next section, the uncertainty inherent in many health economic problems may add further complications.

2.2 Risk and uncertainty

Medical treatment usually does not ensure better health. Rather, it improves the probability that the individual will, in future, be in better health, or it improves the individual's survival probability. Thus, it is often necessary to deal explicitly with the question of risk and uncertainty. A complication is that a certain treatment may affect both the individual's health, given that he is alive, and the probability that he will be alive. In this section, a two-state, two-goods expected utility model is presented, in which one of the goods, a health good, has both these effects. (Details are provided in Appendix A.) Even in this simple model, complications arise, which are not encountered in the deterministic model of section 2.1. It is for example impossible to rule out discontinuous

[3] Notice that we have defined CV and EV in such a way that they will both have the same sign as the underlying change in utility.

[4] Naturally, we may have a situation where the limits in expression (2) exist for $v^r = v^0$, but not for $v^r = v^1$, or the other way around. In that case, only one of the measures will be well defined.

demand, even if the underlying probabilities and conditional utility functions satisfy reasonable regularity assumptions. However, as we will also see, the main results described in Section 2.1 continue to be valid in this setting also.

Call the health good $x_1 \geq 0$ and the other good - the numéraire good - $x_n \geq 0$. Call the price of the health good $p > 0$. The individual's budget restriction can thus be written $px_1 + x_n \leq y$, where $y > 0$ is income. Assume that health is a non-decreasing function, $h(x_1, b)$, of the health good and of a non-market good, b, the latter being exogenously determined. Also, we assume that the probability that the individual will survive is a non-decreasing function of x_1 and b. Call this probability $\pi(x_1, b) \in [0, 1]$. Assume that the individual's cardinal utility, given that he is alive, can be written as a function $u_1(x_n, h(x_1))$, which is strictly increasing in both of its arguments, and that the individual derives a non-negative utility from leaving a bequest. Call this bequest function $q(x_n)$, and assume that it is non-decreasing in x_n. Notice that since "health" has no obvious scale, the inclusion of a health function in a cardinal utility function must also involve some cardinality assumption about the health function.

With these assumptions, if the individual's preferences are rational, and if they satisfy the von Neuman-Morgenstern axioms, then these preferences have an extended expected utility representation, which may be written as:

$$U(x_n, x_1, b) = \pi(x_1, b) u_1(x_n, h(x_1, b)) + (1 - \pi(x_1, b)) q(x_n). \tag{3}$$

(See e.g. Mas-Colell, Whinston and Green[19] for a discussion of the theoretical basis of the expected utility model.) The individual will maximize this expected utility subject to the budget constraint and the two non-negativity constraints. The expected utility will be strictly increasing in the numéraire good, so we may write the utility maximization problem as:

$$\begin{aligned}\max_{x_1} U(y - px_1, x_1, b) = & \\ & = \pi(x_1, b) u_1(y - px_1, h(x_1, b)) + \\ & \qquad + (1 - \pi(x_1, b)) q(y - px_1)\end{aligned} \tag{4}$$

$$s.t. \quad 0 \leq x_1 \leq y/p$$

U is a continuous function, and the budget set is closed and non-empty, which implies that the maximization problem has at least one solution for all $\{p, b, y\}$. Thus, we may define an indirect expected utility function by:

$$v(p, b, y) = U(y - px_1^*, x_1^*, b) \tag{5}$$

where an asterisk refers to a solution to the maximization problem (4). Since we have assumed that u_1 is strictly increasing in x_n, and that q is non-decreasing in x_n, we can be sure that $v(p, b, y)$ will be strictly increasing in y. Thus, it can be inverted to yield the expenditure function:

$$y = m(p, b, v) \tag{6}$$

which gives the minimum income needed to give the individual expected utility v, given prices, p, and the level of the non-market good, b. Thus, we may define a welfare measure for a change in prices or in the provision of the non-market good, from, say $\{p^0, b^0\}$ to $\{p^1, b^1\}$, by:

$$\Delta W = m\left(p^1, b^1, v^r\right) - m\left(p^0, b^0, v^r\right) \tag{7}$$

where v^r is the reference level of (expected) utility. ΔW is thus the change in income needed to keep utility constant when p and b change. Using initial utility as the reference, i.e. setting $v^r = v^0$, this expression defines CV. Congruently, using final utility as the reference, i.e. setting $v^r = v^0$, it defines EV. Notice that (7) defines a state-independent compensated or equivalent variation, since the expenditure function (6) is the expenditure needed to obtain a given value of the function (5), i.e. of expected utility. Thus, the value of (6) will be deterministic, and the welfare measure in (7) will have the same sign as the underlying change in utility.[5] Obviously, for (7) to be defined, the expenditure function must be bounded at the reference level of utility for both the initial and final levels of the non-market commodity. In other words, v^r must be attainable with some finite level of income given either b^0 or b^1. This far, all results are exactly the same as in the deterministic model discussed in the last subsection.

Now, let us assume that we are considering changes which are such that we may reasonably assume that CV and EV are defined. A complication with the utility function in (3) is that even if we assume that the functions u_1, q, h and π are "well behaved"[6], we cannot be sure that U is strictly quasi-concave. We know that in a maximization problem (given that the solution set is non-empty), sufficient conditions for the solution to be continuous in the exogenous variables, in our case p, b and y, are that the value function is continuous in these variables, and that it is strictly quasi-concave in the decision variable(s), in our case in x_1. In other words, if U is not strictly quasi-concave, we cannot be sure that demand is continuous.

In fact, it can be shown that even when the underlying functions are "well behaved", demand may be discontinuous. As is shown in Appendix A, a sufficient condition for the expected utility function to be strictly quasi-concave is:

$$2\pi \frac{\partial^2 u_1}{\partial x_n \partial h} p - 2\frac{1}{\pi}\frac{\partial \pi}{\partial x_1}\frac{\partial q}{\partial x_n} - \pi \frac{\partial u_1}{\partial h}\frac{\partial^2 h}{\partial x_1^2} - \frac{\partial^2 \pi}{\partial x_1^2}\left(u_1 - q\right) > 0 \tag{8}$$

[5] See e.g. Johansson [12] for a discussion of the difference between a state-independent CV or EV, which are sign-preserving welfare measures, and the expected CV and EV, which are not always sign preserving.

[6] In Appendix A we define what we mean by "well behaved".

where the arguments of u_1, q, π and h have been suppressed. Thus, conditions that might give rise to situations with discontinuous demand would be if, for some values of x_1 and b, the marginal effect of a change in x_1 on the survival probability or on the health function were increasing. As will be further discussed in Subsection (2.3), such conditions cannot be ruled out. In addition, if the cross derivative of u_1 is negative, i.e. if health and the numéraire good are substitutes, then the possibility of discontinuous demand cannot be ignored.

However, even if demand is discontinuous, it will be piecewise continuous. (See Appendix A.) Further, using essentially the same approach as Small and Rosen [23], it can be shown that the expenditure function will still be continuous everywhere, and differentiable everywhere except at prices such that the individual is indifferent between two different consumption bundles. At such prices, however, it will still be left and right differentiable and bounded, and thus integrable.

To see the implications of this finding, let us for expositional purposes confine ourselves to the case where demand is discontinuous at one price only. Call this price $\widetilde{p}\left(b, v^r\right)$. Notice that this price will depend on the level of the non-market good, as well as on the reference level of utility. Since, by assumption, this is the only price at which demand is discontinuous, we may write the compensated demand, given $p > \widetilde{p}\left(b, v^r\right)$ and given $p < \widetilde{p}\left(b, v^r\right)$, respectively, as two functions, say, $\widetilde{x}_1^A\left(p, b, v^r\right)$ and $\widetilde{x}_1^B\left(p, b, v^r\right)$. By assumption, the consumer will be indifferent between $\widetilde{x}_1^A\left[\widetilde{p}\left(b, v^r\right), b, v^r\right]$ and $\widetilde{x}_1^B\left[\widetilde{p}\left(b, v^r\right), b, v^r\right]$. (Also by assumption, $\widetilde{x}_1^A\left[\widetilde{p}\left(b, v^r\right), b, v^r\right] < \widetilde{x}_1^B\left[\widetilde{p}\left(b, v^r\right), b, v^r\right]$, otherwise demand would be continuous.)

Thus, we may write the compensated demand as:

$$x_1(p, b, v^r) = \begin{cases} \widetilde{x}_1^A\left(p, b, v^r\right) & \text{if} \quad p > \widetilde{p}\left(b, v^r\right) \\ \widetilde{x}_1^B\left(p, b, v^r\right) & \text{if} \quad p < \widetilde{p}\left(b, v^r\right) \end{cases} \tag{9}$$

where the compensated demand function is thus undefined at $\widetilde{p}\left(b, v^r\right)$. Applying Shephard's lemma to each of these functions, the derivative of the expenditure function can be written:

$$\frac{\partial m(p, b, v^r)}{\partial p} = \begin{cases} \widetilde{x}_1^A\left(p, b, v^r\right) & \text{if} \quad p \geq \widetilde{p}\left(b, v^r\right) \\ \widetilde{x}_1^B\left(p, b, v^r\right) & \text{if} \quad p \leq \widetilde{p}\left(b, v^r\right) \end{cases} \tag{10}$$

where the upper and lower expressions are interpreted as right and left derivatives, respectively, at $\widetilde{p}\left(b, v^r\right)$.

If we consider a change in b, keeping p constant, we may thus use this result to evaluate the integrals in expression (1):

$$\begin{aligned} \int_{p^0}^{\infty} x_1(p^0, b^1, v^r)dp - \int_{p^0}^{\infty} x_1(p^0, b^0, v^r)dp = \\ = \lim_{p \to \infty} m(p, b^1, v^r) - m\left(p^0, b^1, v^r\right) - \\ - \lim_{p \to \infty} m(p, b^0, v^r) + m\left(p^0, b^0, v^r\right) \end{aligned} \tag{11}$$

Under the same assumptions as in the non-stochastic case, the two limits in this expression will have the same value, and thus cancel out, i.e. if b is weakly complementary with x^1, and if x^1 is non-essential. Thus, under these two assumptions, the change in the area to the left of the compensated demand curves will measure the CV or the EV from a change in the provision of the non-market good, b, depending on if we use initial or final utility as our reference. In other words, even if demand for the health good is discontinuous, we can still rely on the results from the deterministic case. However, the empirical analysis may be more complicated.

2.3 Discrete goods, the LRUM model

In many situations medical treatment or health enhancing behavior is best characterized as choices between discrete alternatives. In other words, it is often useful to model a health good as a good that is consumed in a fixed quantity, or not at all. There are several reasons why that may be the case. Sometimes, the relationship between the quantity of the health good consumed and the health benefit will be such that, even though the health good is perfectly divisible, there is a fixed quantity that effects the desired health improvement. Smaller amounts may have no effect, and larger amounts may be outright harmful. Health is improved when you take your medicine, but no additional benefit is obtained by doubling the dose.

In fact, this can be viewed as a special case of the model described in the previous section (altering it slightly to allow for discontinuous derivatives of $\pi(x_1, b)$ and $h(x_1, b)$). Remember from the inequality in (8) that if, for some levels of x_n and x_1, the marginal increase in the probability of survival, or in the health function, is increasing in the health good[7], then we must consider the possibility that demand is discontinuous. In many cases, we may expect that the marginal increase in the probability of survival and in health is close to zero for quantities of the health good below and above certain levels.[8] An extreme case giving rise to a pure discrete choice situation occurs when the marginal effects of the health good are zero everywhere except at exactly one value, say $x_1 = x_1^*$. Then, naturally, all consumers will choose to consume either $x_1 = x_1^*$ or $x_1 = 0$. In other words, there is no point in undergoing a treatment below some critical level, x_1^*, and no, or very small, additional benefit from undergoing more than this amount of the treatment. Thus, considering a single consumer, we may expect demand to be zero at prices over some "switch price" at which demand jumps to this quantity, and then remains there at all lower prices. Thus, the discrete choice model is really a straight forward extension of the continuous model.

An example of such a treatment is a mammographic examination. While an examination is, obviously, discrete in nature, a woman may well vary the frequency of examinations. However, there is no evidence of additional benefit from undergoing ex-

[7] I.e. if $\partial^2\pi(x_1, b)/\partial x_1^2 > 0$ or $\partial^2 h(x_1, b)/\partial x_1^2 > 0$.

[8] Naturally, we must replace the derivatives with differences if the functions are not continuously differentiable.

aminations at a closer interval than about 18 months or two years.[8] Thus, considering a period of one or two years, it is quite reasonable to view the decision of undergoing a mammographic examination as a pure discrete choice situation.

Below, we will use mammography as an example to illustrate some complications with the standard discrete choice formulation, when we set the model in an expected utility framework. Let us first consider the effects on the woman's utility from undergoing mammography. As has been discussed above, a medical treatment will often affect both the individual's survival probability and her health. In the case of mammographic screening, the risk of breast cancer is not reduced[9]. In fact, the radiation to which the patient is subjected causes a slightly increased risk. However, this increase is minuscule compared to the total risk of breast cancer, and will be ignored throughout this paper. The relative increase of the risk is in the order of one tenth of a percent. (Lidbrink [16].) The main benefit is rather that a cancer, should it develop, is discovered earlier, thus reducing the risk of death. In addition, early detection makes it possible to use less invasive treatment, such as breast preserving surgery. In analyzing mammographic screening we thus need to consider at least three future states of the world:

1. The case when the woman does not have breast cancer and stays alive. Call the probability of this contingency $\pi_1(\mathbf{s})$, where $\mathbf{s}$ is a column vector of characteristics of the individual. Notice that this probability does not depend on the health good, or on the non-market good, b. Call health in this state $h_1(x_1, b, \mathbf{s})$. We allow this function to depend on the health good and the non-market good, since it is conceivable that a health check, such as a mammographic screening, reduces anxiety, and thus the individual's perception of her health.

2. The case when the woman has breast cancer, but survives. We call the probability of this contingency, given that the woman has breast cancer, $\pi_2(x_1, b, \mathbf{s})$. This probability depends on x_1, and on b, and is non-decreasing in both. Health in this state is $h_2(x_1, b, \mathbf{s})$, which is also non-decreasing in x_1 and b. Naturally, the unconditional probability that the woman will have breast cancer and survive is $[1 - \pi_1(\mathbf{s})] \cdot \pi_2(x_1, b, \mathbf{s})$.

3. The case when the woman dies, from her breast cancer, or from some other cause. The probability of death is $[1 - \pi_1(\mathbf{s})] \cdot [1 - \pi_2(x_1, b, \mathbf{s})]$. To simplify the exposition, we will disregard any possible utility from leaving a bequest, and normalize by setting the utility from being dead to zero.

[9] This is not entirely true. A mammographic examination may discover "cancer in situ", i.e. a pre-stadium of cancer. If discovered, a cancer in situ is removed by surgery. Sometimes, the patient also undergoes radiotherapy. To somebody who is not a medical specialist, e.g. to an economist, this distinction is not of consequence. We might as well treat cancer in situ as an early discovered cancer.

Assuming that the consumer's preferences meet with von Neuman's and Morgenstern's axioms, we may thus specify an extended expected utility representation of these as:

$$U(x_n, x_1, b, \mathbf{s}) = \pi_1(\mathbf{s})u_1\left[x_n, h_1\left(x_1, b, \mathbf{s}\right), \mathbf{s}\right] + \\ + \left[1 - \pi_1(\mathbf{s})\right]\pi_2\left(x_1, b, \mathbf{s}\right)u_2\left[x_n, h_2\left(x_1, b, \mathbf{s}\right), \mathbf{s}\right] \quad (12)$$

We will now make use of the fact that we have assumed that x_1 is consumed in a fixed quantity or not at all. We normalize this quantity to one. Thus $x_1 \in \{0, 1\}$. Call the price of the health good p and the income of the individual y. In this setting, our consumer thus has a choice between two consumption bundles: $\{x_n, x_1\} \in \{\{y, 0\}, \{y - p, 1\}\}$. One way of viewing the problem is to consider it as a choice between two lotteries.

The expected utility from the lottery defined by $x_1 = 1$ can then be written as:

$$U(y - p, 1, b, \mathbf{s}) = \pi_1(\mathbf{s})u_1\left[y - p, h_1\left(1, b, \mathbf{s}\right), \mathbf{s}\right] + \\ + \left[1 - \pi_1(\mathbf{s})\right]\pi_2\left(1, b, \mathbf{s}\right)u_2\left[y - p, h_2\left(1, b, \mathbf{s}\right), \mathbf{s}\right] \quad (13)$$

Congruently, the expected utility from the lottery defined by $x_1 = 0$ may be written as:

$$U(y, 0, b, \mathbf{s}) = \pi_1(\mathbf{s})u_1\left[y, h_1\left(0, b, \mathbf{s}\right), \mathbf{s}\right] + \\ + \left[1 - \pi_1(\mathbf{s})\right]\pi_2\left(0, b, \mathbf{s}\right)u_2\left[y, h_2\left(0, b, \mathbf{s}\right), \mathbf{s}\right] \quad (14)$$

Apart from individual characteristics and the non-market good, these two functions are functions of price and income only. Thus, we may interpret them as (conditional) indirect utility functions.

In this setting, the weak complementarity assumption amounts to an assumption that:

$$\frac{\partial U|_{x_1=0}}{\partial b} = 0, \forall b$$

As will be discussed below, this should be a reasonably unproblematic assumption, since it is unlikely that the consumer will care about the quality of a treatment that she does not undergo. We will thus assume that this condition is satisfied.

We may then write the conditional indirect utility functions as:

$$\begin{aligned} \widetilde{U}_1(y - p, b, \mathbf{s}) &= U(y - p, 1, b, \mathbf{s}) \\ \widetilde{U}_0(y, \mathbf{s}) &= U(y, 0, b, \mathbf{s}) \end{aligned}$$

The usual way to proceed from this point is to make two assumptions, namely:

1. That the vector $\mathbf{s}$ may be decomposed into a vector, say $\mathbf{r}$, of observable individual characteristics, and a non-observable scalar, ε, which is deterministic, but appears random to an outside observer.

2. That the conditional indirect utility functions are additively separable into two parts, the first of which depends only on $\mathbf{r}$, y, p and b. (By the first assumption, only the conditional utility, given $x_1 = 1$, will depend on b.) The second part is assumed to be linear in ε. We will call the first part "observable utility", and call ε "unobservable utility".

If we call the observable utility given $x_1 = 0$ $\widetilde{V}_0(\mathbf{r}, y)$, and observable utility given $x_1 = 1$ $\widetilde{V}_1(\mathbf{r}, y - p, b)$, we may then write:

$$\widetilde{V}_1(\mathbf{r}, y - p, b) - \widetilde{V}_0(\mathbf{r}, y) = \varepsilon \tag{15}$$

where the scale of ε thus determines the scale of the utility representation. This is the standard linear random utility model, LRUM. (See Anderson, de Palma and Thisse [1] for a discussion of the RUM and LRUM models.) By appropriate normalization, we may then write the utility from undergoing mammography as:

$$\widetilde{U}_1(y - p, b, \mathbf{s}) = \widetilde{V}_1(\mathbf{r}, y - p, b) - \varepsilon \tag{16}$$

and the utility from not undergoing mammography as:

$$\widetilde{U}_0(y, \mathbf{s}) = \widetilde{V}_0(\mathbf{r}, y) \tag{17}$$

In the usual framework these assumptions are perhaps quite innocuous. At least, they are no worse than the usual assumptions behind any model used in applied work. However, in our expected utility model, the implications of this specification are more troublesome. To see this, consider the difference between expressions (13) and (14):

$$\begin{aligned}
&\widetilde{U}_1(y - p, b, \mathbf{s}) - \widetilde{U}_0(y, \mathbf{s}) = \\
&\qquad = \pi_1(\mathbf{s}) u_1^1(y - p, \mathbf{s}) + [1 - \pi_1(\mathbf{s})]\, \pi_2^1(\mathbf{s})\, u_2^1(y - p, \mathbf{s}) - \\
&\qquad\qquad - \pi_1(\mathbf{s}) u_1^0(y, \mathbf{s}) - [1 - \pi_1(\mathbf{s})]\, \pi_2^0(\mathbf{s})\, u_2^0(y, \mathbf{s})
\end{aligned} \tag{18}$$

where we have left out b and x_1, and have used superindices to indicate that the functions are conditional upon $x_1 = 0$ and $x_1 = 1$, respectively. Rearrange to obtain:

$$\begin{aligned}
&\widetilde{U}_1(y - p, b, \mathbf{s}) - \widetilde{U}_0(y, \mathbf{s}) = \\
&\qquad = \pi_1(\mathbf{s}) \left\{ \begin{array}{c} u_1^1(y - p, \mathbf{s}) - u_1^0(y, \mathbf{s}) - \\ - \left[\pi_2^1(\mathbf{s})\, u_2^1(y - p, \mathbf{s}) - \pi_2^0(\mathbf{s})\, u_2^0(y, \mathbf{s})\right] \end{array} \right\} + \\
&\qquad\qquad + \pi_2^1(\mathbf{s})\, u_2^1(y - p, \mathbf{s}) - \pi_2^0(\mathbf{s})\, u_2^0(y, \mathbf{s})
\end{aligned} \tag{19}$$

One natural interpretation of the unobservable part of utility from a mammographic screening is to see it as something which reflects unobservable differences in the (objective or subjective) risk of breast cancer. Thus, ε would influence $\pi_1(\mathbf{s})$. But from (19) it is clear that for this term to be linear, we would need to assume that the terms within the curly brackets are constant, which is hardly reasonable. In fact, it is hard to see any reasonable assumption that would allow ε to enter (19) linearly. Naturally, we may still justify using the linear specification by viewing it as a linear approximation. However, the theoretical awkwardness of the LRUM specification needs to be kept in mind. A linear approximation of (19) does not imply that we approximate a function about which we do not know much. Rather, we approximate a function which we have good reason to believe is non-linear, as a linear function.

One additional comment on the LRUM model is warranted. In the terminology of the random utility framework, $\widetilde{V}_0(\mathbf{r}, y)$ and $\widetilde{V}_1(\mathbf{r}, y-p, b)$ are termed deterministic utility, while ε is viewed as a random part of utility. This is slightly confusing, since $\widetilde{V}_0(\mathbf{r}, y)$ and $\widetilde{V}_1(\mathbf{r}, y-p, b)-\varepsilon$ are really the expected utility from two different lotteries. However, since the value of the functions $\widetilde{V}_0(\mathbf{r}, y)$ and $\widetilde{V}_1(\mathbf{r}, y-p, b)$ will not be random, as long as preferences over such lotteries are well defined, while ε will appear random to an outside observer, this terminology is perhaps not so unnatural. It should be kept in mind, however, that in our interpretation, ε appears random only to the observer, while it is known to the consumer. (See Anderson, de Palma and Thisse [1] for a discussion of the RUM and LRUM models, and about how ε may be interpreted.)

As we will see, the LRUM assumptions are, however, very convenient in applied work. To see this, note first that a rational consumer will choose the lottery that gives her the highest level of (expected) utility. Thus, we may write the unconditional indirect utility of the individual as:

$$V(\mathbf{r}, y, p, b, \varepsilon) = \max\left\{\widetilde{V}_0(\mathbf{r}, y), \widetilde{V}_1(\mathbf{r}, y-p, b) - \varepsilon\right\} \tag{20}$$

The probability that a randomly selected individual with observable characteristics $\mathbf{r}$ and income y will choose $x_1 = 1$ (undergo mammography) given $\{b, p\}$ will then be defined by:

$$\begin{aligned}
\psi&\left[\widetilde{V}_1(\mathbf{r}, y-p, b), \widetilde{V}_0(\mathbf{r}, y)\right] = \\
&= \Pr\left(\widetilde{V}_1(\mathbf{r}, y-p, b) - \varepsilon \geq \widetilde{V}_0(\mathbf{r}, y)\right) = \\
&= \Pr\left(\widetilde{V}_1(\mathbf{r}, y-p, b) - \widetilde{V}_0(\mathbf{r}, y) \geq \varepsilon\right)
\end{aligned} \tag{21}$$

Assume that ε is iid and distributed according to some well behaved probability density function, $f(\varepsilon)$, and define the cumulative distribution function by $F(\varepsilon) = \int_{-\infty}^{\varepsilon} f(t)dt$. We can then write:

$$\begin{aligned}
\psi&\left[\widetilde{V}_1(\mathbf{r}, y-p, b), \widetilde{V}_0(\mathbf{r}, y)\right] = \\
&= \int_{-\infty}^{\widetilde{V}_1(\mathbf{r}, y-p, b)-\widetilde{V}_0(\mathbf{r}, y)} f(\varepsilon)\, d\varepsilon = \\
&= F\left(\widetilde{V}_1(\mathbf{r}, y-p, b) - \widetilde{V}_0(\mathbf{r}, y)\right)
\end{aligned} \tag{22}$$

Given some assumptions about the probability density function and the functional form of $\widetilde{V}_1(\mathbf{r}, y-p, b) - \widetilde{V}_0(\mathbf{r}, y, b)$, we can construct a likelihood function from this expression, and then use it for estimation of the parameters of the model.

One reason for the popularity of the LRUM model is that it gives rise to a very convenient welfare measure. Small and Rosen[23] have shown that, under three conditions, the following expression will measure the compensating variation from a change in prices and in the non-market good, b:

$$CV\left[\{p^0, b^0\}, \{p^1, b^1\}\right] =$$

$$= \frac{1}{\lambda} \int_{\widetilde{V}_1(\mathbf{r}, y-p^0, b^0)}^{\widetilde{V}_1(\mathbf{r}, y-p^1, b^1)} \psi\left[\widetilde{V}_1(\mathbf{r}, y-p, b), \widetilde{V}_0(\mathbf{r}, y)\right] d\widetilde{V}_1 \quad (23)$$

where λ is the marginal utility of income. This brings us to the first of Small's and Rosen's assumptions, which is that the marginal utility of income is approximately independent of p and b. Again, we must note that this assumption is almost certain to be violated in our model. This is perhaps best illustrated by returning to the continuous model in (3). The (expected) marginal utility of income in this model can be written as:

$$\frac{\partial U}{\partial y} = \pi(x_1, b) \frac{\partial u_1}{\partial x_n} + [1 - \pi(x_1, b)] \frac{\partial q}{\partial x_n} \quad (24)$$

since the marginal utility of income is equal to the marginal utility of the numéraire good. From this expression it is obvious that if we believe that b affects the probability of survival, then to assume that the marginal utility of income is independent of b is equivalent to assuming risk neutrality; in our case, assuming that the marginal utility given that the individual is alive will be the same as the marginal utility of income given that she is dead.

Once again, we may be able to live with this assumption, at least for small changes in the probability of death. However, it should be kept in mind that this assumption is most likely violated.

The second of Small's and Rosen's assumptions essentially requires that income effects are sufficiently small for the ordinary demand to be a good approximation of the compensated demand. To understand the implications of this assumption we first note that the demand for our health good, mammography, is either zero or one. The only way by which the Marshallian demand can differ from the Hicksian demand is therefore in having a different "critical price", above which the good will not be consumed. As Mäler [18] has demonstrated, ordinary and Hicksian welfare measures will coincide for a *fall* in the price. The reason for this is that no compensation is required as long as the price is above its critical level, so that the individual does not consume the good. When the price is below the critical level, the compensation cannot alter the amount of the good consumed. However, this result is not symmetric, since for a *rise* in the price, the critical price itself will be altered. If the good we are dealing with is relatively unimportant, this assumption should, however, be relatively innocuous. At least, there are no obvious reasons why it should be more troublesome in our setting than in the "ordinary" setting.

The third assumption that Small and Rosen make to arrive at the expression above consists of the by now familiar non-essentiality and weak complementarity assumptions. These assumptions have been addressed above.

Notice the similarity between this expression and expression (1). In fact, if we interpret the probability function, ψ, as the demand of a representative consumer, then the integral in (23) is the area to the left of this demand curve, between the values given by the function $\widetilde{V}_1$ evaluated at $\{p^0, b^0\}$ and $\{p^1, b^1\}$.[10] Note also that since we assume that the ordinary and compensated demand functions effectively coincide, the compensating variation and the equivalent variation will be equal.

2.4 When can market based methods be used?

Most of the restrictions discussed in the previous subsections apply to any welfare analysis of changes in the provision of health care. Obviously, if the compensating and the equivalent variations are not defined, then any attempt to estimate these welfare measures will be bootless. Compared to the restrictions we need to impose to make sure that our welfare measures are well defined, the additional assumptions, i.e. non-essentiality and weak complementarity, that are needed to use the specific class of methods that are the focus of this paper – market based valuation methods – may not appear so restrictive.

Consider first the weak complementarity assumption. For EV and CV to be well defined, utility v^1 and v^0, respectively, must be attainable at both the initial and final levels of b, with some finite income. Even if this is the case, weak complementarity may not hold if b represents some public good. Naturally, a public good may affect an individual's consumption of many different goods, and thus matter to the individual even if no x_1 is consumed. In many health economic applications, on the contrary, we would expect this restriction to be innocuous. Suppose that b corresponds to the quality of some medical treatment. Disregarding complications arising from altruistic considerations, it is unlikely that an individual cares about the quality of medical treatment that she *does not* undergo.

The non-essentiality condition may, on the other hand, be more troublesome when dealing with health economic problems than it is in the area of environmental economics. Unless we consider very broad categories of environmental goods, it is hard to envisage any market good, complementary to these, which is essential in the sense that the limits in (2) do not exist. When, on the other hand, we are dealing with the health of an individual, or even with the likelihood of his death, the assumption of non-essentiality

[10] See Anderson, de Palma and Thisse for a discussion of the representative consumer interpretation of random utility models.[1]

5

may not always be reasonable. Naturally, it is quite easy to construct examples when either CV or EV is undefined. Suppose, for example, that b^0 represents the case when insulin is available to a consumer with type I diabetes mellitus,[11] and that b^1 represents the case when it is not available. Then CV will obviously not be defined. No possible monetary compensation will restore the initial level of utility of the individual since he will die without insulin. EV may be defined but will hardly be meaningful. However, even in less extreme cases, when both EV and CV are defined, the non-essentiality condition may still be violated.

An example may serve to illustrate this distinction. Consider a new technology that improves the quality of hemodialysis for patients with kidney failure. It may for example reduce the time required and the inconvenience of undergoing hemodialysis. It is reasonable to assume that for most such patients both CV and EV will be defined, unless we are considering extreme changes in technology. A contingent valuation study may thus be used to provide an estimate of the value of this improvement. Even if the treatment is free, or provided at low cost to the patient, the patient would normally incur some private costs to undergo the treatment. For example, he might have to travel to a hospital. Thus the trip to the hospital would be complementary to the quality of the treatment. Further, the weak complementarity condition is likely to be satisfied, since it is reasonable to think that if the patient does not undergo hemodialysis, he will not care about its quality. However, the non-essentiality condition will most likely be violated. Provided an alternative treatment, such as a kidney transplant, is not available,[12] the patient will suffer severe health consequences if he does not undergo hemodialysis. In fact, he is most likely to die within a short period of time. Thus, the limits in expression (2) will not exist, and consequently the integrals in (1) will not measure the CV or EV.

However, if we carefully define what good we are dealing with, the non-essentiality condition need not be very restrictive. In the discussion in the last paragraph, hemodialysis was defined as one single good. Suppose that instead we use a simple household production function approach, where the final commodity which the individual cares about is the "treatment", z, which is produced using either the old technology A, or the new, improved technology B. Suppose, for simplicity, that the user fee for both alternatives is zero. However, some private expenses are incurred by the patient. In particular, the patient needs to travel to the hospital. Also, assume that the hospital where A is performed is closer to the patient than the hospital using technology B. Call travel to the A-hospital x_A and travel to the B-hospital x_B, and define b_A and b_B, respectively, to be the quality of treatment. Define the household production function $z = z(x_A, x_B, b_A, b_B)$. Defined this way, the quality of the new treatment, b_B, will

[11] Type I, or insulin-dependent diabetes mellitus (IDDM), used to be termed juvenile onset diabetes. In contradistinction to patients with type II, or non-insulin-dependent diabetes mellitus (NIDDM), patients with IDDM always require insulin injections.[6]

[12] In practice, the only alternative to hemodialysis is to undergo a kidney transplant. This option is not available to all patients due to limitations set by their own health condition and by the limited availability of suitable donors.

certainly be weakly complementary with travel to hospital B. Also, as long as CV is defined, x_B will be non-essential, since the individual will then use x_A if the price of x_B is sufficiently high.

The main limitation on the use of market based valuation methods is not likely to be that the assumptions underlying how the welfare measures are derived are violated. Rather, the rub will be that no suitable market good that is complementary to the good we want to value is available. In addition, assumptions that are used to translate the model into an applicable tool for welfare analysis, which are rather innocuous in a deterministic world, may become more troublesome when we deal with uncertainty. In particular, the assumptions behind the linear random utility model are likely to be violated in an expected utility model. It should be noted, however, that this is not unique to market based valuation methods. Rather, it will be a feature of any linear random utility model in an expected utility setting.

Another complication when we use market based valuation methods is that we will not in general be able to isolate the effect of a given change in some parameter on the survival probability. As is shown in Appendix B, we cannot derive a value of a statistical life, but only an upper bound for this value. In theory, we would not encounter this problem in a contingent valuation study, since we are then free to construct any hypothetical scenario we like, e.g. a situation where only the survival probability is changed and everything else remains the same. However, it is questionable if such a scenario would be viewed as realistic by the respondents. It is indeed hard to imagine any measure that would improve an individual's chance of survival, but otherwise leave his health unchanged.

3 Operationalization

Ignoring the caveats from the section above, we will now use expression (23) as a measure of welfare from changes in prices or from changes affecting the discrete good, and expression (22) as a definition of the probability that a randomly selected individual will choose to buy the discrete good (undergo mammography). To obtain an operational model, however, we need to specify the probability density function, $f(\varepsilon)$, and the functional forms of $\widetilde{V}_0(\mathbf{r};y)$ and $\widetilde{V}_1(\mathbf{r};y-p^0;b^0)$. In the next subsections we address these questions. In Subsection 3.1 we discuss the choice of distribution of the ε:s, and in Subsection 3.2 the parameterization of the conditional indirect utility functions is specified.

3.1 The logit model

The two most common distributions used in univariate discrete choice models are the standard logistic and the standard normal distributions, resulting in the so called logit and probit models, respectively. There is little theoretical guidance as to which distribution is more appropriate. (See e.g. Kennedy [15] for a discussion.) The logistic and normal distributions are very similar. The difference is that the logistic distribution has a larger probability mass in the tails. As Maddala [17] points out, we can expect logit and probit estimations to give more or less the same results, unless we have a large sample, and thus enough observations at the tails.

In practice, the choice of model is largely a matter of convenience. Since most econometrics packages have both models pre-programmed, both alternatives are often reported when standard models are estimated. In the present study, however, only the logit model will be estimated. That specification, in contradistinction to the probit specification, will make it possible to analytically solve the integral in expression (23), thus saving us some computational trouble.

Assuming that the ε:s follow the logistic distribution, the probability that a random selected individual will chose to undergo mammography, (22), may be written as:

$$\psi\left(\widetilde{V}_1, \widetilde{V}_0\right) = \frac{\exp\left(\widetilde{V}_1 - \widetilde{V}_0\right)}{1 + \exp\left(\widetilde{V}_1 - \widetilde{V}_0\right)} \tag{25}$$

and expression (23) may be written:

$$\begin{aligned} CV\left[\{p^0, b^0\}, \{p^1, b^1\}\right] &= EV\left[\{p^0, b^0\}, \{p^1, b^1\}\right] = \\ &= \frac{1}{\lambda}\left[\ln\left(1 + \exp\left(\widetilde{V}_1 - \widetilde{V}_0\right)\right)\right]_{\widetilde{V}_1(\mathbf{r}, y-p^0, b^0)}^{\widetilde{V}_1\left(\mathbf{r}, y-p^1, b^1\right)} \end{aligned} \tag{26}$$

In both these expressions, the arguments of $\widetilde{V}_0$ and $\widetilde{V}_1$ have been suppressed.

These expressions need to be interpreted carefully. Strictly speaking, it is incorrect to describe (25) as the probability that an individual with observable characteristics $\mathbf{r}$, income y, and facing price p and an exogenously given b, will choose to undergo a mammographic examination. Since we have assumed that ε is not a random term, only that it is unobservable, this is the probability that a randomly drawn individual, from a population of individuals with the same observables, will turn out to be an individual who underwent a mammography. As a consequence, (26) must not be interpreted as the compensating (equivalent) variation of an individual. Instead it is the average compensating (equivalent) variation for individuals with the same observables. Now, if we consider each observation in our data set as a representative of the population of individuals with the same observables, then (26) may be used to provide us with the aggregate compensating (equivalent) variation, if it is summed over all observations in the sample, and divided by the sampling ratio.

3.2 Parameterization of indirect utility

The next step we need to take in order to arrive at an estimatable model is to specify functional forms of the conditional indirect utility functions, $\widetilde{V}_0$ and $\widetilde{V}_1$. This decision will introduce another ad hoc restriction into the model. However, we may keep a large degree of flexibility by applying Taylor series expansions to the two functions. Conceivably, we could apply Taylor series expansions instead on each subfunction in equations (13) and (14). However, this would result in a very complex model.

First, second order Taylor series expansions of $\widetilde{V}_0$ and $\widetilde{V}_1$ are made around some point, say $\{\bar{\mathbf{r}}, \bar{y}\}$, ignoring higher order terms, to obtain an expression for the difference between these two functions. This provides us with a rather unwieldy expression. However, using the assumptions already made, and a few additional assumptions, we obtain[13]:

$$
\begin{aligned}
\widetilde{V}_1(\mathbf{r}; y-p) - \widetilde{V}_0(\mathbf{r}; y) \approx &\\
& \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p}) - \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y}) + \\
& + (\mathbf{r}-\bar{\mathbf{r}})' \left(\frac{\partial \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial \mathbf{r}} - \frac{\partial \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial \mathbf{r}} \right) - \\
& - (p-\bar{p}) \frac{\partial \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial y} + \\
& + \frac{1}{2} (p-\bar{p}) \left[(p-\bar{p}) - 2(y-\bar{y}) \right] \frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial y^2}
\end{aligned} \tag{27}
$$

We may thus write the final model specification as:

$$
\begin{aligned}
\widetilde{V}_1 - \widetilde{V}_0 = \beta_0 + (\mathbf{r}-\bar{\mathbf{r}})' \boldsymbol{\gamma} - \beta_1 (p-\bar{p}) + &\\
& + \beta_2 (p-\bar{p}) \left[(p-\bar{p}) - 2(y-\bar{y}) \right]
\end{aligned} \tag{28}
$$

where the β:s and the column vector $\boldsymbol{\gamma}$ are parameters to be estimated. Notice that β_1 corresponds to the marginal utility of income, and that β_2 corresponds to the second derivative of utility with regard to income, i.e. the change in the marginal utility of income. For the welfare measure defined in (23) to be valid, we have assumed that the discrete good is sufficiently unimportant for us to ignore income effects, and that the marginal utility of income is approximately independent of the price of the discrete good. That does not imply, however, that we cannot allow the marginal utility of income to differ between consumers with different levels of income.

Instead of using a Taylor series expansion, we could instead specify e.g. a linear or logarithmic model. (See e.g. Hanemann and Kanninen [9] for a discussion of these two formulations.) However, this would put limitations on the model which, if violated, may bring us to erroneous conclusions. As will be discussed below, we would expect income to influence the price, p, as it is one determinant of the opportunity cost of time. In addition, we may want to include income in the $\mathbf{r}$-vector, as a proxy for socioeconomic factors that may influence the likelihood of undergoing mammography. If income is ignored in $\mathbf{r}$, as a determinant of the opportunity cost of time, or in the specification of the conditional indirect utility, this may result in what amounts to an omitted variables bias.

[13]Details are provided in Appendix C.

The consequences may be particularly severe if we use the logarithmic specification of the model. Income and price would then enter as $\ln(1 - p/y)$[9]. Thus, if the coefficient on this variable is significant in a regression, it may reflect the fact that socioeconomic variables, correlated with income, influence the probability that a woman will undergo mammography, and not that this probability is related to the price. This would invalidate any welfare estimate. Since it seems exceedingly likely that income will in fact be correlated to socioeconomic factors that do influence the likelihood that a woman will undergo mammography, we must allow enough flexibility in the model to accommodate such an influence. Further, it is not clear what the logarithmic form implies in an expected utility model. Naturally, any specification of a functional form imposes restrictions on the model. However, a Taylor series expansion is simple, straightforward and reasonably flexible.

Given a specification of $\mathbf{r}$, it would now be possible to estimate the model defined by (25) and (28) under normal circumstances. However, a complication in travel cost studies is that we cannot directly observe the price each consumer faces. (See Randall[20] for a discussion.) Rather, the price, or perhaps more correctly, the cost, of undergoing mammographic screening will be influenced by several factors. Naturally, the user fee is part of the cost. In addition, there will be a cost associated with traveling to the location where the mammography is performed. This cost will consist of the out of pocket cost of travel, and the opportunity cost of time. Unfortunately, none of these are readily observable. In a detailed survey, we may perhaps obtain good data on the out of pocket cost of travel, as well as on travel time. However, the cost of time will still be illusive.

An alternative approach will be attempted in this study.[14] Instead of computing the price, based on stated out of pocket cost, travel time, and some assumption about the opportunity cost of time, we will assume that the cost of a trip to the place where mammography is performed can be written as a function of the user fee, the distance and some other factors that may influence the cost of undergoing mammographic screening. Specifically, we assume that the price can be expressed as

$$p = \alpha_0 FEE + \alpha(\mathbf{x}) \tag{29}$$

where FEE is the user fee, α_0 is a parameter, $\mathbf{x}$ is a column vector of factors influencing the cost to the consumer of undergoing mammography and $\alpha(\mathbf{x})$ is some function of these factors. Normalizing by setting $\alpha_0 = 1$ has the effect of denominating this estimated cost in monetary units. (In the present study, hundreds of SEK.)

We may again use a Taylor series expansion, to approximate $\alpha(\mathbf{x})$. A second order Taylor series expansion around $\bar{\mathbf{x}}$, ignoring higher order terms, yields:

$$\alpha(\mathbf{x}) \approx \alpha(\bar{\mathbf{x}}) + (\mathbf{x} - \bar{\mathbf{x}})' \frac{\partial \alpha(\bar{\mathbf{x}})}{\partial \mathbf{x}} + (\mathbf{x} - \bar{\mathbf{x}})' \frac{\partial^2 \alpha(\bar{\mathbf{x}})}{\partial \mathbf{x} \partial \mathbf{x}'} (\mathbf{x} - \bar{\mathbf{x}}) \tag{30}$$

[14]This approach is similar in spirit, though not in method, to the methodology employed by Englin and Shonkwiler [7].

Thus, we can write:

$$(p - \bar{p}) = \left(FEE - \overline{FEE}\right) + (\mathbf{x} - \bar{\mathbf{x}})' \frac{\partial \alpha (\bar{\mathbf{x}})}{\partial \mathbf{x}} + (\mathbf{x} - \bar{\mathbf{x}})' \frac{\partial^2 \alpha (\bar{\mathbf{x}})}{\partial \mathbf{x} \partial \mathbf{x}'} (\mathbf{x} - \bar{\mathbf{x}}) \tag{31}$$

In the present study, we will let $\mathbf{x}$ consist of three variables. We include the distance to the closest mammography unit and the respondent's income. The reason for including the distance should be obvious. Income is included to obtain a measure of the opportunity cost of time. In addition we include as an explanatory variable the proportion of the population living in multi-dwelling areas[15], in the parish in which the respondent lives. The reason why we want to include this variable is that it may serve as a proxy for the availability and quality of public transport.

In practice, we will have to constrain some of the derivatives in equation (31) to zero, to make estimation possible. In all estimations presented in this study, the diagonal elements of the second derivative will be constrained to zero. Thus, we assume that the second derivatives of the function with regard to all arguments are zero. However, we allow the cross derivatives to be non-zero. The final cost function may thus be written:

$$\begin{aligned}(p - \bar{p}) = \left(FEE - \overline{FEE}\right) &+ \alpha_1 \left(DIST - \overline{DIST}\right) + \\ &+ \alpha_2 (y - \bar{y}) + \alpha_3 \left(TOWNP - \overline{TOWNP}\right) + \\ &+ \alpha_4 \left(DIST - \overline{DIST}\right)\left(TOWNP - \overline{TOWNP}\right) + \\ &+ \alpha_5 \left(TOWNP - \overline{TOWNP}\right)(y - \bar{y}) + \\ &+ \alpha_6 \left(DIST - \overline{DIST}\right)(y - \bar{y})\end{aligned} \tag{32}$$

where $DIST$ is the distance to the closest mammography unit, $TOWNP$ is the proportion of the parish population living in multi-dwelling areas, and the α:s are parameters to be estimated. With this specification, we may still expect multicollinearity between the variables. However, since we are not primarily interested in the coefficients of the cost equation per se, but rather in the coefficients of the expression for the indirect utility, (28), this is not a serious problem, provided that we are able to obtain convergence of the likelihood function.

Using expression (32) together with expression (28), and plugging them into a log-likelihood function constructed from the probabilities (25), we have a model that can in principle be estimated.

4 Data and estimation

In the first subsection below, we describe the data. In Subsection 4.2, the estimation of the model and the results are discussed. The following section, 4.3, presents estimates of

[15] A "multi-dwelling area" corresponds to what Statistics Sweden terms a "locality". This term is defined as a "group of buildings normally not more than 200 metres apart from each other, and has to fulfill a minimum criterion of having at least 200 inhabitants." (Statistics Sweden [24], p. 4.)

the welfare derived from mammographic screening. In the last subsection, some possible policy implications are discussed, and the results are compared to other studies.

4.1 The data

The data used in this study were obtained from a telephone survey conducted by Skandinavisk Opinion AB, a survey firm. Data were collected during the periods 15 May through 4 June, and 6 through 30 June 1996. In the random sampling, specific individuals were selected. Women who could not be interviewed over the telephone were sent mail questionnaires. The non-response rate was 17 percent during the first period, and 19 percent during the second period. Women were asked if they had undergone mammography during the past two years and about their age, income and employment status. Information is also available on the parish where the respondent lives, the parish being the smallest geographical region in Sweden.

A survey was sent out to the physician responsible for the screening programs in each medical district[16] in Sweden. Answers were received from all of them. At the time of the survey, all 29 medical districts except Gotland had screening programs in operation. The screening programs differ somewhat between the different medical districts. The lower limit of the age groups included in the screening program is between 40 and 50 years, and the upper limit is either 69 or 74 years. The interval between invitations to mammography also differs, from 11 months to 24 months. The patient fee varies from zero to 200 SEK (≈25USD). In 15 of the medical districts, mobile mammography units are used to make it simpler for women living far from a hospital to undergo mammography.

A distance table from the Swedish road authority (Vägverket [25]), and an ordinary road map, were used to compute the distance from the largest multi-dwelling area in each parish to the closest mammography unit. This distance varies from 0.5 km to a little over 100 km. Demographic data on the parishes was obtained from Statistics Sweden.

Of the total sample of 1 045 observations, 110 observations had to be discarded due to missing data. In addition, data for interviewees living in Stockholm county were discarded. The reason for excluding women from Stockholm is that it is unlikely that distance is a good measure of travel time and cost in a large city like Stockholm. Rather, easy access to the subway or other public means of transportation is likely to be more important. This brings down the final sample size to 761 observations.

Of these women, 344 answered that they had undergone mammography during the past two years. Of the whole sample, 373 women were covered by a screening program. In this subsample, 292 women had undergone mammography.

[16]With a few exeptions, the medical districts coincide with counties.

4.2 Estimation and results

The Taylor expansions are made around an arbitrary point. In all estimations, this point is the mean value for all variables, except for a dummy variable. A variable with an upper bar should thus be interpreted as a mean. Estimation is carried out on a transformed data set consisting of deviations from the mean.

Estimation is complicated by the fact that the individual's income enters both in the indirect utility functions and in the price equation, through the vector $\mathbf{x}$. A further complication is that we can expect income to be highly correlated with socioeconomic factors influencing the likelihood that a woman will undergo mammography. Thus, we would like to include income also in the vector $\mathbf{r}$, as a proxy for such socioeconomic factors. In principle, the model is still possible to estimate, with the restrictions already made, since $\mathbf{r}$, $\mathbf{x}$ and $\widetilde{V}_1 - \widetilde{V}_0$ involve different functions of income. Age is also likely to influence the utility from undergoing mammography. In addition, we include a dummy variable equal to one if the individual is covered by a screening program, and zero otherwise. Thus, we define the vector $\mathbf{r}$ by: $\mathbf{r} = \{PCOV, AGE, y\}$, where $PCOV$ is the dummy variable.

The results from the estimation of this model can be viewed in Table 1. (Model A.) The results are not particularly encouraging. In the cost equation, only $TOWNP$ is significant, and then only at the 10 percent level of significance. Not surprisingly, $PCOV$ and AGE are significantly different from zero (at the 1 percent level of significance). Income also appears to have a significant effect (at the 10 percent level) on the probability that the individual undergoes mammography. A disturbing result in model A is that β_1 and β_2 have the wrong signs. Remember from equation (28) that β_1 corresponds to the marginal utility of income, evaluated at $\bar{y}$. Thus, it ought to be positive. Since β_2 corresponds to the change in the marginal utility of income, when income changes, we would expect it to be negative, or at least non-positive. The result from the estimation of model A is thus unreasonable. Even more troubling is the fact that this makes it impossible to obtain a measure of the compensating variation, since any such exercise requires that we have an estimate of the marginal utility of income. In fact, marginal utility of income calculated on the basis of the estimates from model A turns out to be negative for a large number of individuals. Also, the absolute value of the calculated compensating variation is huge for some individuals, since the calculated marginal utility of income is close to zero.

/ Table 1 about here. /

Imposing the restrictions $\alpha_2 = 0$ and $\beta_2 = 0$, i.e. leaving out the linear income term from the price equation, and the term corresponding to the second derivative of the indirect utility with regards to income, provides us with more reasonable results. (Model B.) All three γ-parameters, i.e. parameters corresponding to the vector $\mathbf{r}$, are significant at least on a 5 percent level of significance, which is basically the same result as in model A. In model B, however, none of the parameters in the cost function are significant, nor is the parameter on price, β_1. However, this coefficient, which we remember from (27) also corresponds to the marginal utility of income, has the expected sign, i.e. it is positive.

Models A and B are estimated on the full data set. What we are really interested in, however, is how the women who are covered by a screening program act. Thus, a model is estimated on this subsample of the data. (Model C.) Since this data set is rather small, all α-parameters except α_1 and α_6, i.e. the coefficients corresponding to distance and the product of distance and income, respectively, are constrained to zero, to simplify estimation. Obviously, the dummy-variable indicating if the individual is covered by a screening program is left out, as is the age variable, since only a rather small age group is covered by screening programs. In this regression, no coefficients are significant, except the constant. However, the price coefficient has the right sign, i.e. the estimated marginal utility of income is positive.

A number of alternative specifications were estimated. In particular, alternative cost functions were used. As can be expected for models with so many insignificant coefficients, the estimates vary to a certain degree. However, the qualitative results are unaltered in most specifications.

4.3 Benefit estimates

Ignoring the inadequacies of the model and the data, the results from model B and model C can be used to obtain estimates of the welfare derived from mammographic examinations. Since we have assumed that the Hicksian and Marshallian demand curves are essentially identical, estimates of CV and EV will be identical. If we choose to interpret our welfare measure as a compensated variation, we are in effect making the thought experiment that mammography was made unavailable. We are then attempting to measure what amount of money would be needed to compensate the women for the consequent loss of utility. It is perhaps more natural to interpret the welfare estimate as an equivalent variation. Our welfare measure can then be seen as the highest price a woman would be willing to pay to undergo a mammographic screening. Using (26) and the fact that β_1 is our estimate of the marginal utility of income, our formula for the welfare from mammography is:

$$\Delta W = CV = EV = -\frac{\ln\left(1 + \exp(\widetilde{V}_1 - \widetilde{V}_0)\right)}{\beta_1} \tag{33}$$

where we use the estimated models to calculate $\widetilde{V}_1 - \widetilde{V}_0$. Summary statistics for the estimates are displayed in Table 2.

/ Table 2 about here./

In model B, (the absolute value of) the mean estimated compensated/equivalent variation per woman is 331 SEK. For the subsample of women who are covered by screening programs, the mean is 615 SEK. A more interesting measure than CV or

EV per woman is perhaps the estimated welfare per mammographic examination undertaken. This figure is obtained by dividing ΔW for each woman by the estimated probability of undergoing mammography.[17] The corresponding means of the estimates are 597 SEK and 780 SEK.

The estimated values from model C are all considerably higher. The mean estimated CV or EV per woman is 1 710 SEK, and the mean per mammography is 2 180 SEK. This difference is entirely due to the lower estimate of the marginal utility of income in model C. In fact, using the estimate of marginal utility of income from model B, and the nominator of expression (33), from model C to compute ΔW, gives values that are almost identical to the values obtained from model B. The estimate per woman is 605 SEK. Per mammography it is 770 SEK. This illustrates how sensitive the analysis is to the estimate of the marginal utility of income.

4.4 Policy implications

Naturally, it is of interest to know how these figures compare to the cost of screening programs. Calculating the cost of a screening program is not entirely straight forward. It is possible, for example, that early detection of tumours affects the cost of medical treatment, providing an additional benefit from the screening program, and thus affecting the net cost. A study of the screening program in Kopparbergs län (Kopparberg County) provides us with some indication of the cost. The extra cost, compared to a control group, of inviting 39 000 women to mammography was estimated at 34 million SEK. Around 90 percent of these women underwent mammography. The cost per mammography would thus be around 970 SEK.[8] It is not reasonable to draw any conclusions about whether screening programs are justified from a social cost-benefit point of view, based on the present study. However, it is somewhat comforting to note that the benefit estimates are of the same magnitude as the costs.

As pointed out above, in Section 2.2, it is not reasonable to assume that the value a woman derives from undergoing mammography is entirely attributable to the reduced risk of death. However, since we have shown that dividing ΔW by the change in probability of death provides us with an upper bound on the value of a statistical life,[18] we can use it as a means of assessing whether our benefit estimates are reasonable. Wright and Mueller [30] summarize six studies of the health benefits from mammography. In these studies, the estimated ratio of screenings per lives saved per year ranges from 7 086 to 63 264, excluding one study which found no risk reduction at all. Using these figures and our estimates of ΔW would provide us with (an upper bound on) the mean values of the

[17] Remember that with the representative consumer interpretation of random utility models, we can view the probabilities as the proportion of women with given values of the observable variables who will undergo mammography.

[18] It should be noted that the claim that we obtain an upper bound on the value of a statistical life rests on the assumption that the model is correctly specified. In particular, if there are some potential disadvantages of undergoing mammography, such as a risk for false positives, false negatives, or some discomfort from undergoing the examination, the result may not hold.

implicit value of a statistical life ranging from 5.5 million SEK to 49 million SEK, and from 15 million SEK to 140 million SEK using, respectively, the estimates from model B and model C.[19] Estimates of the value of a statistical life vary greatly. Viscusi [28] summarizes the results from 44 different studies. In these studies estimates of the value of a statistical life range from USD 70 000 (SEK 56 000) to USD 16.2 million (SEK 130 million). The mean value is USD 5.09 million (SEK 41 million). However, the median value is considerably lower, USD 3.7 million (SEK 30 million). Thus, except for the highest estimate, the values obtained in this study seem to be reasonably in line with other estimates of the value of a statistical life.[20]

5 Concluding remarks

There is really no reason to believe that the theoretical assumptions behind market based valuation methods should put more severe limitations of the applicability of market based valuation methods when these are applied in the field of health economics, than when they are used in a more traditional setting to value public goods. The problems with using such methods are rather likely to be the same as in the field of environmental economics. Firstly, there may not exist any market good which is meaningfully linked with the non-market good we want to value. Secondly, we will encounter numerous practical problems in estimating our model. One of the most important of these problems is to define the "true cost" of undertaking a treatment, or for that matter of traveling to some recreation site. However, an additional problem when we leave the "ordinary", deterministic setting, is that the linear random utility model implies some awkward assumptions when set within an expected utility framework.

The empirical study presented in this paper serves well to illustrate some of these problems. The results may appear quite discouraging. However, this may at least partly be a result of the data employed. A richer data set, including information on e.g. known risk factors for breast cancer and more detailed socio-economic data, may have a greater explanatory power. A key source of the estimation difficulties is probably the fact that income is correlated with many such variables. Controlling for such factors may improve the results.

In addition, we may obtain greater variability in our data by collecting them in a region where the difference in travel distances is likely to be large between individuals, since a prime difficulty in our attempts to estimate welfare measures is that we do not obtain statistically significant estimates of the marginal utility of income. This may indicate either that our sample is just too small, or that the variation in the cost of

[19] The mean of the estimated compensating variation per mammography for women covered by a screening program is 780 SEK, using model B. 780*7 086≈5.5 millions. 780*63 265≈49 millions. The mean of the compensating variation per mammography from model C is 2 180 SEK. 2 180*7 089≈15 millions. 2 180*63 264≈140 millions.

[20] Since Clarke [5] values changes in access cost, rather than mammography per se, comparisons with his results are not meaningful.

undergoing mammography is not large enough to have any detectable influence on the likelihood that a woman will choose to undergo an examination.

In spite of all these empirical problems, the benefit estimates appear to be of a reasonable magnitude, even though they must definitely be taken with a good pinch of salt. Obviously, a number of questions need to be answered before we would be comfortable in using an analysis such as the one presented in this paper as the basis for policy decisions. However, there is reason to believe that the development of market based valuation methods in the area of health economics will be a fruitful task.

References

[1] Anderson, Simon P., André de Palma and Jacques-François Thisse (1992) *Discrete choice theory of product differentiation.* Cambridge, Mass. and London: MIT Press.

[2] Berck, Peter and Knut Sydsæter (1993) *Economists' Mathematical Manual.* 2nd ed. Heidelberg: Springer Verlag.

[3] Bockstael, Nancy and Kenneth E. McConnell (1993) "Public goods as characteristics of non-market commodities." *The Economic Journal.* 103:1244-1257.

[4] Bockstael, Nancy and Kenneth E. McConnell (1991) "Recreation." in John B. Braden and Charles D. Kolstad, eds. *Measuring the demand for environmental quality.* Amsterdam: North-Holland.

[5] Clarke, Philip M. (1996) "Cost-benefit analysis and mammographic screening: a travel cost approach." *Journal of Health Economics.* 17:767-787.

[6] "diabetes mellitus" *Britannica Online.* <http://www.eb.com:180/cgi-bin/g?DocF=micro/711/75.html> [Accessed 01 December 1998].

[7] Englin, Jeffrey and J.S. Shonkwiler (1995) "Modeling recreational demand in the presence of unobservable travel cost: towards a travel price model." *Journal of Environmental Economics and Management.* 29:368-377.

[8] *Hälsoundersökning med mammografi.* (1998) SoS-rapport 1998:17. Spånga: Socialstyrelsen.

[9] Hanemann, W. Michael and Barbara Kanninen (1996) "The statistical analysis of discrete-response CV data." Department of Agricultural and Resource Economics, University of California at Berkley, Working paper no 798.

[10] Hotelling, Harold (1947) Letter to the US National Park Service.

[11] Johansson, Per-Olov (1997) "On the use of market prices to evaluate medical treatments." *Journal of Health Economics.* 16:609-615.

[12] Johansson, Per-Olov (1995) *Evaluating Health Risks.* Cambridge: Cambridge University Press.

[13] Journal of Economic Perspective (1994) "Symposia on the contingent valuation method." *Journal of Economic Perspectives*, 8.

[14] Kahneman, D. and A. Tversky (1979) "Prospect theory: an analysis of decisions under risk." *Econometrica.* 47:263-91.

[15] Kennedy, Peter (1992) *A guide to econometrics.* Oxford, UK and Cambridge, Mass.: Blackwell.

[16] Lidbrink, Elisabeth (1995) "Mammographic screening for breast cancer. Aspects on benefits and risks." Stockholm: Department of Clinical Oncology, Södersjukhuset and Radiumhemmet, Karolinska Institute.

[17] Maddala, G.S. (1983) *Limited-dependant and qualitative variables in econometrics.* Cambridge, UK: Cambridge University Press.

[18] Mäler, Karl-Göran (1974) *Environmental economics: a theoretical inquiry.* Baltimore: John Hopkins University Press.

[19] Mas-Collei, Andreu, Michael D. Whinston and Jerry R. Green (1995) *Microeconomic Theory.* New York and Oxford: Oxford University Press.

[20] Randall, Alan (1994) "A difficulty with the travel cost method." *Land Economics.* 70:88-96.

[21] Rosen, S. (1986) "The theory of equalizing differences." in O. Ashenfelter and R. Layard (eds.) *Handbook of labor economics.* Amsterdam: Elsvier Science Publishers.

[22] Shogren, Jason F. and Thomas D. Crocker (1991) "Risk, self-protection, and ex-ante economic value." *Journal of environmental economics and management.* 20:1-15.

[23] Small, Kenneth A. and Harvey S. Rosen (1981) "Applied welfare economics with discrete choice models." *Econometrica.* 49:1:105-130.

[24] *Tätorter (localities) 1995* (1996) Statistiska meddelanden. Be 16 SM 9601. Stockholm: Statistics Sweden.

[25] Vägverket (1996) *Vägavstånd i Sverige.* Borlänge: Vägverket.

[26] Varian, Hal R. (1992) *Microeconomic analysis.* 3rd ed. New York and London: W.W. Norton & Company.

[27] Viscusi, W.K. (1993) "The value of risks to life and health." *Journal of Economic Literature*, 31:1912-46.

[28] Viscusi, W.K. (1992) *Fatal tradeoffs. Public and private responsibilities for risk.* New York and Oxford: Oxford University Press.

[29] Weibull, Jörgen (1992) "Lecture notes in Mathematics for Economists." Stockholm: Department of Economics, University of Stockholm.

[30] Wright, Charles J. and C. Barber Mueller (1995) "Screening mammography and public health policy: the need for perspective." *Lancet.* July 1, 346:29-32.

A An expected utility model

A.1 The model

Define a two state model, and assume that in state one, the individual is alive, and in state two she is dead. Call the probability that state one will be realized $\pi(x_1, b)$, where x_1 is a "health good" purchased at the market at price $p > 0$, and b is a non-market good which is outside the individual's decision set. The consumer's health, given that she is alive, can be described by a cardinal health function, $h(x_1, b)$. Thus, health as well as the probability of being alive depends on x_1 and b. Assume that all consumption other than x_1 can be aggregated into a single "numéraire good". The budget constraint can thus be written $x_n + px_1 \leq y$. Define cardinal utility conditional upon being alive $u_1(x_n, h(x_1, b))$, and define a likewise cardinal "bequest function" $q(x_n)$.

If the consumer's behavior satisfies the von Neuman-Morgenstern assumptions, then the utility maximization problem can be written:

$$\max_{x_1, x_n} \{U(x_1, x_n, b) = \pi(x_1, b) u_1 [x_n, h(x_1, b)] + [1 - \pi(x_1, b)] q(x_n)\}$$
$$s.t. \begin{cases} x_1 \geq 0 \\ x_n \geq 0 \\ x_n + px_1 \leq y \end{cases} \tag{34}$$

We will assume that the functions $u_1(x_n, h(x_1))$, $q(x_n)$, $h(x_1, b)$ and $\pi(x_1, b)$ are "well behaved" in the sense that they are twice continuously differentiable, and that they satisfy the following conditions:

$$\frac{\partial u_1}{\partial x_n} > 0 \tag{35a}$$
$$\frac{\partial u_1}{\partial h} > 0 \tag{35b}$$
$$\frac{\partial h}{\partial x_1} \geq 0 \tag{35c}$$
$$\frac{\partial \pi}{\partial x_1} \geq 0 \tag{35d}$$
$$\frac{\partial q}{\partial x_n} \geq 0 \tag{35e}$$
$$u_1 > q, \frac{\partial u_1}{\partial x_n} > \frac{\partial q}{\partial x_n} \forall x_n, x_1, b, \pi > 0 \tag{35f}$$
$$\frac{\partial^2 u_1}{\partial h^2} \leq 0, \frac{\partial^2 u_1}{\partial x_n^2} \leq 0, \frac{\partial^2 q}{\partial x_n^2} \leq 0 \tag{35g}$$

Assumptions (35a) and (35b) are assumptions of local non-satiation. The assumption (35c) implies that the marginal return from investing in health is always non-negative, and assumption (35d) congruently ensures that the probability of ending up in state

$r = 1$ is always non-decreasing in x_1. Assumption (35e) implies that the marginal utility of leaving a bequest is non-negative. By assumption (35f) it is always better to be alive than to be dead. Also, as long as there is a positive probability that we will be alive, the marginal utility of the numéraire good is higher when we are alive than when we are dead. Assumptions (35g), finally, imply that the marginal utility of health as well as the marginal utility of the numéraire good are non-increasing. These assumptions all appear quite uncontroversial. In all of the following, we will also assume that $\pi > 0$. We thus assume that there is a positive probability that the individual will be alive.

A.2 The indirect utility function

Assuming $y > 0$ and $p > 0$, the budget set will be a closed and non-empty set. Since we have assumed that $u_1[y - px, h(x,b)]$, $q(y - px)$, $h(x,b)$ and $\pi(x,b)$ are continuous, twice differentiable functions, $U(x,b)$ will also be continuous and twice differentiable. Since $U(x,b)$ is continuous, and the budget set is closed and non-empty, we can be sure, by Weierstrass' Maximum Theorem, that the solution set of (34) is non-empty for all $\{p, b, y\}$.

We can thus define the indirect utility function:

$$v(p,b,y) = \max_{x_1, x_n} \{U(x_1, x_n, b) = \pi(x_1, b)\, u_1[x_n, h(x_1, b)] + [1 - \pi(x_1, b)]\, q(x_n)\}$$
$$s.t. \begin{cases} x_1 \geq 0 \\ x_n \geq 0 \\ x_n + px_1 \leq y \end{cases} \tag{36}$$

A.3 Is the indirect utility function continuous?

By assumptions (35a) and (35e), U is strictly increasing in x_n. Thus, we can reformulate the utility maximization problem and the value function (i.e. the indirect utility function):

$$v(p,b,y) = \max_{x_1} \left\{ \begin{array}{c} U(x_1, y - px_1, b) = \pi(x_1, b)\, u_1[y - px_1, h(x_1, b)] + \\ + [1 - \pi(x_1, b)]\, q(y - px_1) \end{array} \right\}$$
$$s.t. \begin{cases} x_1 \geq 0 \\ x_n = y - px_1 \geq 0 \end{cases} \tag{37}$$

Since, by assumption, u_1 and q are continuous in x_n, we immediately see that U is continuous in prices and income. We have also assumed that π is continuous in b so that U will also be continuous in b. Thus, the indirect utility function, V, will be continuous in p, y and b.

A.4 Is demand continuous?

Since the budget set is convex, non-empty and compact, we know that if U is strictly quasi-concave in x_1, the solution to the maximization problem (34) is unique for all $p > 0$, $y > 0$ and b.[29] Also, if the solution to the maximization problem is unique, then the solution will be continuous in $\{p, y, b\}$, since U is continuous in $\{p, y, b\}$.[2] In other words, if U is strictly quasi-concave, the demand for x_1 will be continuous in $\{p, y, b\}$. In this section, we explore if U is strictly quasi-concave under reasonable assumptions.

U will be strictly quasi-concave in x_1 iff[2]:

$$\frac{\partial U(x_1, y - px_1, b)}{\partial x_1} = 0 \Rightarrow \frac{\partial^2 U(x_1, y - px_1, b)}{\partial x_1^2} < 0 \tag{38}$$

From (37) we obtain the first derivative:

$$\frac{\partial U(x_1, y - px_1, b)}{\partial x_1} = \\ = \frac{\partial \pi}{\partial x_1}(u_1 - q) + \pi \left[\frac{\partial u_1}{\partial h}\frac{\partial h}{\partial x_1} - \frac{\partial u_1}{\partial x_n}p\right] - (1 - \pi)\frac{\partial q}{\partial x_n}p \tag{39}$$

where the arguments of the functions π, u_1, q and h have been suppressed. The second derivative will be:

$$\begin{aligned} \frac{\partial^2 U(x_1, y - px_1, b)}{\partial x_1^2} = & \frac{\partial^2 \pi}{\partial x_1^2}(u_1 - q) + \\ & + 2\frac{\partial \pi}{\partial x_1}\left[\frac{\partial u_1}{\partial h}\frac{\partial h}{\partial x_1} - \frac{\partial u_1}{\partial x_n}p\right] + 2\frac{\partial \pi}{\partial x_1}\frac{\partial q}{\partial x_n}p \\ & + \pi \left[\begin{array}{c} \frac{\partial^2 u_1}{\partial h^2}\left(\frac{\partial h}{\partial x_1}\right)^2 - 2p\frac{\partial^2 u_1}{\partial x_n \partial h}\frac{\partial h}{\partial x_1} + \\ + \frac{\partial u_1}{\partial h}\frac{\partial^2 h}{\partial x_1^2} + p^2 \frac{\partial^2 u_1}{\partial x_n^2} \end{array}\right] \\ & + (1 - \pi)\frac{\partial^2 q}{\partial x_n^2}p^2 \end{aligned} \tag{40}$$

Assuming $\pi > 0$, $\partial U(x_1, y - px_1, b)/\partial x_1 = 0$ implies:

$$\frac{\partial u_1}{\partial h}\frac{\partial h}{\partial x_1} - \frac{\partial u_1}{\partial x_n}p = \frac{(1 - \pi)\frac{\partial q}{\partial x_n}p}{\pi} - \frac{\frac{\partial \pi}{\partial x_1}(u_1 - q)}{\pi} \tag{41}$$

Inserting this expression into (40), we obtain the following condition for U to be strictly quasi-concave:

$$\left.\frac{\partial^2 U(x_1, y - px_1, b)}{\partial x_1^2}\right|_{\frac{\partial U}{\partial x_1}=0} =$$

$$= \frac{\partial^2 \pi}{\partial x_1^2}(u_1 - q) - 2\frac{(\partial \pi/\partial x_1)^2}{\pi}(u_1 - q) + 2\frac{1}{\pi}\frac{\partial \pi}{\partial x_1}\frac{\partial q}{\partial x_n}p$$
$$+ \pi \left[\begin{array}{c} \frac{\partial^2 u_1}{\partial h^2}\left(\frac{\partial h}{\partial x_1}\right)^2 - 2p\frac{\partial^2 u_1}{\partial x_n \partial h}\frac{\partial h}{\partial x_1} + \\ + \frac{\partial u_1}{\partial h}\frac{\partial^2 h}{\partial x_1^2} + p^2\frac{\partial^2 u_1}{\partial x_n^2} \end{array} \right]$$
$$+ (1 - \pi)\frac{\partial^2 q}{\partial x_n^2}p^2 < 0 \quad (42)$$

Under assumptions (35a) through (35g), we can be sure that the second and the last terms on the first line of expression (42) are non-positive, as well as the first and last terms within square brackets. Rearranging this expression so that we have all terms that are always non-positive on the left hand side, and all other terms on the right hand side, yields:

$$- 2\frac{(\partial \pi/\partial x_1)^2}{\pi}(u_1 - q)$$
$$+ \pi \left[\frac{\partial^2 u_1}{\partial h^2}\left(\frac{\partial h}{\partial x_1}\right)^2 + \frac{\partial^2 u_1}{\partial x_n^2}p^2\right] +$$
$$+ (1 - \pi)\frac{\partial^2 q}{\partial x_n^2}p^2 <$$
$$< 2\pi\frac{\partial^2 u_1}{\partial x_n \partial h}\frac{\partial h}{\partial x_1}p - 2\frac{1}{\pi}\frac{\partial \pi}{\partial x_1}\frac{\partial q}{\partial x_n}p$$
$$- \pi\frac{\partial u_1}{\partial h}\frac{\partial^2 h}{\partial x_1^2} - \frac{\partial^2 \pi}{\partial x_1^2}(u_1 - q) \quad (43)$$

From the restrictions, (35a) through (35g), the maximum value of the left-hand side of the inequality is zero. A sufficient condition for uniqueness of the maximum, and thus for continuity of the demand, is thus that:

$$2\pi\frac{\partial^2 u_1}{\partial x_n \partial h}\frac{\partial h}{\partial x_1}p - 2\frac{1}{\pi}\frac{\partial \pi}{\partial x_1}\frac{\partial q}{\partial x_n}p - \pi\frac{\partial u_1}{\partial h}\frac{\partial^2 h}{\partial x_1^2} - \frac{\partial^2 \pi}{\partial x_1^2}(u_1 - q) > 0 \quad (44)$$

Suppose that the bequest function is unimportant enough to be ignored. Then we can be sure that the condition is fulfilled if $\partial^2 u_1/\partial x_n \partial h$ is non-negative, and $\partial^2 h/\partial x_1^2$ and $\partial^2 \pi/\partial x_1^2$ are non-positive. Put differently, this requires that the numéraire good and health are not substitutes in the utility function, and that the marginal increase in health and in the probability of survival are not increasing in the health good.

Thus, if: 1) The underlying functions - utility when the individual is alive, the bequest function, the health function and the probability function - are well-behaved in the sense that they are continuous, twice continuously differentiable, and satisfy the assumptions (35a) through (35g), 2) health and the numéraire good are complements in the utility function, 3) the marginal increase in health and in the probability of survival are non-increasing in the health good and 4) the marginal utility of income in the bequest function is negligible, then we can be sure that demand for x_1 is continuous. If, on the other hand, health and the numéraire good are substitutes for some values of $\{p, y, b\}$, or if for some values of x_1 the marginal increase in health or in the probability of survival are increasing, then demand may be discontinuous, even if the underlying functions are well behaved.

B The value of a statistical life

To set priorities between different costly measures to reduce health risks, one proposed yardstick is the so called "value of a statistical life". Used as such, this value is the amount of society's resources a decision maker should be willing to commit to a project that reduces by one the number of people likely to die from a given health hazard. Thus if the cost of a project, divided by the number of lives it is likely to save, is less than the value of a statistical life, then by this decision criterion, it should be undertaken. If not, it should not be undertaken. To use the value of a statistical life in this fashion, it is of course of crucial importance that we are able to obtain reasonable estimates of this value. In other words, we need to know what value people put on reductions in the risk of death. In practice, estimates of the value of a statistical life are obtained by dividing by the accomplished risk reduction the cost people voluntarily incur, in various circumstances, in order to reduce their risk of death.

It is thus noteworthy that we will usually not be able to obtain an estimate of the value of a statistical life from a valuation study based on market data, when the health good affects both the survival probability and the individual's health. Suppose that the probability of survival is exogenously changed by an infinitesimally small amount, say $d\pi$. The compensating variation[21] for this change is equal to the negative of the change in income necessary to bring the individual back to his initial level of utility. Call this income change dy_π. One measure of the value of a statistical life would then be $-dy_\pi/d\pi$. (See Johansson [12] for a discussion.) Assume that the conditions for the utility function (3) to be strictly quasi-concave are fulfilled, so that demand is continuous. Then, by differentiating this utility function, substituting for x_n using the budget restriction, and

[21]The alterations necessary for the reasoning to apply to the equivalent variation are obvious.

keeping b and p fixed, we will have (by the envelope theorem):

$$dU = [u_1(y - px_1^*, h(x_1^*, b)) - q(y - px_1^*)]\, d\pi + \\ + \left(\begin{array}{c} \pi(x_1^*, b)\dfrac{\partial u_1(y - px_1^*, h(x_1^*, b))}{\partial x_n} + \\ + (1 - \pi(x_1^*, b))\, \dfrac{\partial q(y - px_1^*)}{\partial x_n} \end{array} \right) dy_\pi \quad (45)$$

where x_1^* is the optimal demand of x_1, given p, b and y. Setting $dU = 0$, we may then define the value of a statistical life, $VOSL$, by:

$$VOSL = -\frac{dy_\pi}{d\pi} = \frac{u_1 - q}{\pi \dfrac{\partial u_1}{\partial x_n} + (1 - \pi)\, \dfrac{\partial q}{\partial x_n}} \quad (46)$$

where the arguments of u_1, q and π have been subsumed. Notice that the denominator of this expression is the expected marginal utility of income. The value of a statistical life will thus be the difference in utility between the two states - alive and dead - transformed into a money measure by dividing by the expected marginal utility of income.

Now, consider a change in b. We repeat the exercise, this time keeping p and v fixed. Call the change in income needed to preserve initial utility, dy_b. Thus:

$$dU = \left[\frac{\partial \pi}{\partial b}(u_1 - q) + \pi \frac{\partial u}{\partial h}\frac{\partial h}{\partial b} \right] db + \left(\pi \frac{\partial u_1}{\partial x_n} + (1 - \pi)\, \frac{\partial q}{\partial x_n} \right) dy_b \quad (47)$$

Obviously, if we set $dU = 0$, dy_b will be equal to the negative of the compensating variation for this change in b.

The change in the survival probability, induced by a change in b, is $d\pi = (\partial \pi / \partial b) \cdot db$. Rearranging, we obtain:

$$-\frac{dy_b}{(\partial \pi / \partial b) \cdot db} = \frac{CV}{d\pi} = \frac{(u_1 - q) + \dfrac{\pi}{db}\dfrac{\partial u}{\partial h}\dfrac{\partial h / \partial b}{\partial \pi / \partial b}}{\pi \dfrac{\partial u_1}{\partial x_n} + (1 - \pi)\, \dfrac{\partial q}{\partial x_n}} \neq VOSL \quad (48)$$

Thus, if we estimate the compensated variation for a change in b, and then divide this estimate by the induced change in the probability of survival, we will *not* in general obtain an estimate of the value of a statistical life. Obviously, an exception occurs if $\partial h / \partial b = 0$. From expression (48), it is clear that the bias will be larger, the larger is the ratio between the marginal effects of b on health and on the survival probability, respectively. Since, by assumption, the second term in the numerator will always have the same sign as db, and because db and $d\pi$ will also always have the same sign, expression (48) will, however, provide us with an upper bound on the value of a statistical life.

C The conditional indirect utility functions

The functional forms of the conditional indirect utility functions, $\widetilde{V}_0$ and $\widetilde{V}_1$, are specified by taking second order Taylor series expansions of each of the functions around some point, say $\{\bar{\mathbf{r}}, \bar{y}\}$. Ignoring higher order terms would yield, for $\widetilde{V}_0$:

$$\begin{aligned}\widetilde{V}_0(\mathbf{r}; y) \approx \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y}) + (\mathbf{r} - \bar{\mathbf{r}})' \frac{\partial \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial \mathbf{r}} + \\ + (y - \bar{y}) \frac{\partial \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial y} + \frac{1}{2} (\mathbf{r} - \bar{\mathbf{r}})' \frac{\partial^2 \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial \mathbf{r} \partial \mathbf{r}'} (\mathbf{r} - \bar{\mathbf{r}}) + \\ + \frac{1}{2} (y - \bar{y})^2 \frac{\partial^2 \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial y^2} + (\mathbf{r} - \bar{\mathbf{r}})' \frac{\partial^2 \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial \mathbf{r} \partial y} (y - \bar{y})\end{aligned} \quad (49)$$

Similarly, a second order Taylor series expansion of $\widetilde{V}_1$ around $\{\bar{\mathbf{r}}, \bar{y} - \bar{p}\}$ yields (again ignoring higher order terms):

$$\begin{aligned}\widetilde{V}_1(\mathbf{r}; y - p) \approx \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y} - \bar{p}) + (\mathbf{r} - \bar{\mathbf{r}})' \frac{\partial \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y} - \bar{p})}{\partial \mathbf{r}} + \\ + [(y - \bar{y}) - (p - \bar{p})] \frac{\partial \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y} - \bar{p})}{\partial y} + \frac{1}{2} (\mathbf{r} - \bar{\mathbf{r}})' \frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y} - \bar{p})}{\partial \mathbf{r} \partial \mathbf{r}'} (\mathbf{r} - \bar{\mathbf{r}}) + \\ + \frac{1}{2} [(y - \bar{y}) - (p - \bar{p})]^2 \frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y} - \bar{p})}{\partial y^2} + \\ + (\mathbf{r} - \bar{\mathbf{r}})' \frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y} - \bar{p})}{\partial \mathbf{r} \partial y} [(y - \bar{y}) - (p - \bar{p})]\end{aligned} \quad (50)$$

Subtracting (49) from (50) yields:

$$
\begin{aligned}
\widetilde{V}_1(\mathbf{r}; y-p) - \widetilde{V}_0(\mathbf{r}; y) \approx & \\
& \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p}) - \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y}) + && \text{(51a)}\\
& + (\mathbf{r}-\bar{\mathbf{r}})' \left(\frac{\partial \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial \mathbf{r}} - \frac{\partial \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial \mathbf{r}} \right) + && \text{(51b)}\\
& + (y-\bar{y}) \left(\frac{\partial \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial y} - \frac{\partial \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial y} \right) - && \text{(51c)}\\
& - (p-\bar{p}) \frac{\partial \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial y} + && \text{(51d)}\\
& + \frac{1}{2} (\mathbf{r}-\bar{\mathbf{r}})' \left(\frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial \mathbf{r} \partial \mathbf{r}'} - \frac{\partial^2 \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial \mathbf{r} \partial \mathbf{r}'} \right) (\mathbf{r}-\bar{\mathbf{r}}) + && \text{(51e)}\\
& + \frac{1}{2} (y-\bar{y})^2 \left(\frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial y^2} - \frac{\partial^2 \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial y^2} \right) + && \text{(51f)}\\
& + \frac{1}{2} (p-\bar{p}) \left[(p-\bar{p}) - 2(y-\bar{y}) \right] \frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial y^2} + && \text{(51g)}\\
& + (\mathbf{r}-\bar{\mathbf{r}})' \left(\frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial \mathbf{r} \partial y} - \frac{\partial^2 \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial \mathbf{r} \partial y} \right) (y-\bar{y}) - && \text{(51h)}\\
& - (\mathbf{r}-\bar{\mathbf{r}})' \frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial \mathbf{r} \partial y} (p-\bar{p}) && \text{(51i)}
\end{aligned}
$$

The first term of this expression, (51a), is just the mean difference in utility between choosing to consume x_1 and not consuming it, for individuals with characteristics $\{\bar{\mathbf{r}}, \bar{y}, \bar{p}\}$. The next row, (51b), can be interpreted as the influence of individual characteristics, $\mathbf{r}$, on the utility from consuming x_1, or rather on the difference in utility from consuming x_1 and consuming the numéraire good, x_0. The term in (51c) is the difference in the marginal utility of income when x_1 is bought at price $\bar{p}$ and when it is not bought. This term can also be described in terms of the marginal utility of consuming the numéraire good. This reasoning also applies to (51d). Thus, (51d) measures the sacrifice, in utility terms, from reducing the consumption of the numéraire good necessary to purchase the discrete good. Thus instead of interpreting $\partial \widetilde{V}_1 / \partial y$ as marginal utility of income per se, we can see it as a measure of the value to the individual from consuming the discrete good, in relation to the utility from consuming the numéraire good. The interpretation of the second order terms, i.e. (51e) to (51i), follow quite naturally from the interpretation of the corresponding first order terms, (51b) to (51d).

To reduce the number of parameters we need to estimate, a few simplifying assumptions are needed. We will assume:

$$\frac{\partial \widetilde{V}_1(\bar{\mathbf{r}}; y-p)}{\partial y} \approx \frac{\partial \widetilde{V}_0(\bar{\mathbf{r}}; y)}{\partial y} \tag{52}$$

$$\frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; y-p)}{\partial \mathbf{r} \partial \mathbf{r}'} \approx \frac{\partial^2 \widetilde{V}_0(\bar{\mathbf{r}}; y)}{\partial \mathbf{r} \partial \mathbf{r}'} \tag{53}$$

$$\frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; y-p)}{\partial y^2} \approx \frac{\partial^2 \widetilde{V}_0(\bar{\mathbf{r}}; y)}{\partial y^2} \tag{54}$$

$$\frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; y-p)}{\partial \mathbf{r} \partial y} \approx \frac{\partial^2 \widetilde{V}_0(\bar{\mathbf{r}}; y)}{\partial \mathbf{r} \partial y} \approx 0 \tag{55}$$

The first of these assumptions implies two things: 1) The marginal utility of income should be the same, whether or not the discrete good is consumed, and 2) $\bar{p}$ should be small enough to have a negligible effect on the marginal utility of income. These are basically the assumptions made to arrive at the welfare measure (23).

Assumption (54) follows from assumption (52). Assumptions (53) and (55) are ad hoc assumptions made so that it will be feasible to carry out estimation with a rather small data set. Without these assumptions, multicollinearity is bound to be a severe problem.

With these assumptions, expression (51) can be written:

$$\begin{aligned} \widetilde{V}_1(\mathbf{r}; y-p) - \widetilde{V}_0(\mathbf{r}; y) \approx & \\ & \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p}) - \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y}) + \\ & + (\mathbf{r}-\bar{\mathbf{r}})' \left(\frac{\partial \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial \mathbf{r}} - \frac{\partial \widetilde{V}_0(\bar{\mathbf{r}}; \bar{y})}{\partial \mathbf{r}} \right) - \\ & - (p-\bar{p}) \frac{\partial \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial y} + \\ & + \frac{1}{2} (p-\bar{p}) \left[(p-\bar{p}) - 2(y-\bar{y}) \right] \frac{\partial^2 \widetilde{V}_1(\bar{\mathbf{r}}; \bar{y}-\bar{p})}{\partial y^2} \end{aligned} \tag{56}$$

We may thus write the final model specification as:

$$\begin{aligned} \widetilde{V}_1 - \widetilde{V}_0 = \beta_0 + (\mathbf{r}-\bar{\mathbf{r}})' \gamma - \beta_1 (p-\bar{p}) + & \\ + \beta_2 (p-\bar{p}) \left[(p-\bar{p}) - 2(y-\bar{y}) \right] & \end{aligned} \tag{57}$$

where the β:s and the column vector γ are parameters to be estimated.

Model		A	B	C
Nobs		761	761	373
LogL		-334.18	-336.85	-193.84
α_1	*DIST*	-1.18 (1.55)	0.812 (4.26)	-6.00 (27.3)
α_2	y	-4.38 (3.57)	-	-
α_3	*TOWNP*	1.25^{*} (0.667)	-0.730 (1.82)	-
α_4	*DIST·TOWNP*	-1.12 (5.44)	6.13 (16.7)	-
α_5	*TOWNP·y*	4.27 (14.4)	41.7 (63.2)	-
α_6	*DIST·y*	0.229 (39.6)	55.8 (123)	267 (1260)
β_0	*CONST*	-1.90^{***} (0.186)	-1.64^{***} (0.164)	1.28^{***} (0.126)
γ_1	*PCOV*	2.67^{***} (0.244)	2.52^{***} (0.234)	-
γ_2	*AGE*	0.0317^{***} (0.00924)	0.0377^{***} (0.00900)	-
γ_2	y	5.33^{*} (2.75)	4.97^{**} (2.40)	2.63 (2.88)
β_1	$-p$	-0.161 (0.242)	0.255 (0.278)	0.0902 (0.399)
β_2	$p(p-2y)$	0.697^{*} (0.413	-	-

Table 1 – The table displays the coefficients from three of the estimated models. Asymptotic standard errors are displayed within parentheses. All variables should be interpreted as deviations from the sample mean. One, two and three asterisks imply that the estimate is significantly different from zero in a two-tailed t-test, at least on a 10 percent, 5 percent and 1 percent level of significance, respectively.

SEK (USD)	Mean	St.dev.	Minimum	Maximum
Model B				
per woman, full sample	*331* *(41.4)*	*290* *(36.2)*	*16.2* *(2.02)*	*1190* *(149)*
ΔW per woman, only scr. prog.	*615* *(76.9)*	*114* *(14.2)*	*207* *(25.9)*	*1190* *(149)*
ΔW per mammo., full sample	*597* *(74.6)*	*189* *(23.6)*	*401* *(50.1)*	*1249* *(156)*
ΔW per mammo., only scr. prog.	*780* *(97.5)*	*85.0* *(10.6)*	*505* *(63.1)*	*1249* *(156)*
Model C				
ΔW per woman	*1710* *(214)*	*191* *(23.9)*	*1320* *(165)*	*2780* *(348)*
ΔW per mammo	*2180* *(272)*	*143* *(17.9)*	*1890* *(236)*	*3030* *(379)*

Table 2 – The table displays summary statistics of the compensated/equivalent variation, estimated from model B and model C. It should be noted that the standard deviations are standard deviations of the estimates, not the estimated standard deviations. The CV/EV per mammography is obtained by dividing the estimated welfare measure by the estimated probability. The values are expressed in SEK and, within parentheses, USD. (Exchange rate, 1 USD ≈ 8 SEK.)

Essay 3

Green taxes, eco-labeling and the auto market

Green taxes, eco-labeling and the auto market

F. Mikael Sandström*
Stockholm School of Economics

Abstract

We study the effects of eco-labeling and tax incentives on the composition of the car park. To this end, we model the utility to the consumer from owning a given car model, taking into account that cars are consumer durables. A discrete choice model is used to obtain functions for the aggregate demand for each car model. The auto producers are modelled as price setting producers of differentiated goods.

The model is estimated on a comprahensive and detailed data set over the sales of different car models in Sweden over the period 1991-1997, using instrumental variable estimation in a panel data setting. Taxes, exept for the fuel tax, seem to have had a limited role in changing the composition of the car park in a "green" direction. In contradistinction, an attempt to use "eco-labeling" to reach this objective appears to have been successful.

*I am grateful to assistance and valuable comments from Magdalena Andersson, Marcus Asplund, Per Eskilsson, Richard Friberg, Almas Heshmati, Per-Olov Johansson, Dennis Nordberg, Kasper Roszbach, Henrik Sandström and seminar participants at the Stockholm School of Economics. This research has been financed by the Stockholm School of Economics and the Government Committee on Automobile Taxation (trafikbeskattningsutredningen).

1 Introduction

This paper examines the effects of automobile taxation on the composition of the car park. In particular, we are interested in if taxes are effective instruments to give the automobile park a more "environmentally friendly" composition. We also consider the effects of "eco-labeling" of cars.[1]

Historically, taxes have been levied on automobiles mainly for fiscal reasons. Lately, however, the harmful effects of road traffic to the environment have been used as argument for increases in automobile taxation, and for alterations to the tax structure. In Sweden, environmental considerations have played a role in the design of the system for taxation of automobiles at least since 1986, when the registration tax for cars with cathalytic converters was reduced. In 1991, an environmental classification system was introduced for cars of model year 1993 and later. Under this system "environmentally friendly" automobiles obtain tax credits. It also has features of an eco-labeling system. The system has later been altered several times, and has affected, during different periods, both the registration tax and the road tax. The taxation system has also undergone other important changes during the 1990s. We analyze what effects taxes and the eco-labeling system have had on the composition of automobile sales using comprehensive and detailed Swedish data for the years 1991 through 1997.

The environmental classification system has some interesting features. In part, it is an eco-labeling system, by which consumers are informed about which products are less harmful to the environment. However, it is also combined with economic incentives, in the form of tax rebates to consumers who buy "environmentally friendly" automobiles. The system by which these tax rebates have been given has changed during the time the system has been in operation. This provides an opportunity to distinguish between the "eco-labeling" effect and the tax incentive effect. Thus, we may be able to determine if consumers primarily respond to the economic incentives, or if environmental considerations in themselves have a discernible effect on the car purchasing decision.

Many studies that analyze taxes assume, implicitly or explicitly, that the market is characterized by perfect competition. This is also the case with most studies of automobile demand.[2] However, as Goldberg [16] points out, such studies are likely to be of limited relevance when we want to analyze the effects of policy, since they will not capture how the market as a whole works. To remedy this problem we borrow a model from the industrial organization literature. We use Berry's [4] model of oligopolistic competition on the market of a differentiated product. To my knowledge, the only study previous to the present one that has studied the effects of automobile taxation in a similar framework was made by Fershtman, Gandal and Markovich [13].[3]

[1] Under an eco-labeling scheme, an authorizing body – a government agency or an NGO – grants firms the right to use a label on products that satisfy certain environmental requirements. The purpose of such a scheme is to facilitate for consumers to choose products that are less harmful to the environment.

[2] Train[28] summarizes a few such studies of automobile demand. Two quite recent examples, on Norwegian data, are Brandemoen[7] and Wetterwald[32].

[3] Fershtman, Gandal and Markovich[13] claims that only two studies previous to theirs analyze taxes

A pertinent issue in empirical applications of any model with imperfect competition is that prices and quantities will be endogenous, since they are simultaneously determined in the model. Estimation techniques that do not take account of this will produce inconsistent estimates. It is therefore important to find appropriate instruments. Previous studies using Berry's [4] approach have used cost shifters and the characteristics of the automobiles to construct instruments.[4] In this paper, a slightly different approach will be used. We construct a model of how a consumer's utility from owning a car of a given model is determined. In particular, we acknowledge that an automobile is a consumer durable, and that the utility will thus depend on the net present value of costs and benefits derived from the car during a considerable time into the future. From this model, we obtain restrictions on the interaction between different variables in the demand functions for the different car models. These are then used to construct instruments. Thus, an additional objective with the paper is to show how consumer theory may be used to solve the endogeneity problem intrinsic to models with imperfect competition.

Changes in taxation simplify the task of identification since the different forms of taxes levied on automobiles will qualitatively differ in their effects on the cost of owning a car. In most industrialized countries, automobile ownership and use are subject mainly to three different classes of taxes. Firstly, taxes are levied on the sales of automobiles as some form of registration tax. This tax thus increases the capital cost of owning a car. Secondly, owners of automobiles pay some periodic tax, such as a road tax, which increases the fixed cost of being a car owner. Finally, car fuel is taxed, thus increasing the flexible cost of car use. Since road taxes and registration taxes are usually differentiated over car models, such taxes may be used as "demand shifters". This is also the case with the fuel tax, which is essentially the same for large groups of cars, but affects the cost of running them differently, since the gas milage differ between different car models.

The paper is organized as follows. We begin by constructing a theoretical model of the car market, in Section 2. In Section 3 the data are described and some estimation issues are discussed. In that section, we also describe the Swedish system for taxation of automobiles, and the environmental classification system. The estimation results are presented and interpreted in Section 4. In the final section, the main conclusions are summarized.

in an oligopoly framework. Bernett, Keeler and Hu[3] analyze the incidence of cigarette excise taxes, and Levinsohn[23] studies the effects of tariffs on the American automobile industry. The methodology of these two papers, however, differ considerably from Berry's [4] approach. In addition to these two studies, Doyle[9] has performed a study of how the sales tax is shifted to buyers of cars, using household data on car purchases.

[4]Both Fershtman, Gandal and Markovich[13] and Verboven[30],[31] make use of Berry's [4] observation that in models with product differentiation, and with exogenous product characteristics, the product characteristics of all firms other than firm i will constitute appropriate instruments for the price(s) of the product(s) of firm i.

2 A theoretical model of the automobile market

The setup of the model is as follows. We assume that there are $N+1$ mutually exclusive alternatives open to the consumer. He can either buy one of the N different car models, or he can choose not to be a car owner. The model consists of three parts. Firstly, in Subsection 2.1, we model the consumers conditional utility from each of the available alternatives. Essentially we construct a model of the net present value to the consumer from each alternative. Secondly, we model how the choice between the different alternatives is made. (Subsection 2.2.) Following Berry[4], we use a multinomial logit setting. From the individual choices, aggregate demand functions for each car model can be constructed. Finally, in Subsection 2.3, we consider the automobile producers, who are assumed to act as price setters.

2.1 The consumers' conditional utility

In this subsection, we will derive an expression for the consumer's utility given that he makes a certain choice. Thus, we will not yet consider how the choice between different alternatives is made.

When a consumer purchases an automobile, we must assume that he is interested in the total value of the car during its entire life. Even if he does not intend to keep the car himself all the time until it is scrapped, he will be interested in the price he will receive when the car is sold. This price, in its turn, will depend on the utility the next owner will derive from the car. To simplify, we assume that each owner in fact keeps the car for its entire life. Thus, we ignore the market for second hand cars. When buying a new car, the consumer will also be interested in what happens after he disposes of it, i.e. when the car reaches the end of its life-length. When the car is scrapped, the consumer will either buy a new car, or he will choose not to be a car owner. To take account also of this fact, a version of the chain-of-replacement model (see e.g. Massé[24]) will be used. We will thus assume an infinite horizon model.

Consider the utility that individual j derives given that he owns a car of model i. Both the benefit and the cost of owning the automobile may consist of elements that are related to the age of the car, as well as of elements that are unrelated to the age of the car. We will also assume that utility may be divided into one part that is common to all consumers, and one part that differs over consumers. We specify the instantaneous utility to individual j from owning a car of model $i = 1, ..., N$ and of age r as follows:

$$u_{i,j}(r) = v_i + w_i(r) + a_{i,j}$$

In other words, we assume that the instantaneous utility is additive in, firstly, a term v_i, which is invariant across individuals, and also is independent of the age of the car, secondly, in a term $w_i(r)$, which is invariant across individuals, but which depends on the age of the car, and thirdly, in a term $a_{i,j}$, which may differ between individuals, but which is independent of the age of the car. We will assume that the latter term is deterministic and known to the individual, but unobservable. Since the instantaneous utility from owning the car is likely to be lower, the older is the car, we will assume that $w_i(r)$ is strictly decreasing. Below, in Subsection 2.2, we will specify v_i and $w_i(r)$ to be functions of vehicle characteristics, taxes, etc. For now, however, we treat v_i as a constant and $w_i(r)$ as a function of the age of the car only.

Next, assume that the instantaneous utility of not being a car owner can be written:

$$u_{0,j} = v_0 + a_{0,j} \tag{1}$$

where v_0 is thus common to all individuals while $a_{0,j}$ is an individual specific constant. In subsection 2.2, we will define v_0 as a function of time dependent variables. For now, we may treat it as a constant.

When the automobile is purchased the consumer will pay the retail price and the registration tax. On the total purchase price the consumer will also have to pay a value added tax. Thus, the price of the car will be:

$$p_i = (1 + \tau^v)(q_i + \tau_i^s) \tag{2}$$

where q_i is the retail price, exclusive of taxes, τ^v is the value added tax rate, and τ^s is the registration tax.

Assuming a simple utility function, which is linear in income, the net present value of the benefit to individual j from owning a car of model i, $i = 1, ..., N$, over its entire life, may then be written:

$$\begin{aligned} B_{i,j} &= -\lambda(1 + \tau^v)(q_i + \tau^s) + \int_0^{T_j} e^{-\rho r}(v_i + w_i(r) + a_{i,j})\, dr = \\ &= -\lambda(1 + \tau^v)(q_i + \tau_i^s) + \frac{1}{\rho}\left(1 - e^{-\rho T_j}\right)(v_i + a_{i,j}) + \int_0^{T_i} e^{-\rho r} w_i(r)\, dr \end{aligned} \tag{3}$$

where λ is the marginal utility of income, which we assume is constant and identical for all individuals. T_i is the number of years the car is kept before it is scrapped and ρ is the discount rate.

Assume that the choices available to the consumer will be the same every time he buys a new car. In an infinite horizon model, then, if a car of model i is the best choice for individual j at one choice occasion, it will be so also on all other choice occasions, since the decision problem will be identical every time a new car is bought. We may thus view a consumer who buys a car of model i as a permanent "model i driver", and consider the net present value of owning car i over an infinite horizon, with periodic replacements of the vehicle.

Suppose that the total net present value from being a "model i driver", for $i = 1, ..., N$, to individual j is $\widetilde{U}_{i,j}$. Since this utility will be the same every time the car is replaced we may write:

$$\begin{aligned}\widetilde{U}_{i,j} &= B_{i,j} + e^{-\rho T_i}\widetilde{U}_{i,j} \\ &\Rightarrow \\ \widetilde{U}_{i,j} &= \frac{1}{1-e^{-\rho T_i}}B_{i,j}\end{aligned} \tag{4}$$

Using (3), this may be written:

$$\widetilde{U}_{i,j} = -\lambda\frac{1+\tau^v}{1-e^{-\rho T_i}}\left(q_i+\tau_i^s\right) + \frac{1}{\rho}\left(v_i + a_{i,j}\right) + \frac{\int_0^{T_i} e^{-\rho r} w_i\left(r\right) dr}{1-e^{-\rho T_j}} \tag{5}$$

which defines a version of the "chain-of-replacement" model. (See e.g. Massé[24].) In such a model, T_i is chosen to maximize the utility from alternative i. (See Sandström[25] for a discussion, and an application to automobiles.) However, in this paper, we will assume that it is taken as given by the consumers.

The net present value utility of not owning a car will simply be the value of an annuity equal to $u_{0,i}$ as defined in (1):

$$\widetilde{U}_{0,j} = \frac{1}{\rho}\left(v_0 + a_{0,j}\right) \tag{6}$$

Since the $\widetilde{U}_{i,j}$:s, $i = 0, ..., N$ will be the utility *given* that i is chosen among the $N+1$ alternatives, we may term these functions conditional utility functions. We have assumed that the consumer will always choose exactly one of the alternatives. Thus, given all variables and parameters in the model, a representation of the unconditional utility of individual j may be written:

$$U_j = \sum_{i=0}^{N} a_i \widetilde{U}_{i,j} \tag{7}$$

where a_i is an index variable equal to one if i is chosen and zero otherwise.

A rational consumer will choose the alternative that maximizes this expression. Given a suitable parameterization of the $u_{i,j}\left(r\right)$:s, $i = 1, ..., N$ and of $u_{0,j}$, and some assumption of the distribution of the unobservable $a_{i,j}$:s, $i = 0, ..., N$ over the population, this specification of utility can then be used to construct a demand system that may be tested empirically. However, a complication is the interaction in the conditional utility functions between the discount factor and the unobservables. This problem is easily solved if we take advantage of the fact that if (7) is a valid utility function, then any monotonous transformation is also a valid utility function. Thus, we may simply multiply through with the discount factor to obtain functions which are additive in the unobservable factors, $a_{i,j}$:s, $i = 0, ..., N$. In other words, we may as well write the conditional utility functions as:

$$\widetilde{U}'_{i,j} = \rho\widetilde{U}_{i,j}, i = 0, ..., N \tag{8}$$

and the unconditional utility as:

$$U'_j = \rho U_j = \sum_{i=0}^{N} a_i \widetilde{U}'_{i,j} \tag{9}$$

If we then define the individual invariant part of the conditional utility by:

$$\begin{cases} V_0 = v_0 \\ V_i = -\lambda \frac{\rho(1+\tau^v)}{1-e^{-\rho T_i}} \left(q_i + \tau_i^s\right) + v_i + \frac{\rho \int_0^{T_i} e^{-\rho r} w_i(r) dr}{1-e^{-\rho T_i}}, i = 1, ..., N \end{cases} \tag{10}$$

we obtain a model which is strikingly similar to the formulation used in the standard linear random utility maximization (LRUM) model, since we may write:

$$\widetilde{U}'_{i,j} = V_i + a_{i,j}, i = 0, ..., N \tag{11}$$

The V_i may then be interpreted as the mean utility across all consumers, from choosing alternative i. (See Anderson, de Palma and Thisse [1] for an extensive discussion of alternative interpretations of the LRUM model.) From this model of the consumer's utility given that a certain alternative is chosen, we can derive a system of demand equations. How that may be accomplished is the subject of the next subsection.

2.2 Demand equations

In this subsection we will use the conditional utility functions constructed in the last subsection to construct aggregate demand functions for the different car models. We proceed by first assuming a distribution for the individual specific factors, the $a_{i,j}$:s, and then by defining the remaining components of the instantaneous utility functions as functions of automobile characteristics, taxes and of the price of the car. The utility of not owning a car is defined as a function of time dependent variables.

Assume that the $a_{i,j}$:s follow the same, known and well behaved[5] distribution $f(a_{i,j})$. Then the probability that a randomly selected individual will choose alternative $i, \forall i = 0, ..., N$ will be:

$$\psi_i = prob\left\{V_i + a_{i,j} > V_k + a_{k,j}, \forall\left(k = 0, ..., N \wedge i \neq k\right)\right\} = \\ = \int_{-\infty}^{\infty} \prod_{k \neq i} F\left(V_i - V_k + a_{i,j}\right) f\left(a_{i,j}\right) da_{i,j} \tag{12}$$

where $F(s) = \int_{-\infty}^{s} f(a_{i,j})\, da_{i,j}$ is the cumulative distribution of $a_{i,j}$. Given a parameterization of v_i and $w_i(r)$, and a specific assumption about the distribution $f(a_{i,j})$, the ψ_i:s can be used to construct a likelihood function, which may be used in estimation.

[5] By "well behaved" we mean that $f(a_{i,j})$ is integrable, continuous and everywhere non-negative, and that $\lim_{s \to -\infty} f(s) = \lim_{s \to \infty} f(s) = 0$.

But this would imply that the only unobservable factors affecting demand are the individual specific effects, i.e. the $a_{i,j}$:s. This is hardly realistic. Rather, we must allow for unobservable factors to affect either v_i or $w_i(r)$, or both. However, this implies that the probabilities (12) are non-linear in the unobservables.

Berry[4] suggests a solution to this problem. Let us first assume that the $a_{i,j}$:s follow the type I extreme value (double exponential) distribution. This gives us the familiar multinomial logit specification of the probabilities:

$$\psi_i = \frac{e^{V_i}}{\sum_{k=0}^{N} e^{V_k}}, \forall i = 0, ..., N \tag{13}$$

If the size of the market is M, then the aggregate demand for cars of models $i = 1, ..., N$ is:

$$y_i = M\psi_i = M\frac{e^{V_i}}{\sum_{k=0}^{N} e^{V_k}} \tag{14}$$

and the number of consumers choosing the zero alternative will be:

$$y_0 = M\frac{e^{V_0}}{\sum_{k=0}^{N} e^{V_k}} = M - Y \tag{15}$$

where we have defined $Y = \sum_{i=1}^{N} y_i$ to be the total number of cars sold of all models.

Now, by a simple transformation (see Berry[4] for details) we may use (14) and (15) to obtain functions of the aggregate demand for the different car models, which are linear in the functions V_i, $i = 0, ..., N$. If we divide (14) by (15), the denominators will cancel out. Taking the natural logarithm of the resulting expression, we obtain, for each $i = 1, ..., N$:

$$\ln \frac{y_i}{M - Y} = V_i - V_0 \tag{16}$$

If we assume that the unobservables enter linearly in the $V_i : s$, $i = 0, ..., N$, this expression will thus also be linear in the unobservables.

To transform these functions into demand functions that may be estimated, we first consider the v_i:s, $i = 1, ..., N$. We now treat these as functions. Recall that v_i is the part of the instantaneous utility from owning a car of model i that is independent of the age of the car. We assume that this function can be written as a linear function of the characteristics of the automobile model, say a column vector $\mathbf{x}_i$, of the annual road tax, τ_i^r, the cost of gasoline per kilometer driven, i.e. the product of the price of gasoline and the car's gas use per kilometer, which we call by g_i, and of a model specific "fixed effect", η_i. The latter is supposed to capture unobservable model specific characteristics. In addition, we allow the function v_i to contain a purely random component, ε_i. Adding an index for time, since we will need to use time series data to obtain estimates of the fixed effects,[6] we obtain:

$$v_{i,t} = v\left(\mathbf{x}_{i,t}, \tau^{r}_{i,t}, \tau^{m}_{i,t}, g_{i,t}; \boldsymbol{\alpha}, \lambda, \gamma, \eta_i, \varepsilon_{i,t}\right) = \\ = \boldsymbol{\alpha}'\mathbf{x}_{i,t} - \lambda\left(\tau^{r}_{i,t} - \left(1 - e^{-5\rho}\right)\tau^{m}_{i,t}\right) + \gamma g_{i,t} + \eta_i + \varepsilon_{i,t} \quad (17)$$

where the column vector $\boldsymbol{\alpha}$ and the scalars λ, γ, η_i are parameters to be estimated. The term $(1 - e^{-5\rho})\,\tau^{m}_{i,t}$ requires explanation. As will be discussed below, automobiles meeting certain environmental standards have been exempted from the road tax during the first five years. Thus, we need this term to account for this tax credit.[7]

Next, let us consider the part of instantaneous utility that depends on the age of the car, i.e. $w_i(r)$. This term may be interpreted as the depreciation of the car. Assume that this depreciation is linear, in other words that we may write:

$$w_i(r) = w(r, d_{i,t}; \theta, \delta) = -\delta \cdot r \cdot d_{i,t} \quad (18)$$

where d_i is some model specific measure of the depreciation rate of the car, and δ is a parameter to be estimated. With this assumption, we can solve the integral in expression (10).

Finally, consider the individual independent instantaneous utility from not owning a car, i.e. the term v_0. Berry[4] normalizes this utility to zero. However, since we use data from several time periods, we want to allow it to vary over time. We will use the following linear specification:

$$v_{0,t} = v_0(\mathbf{s}_t; \boldsymbol{\theta}) = \boldsymbol{\theta}'\mathbf{s}_t \quad (19)$$

where $\mathbf{s}_t$ is a column vector of time dependent variables, and $\boldsymbol{\theta}$ is a column vector of parameters.

By inserting (17) and (19) into (10), and then inserting the resulting expressions into (16), we may write the final demand equation for a car of model $i = 1, ..., N$ as:

$$\ln\frac{y_{i,t}}{M - Y_t} = \\ = -\lambda\frac{\rho(1+\tau^{v})}{1 - e^{-\rho T_i}}\left(q_{i,t} + \tau^{s}_{i,t}\right) + \boldsymbol{\alpha}'\mathbf{x}_{i,t} - \lambda\left(\tau^{r}_{i,t} - \left(1 - e^{-5\rho}\right)\tau^{m}_{i,t}\right) + \gamma g_{i,t} - \\ - \delta\frac{1 - (1 + \rho T_i)\,e^{-\rho T_i}}{\rho\left(1 - e^{-\rho T_i}\right)}d_i - \boldsymbol{\theta}'\mathbf{s}_t + \eta_i + \varepsilon_{i,t} \quad (20)$$

[6] An alternative would be to use data from several markets. See e.g. Verboven [30].

[7] This term is calculated as the net present value of the tax exemption:

$$\int_0^5 e^{-\rho r}\tau^{m}_{i,t}dr = \frac{1 - e^{-5\rho}}{\rho}\tau^{m}_{i,t}$$

Since we multiply the conditional utility function by ρ to obtain expression (8), the denominator will cancel out.

where we have used (18) to solve the integral from expression (10).

Since this function is linear in parameters, it can be estimated using standard panel data estimation techniques. However, since prices, the $q_{i,t}$:s, and quantities, the $y_{i,t}$:s, are determined simultaneously, we must find some suitable instrument for the $q_{i,t}$:s. We will assume that all other variables in the demand equation are exogenous. This issue will be further addressed below, in subsection 3.3. Before that, however, we will consider the supply side of the market.

2.3 The firms' pricing decisions

Following Berry[4], we assume that the firms are price setters. We also assume that all car models are produced by single product firms. Finally, we assume that the marginal cost of producing a car is constant and can be described as a linear function of its characteristics, including an unobservable fixed effect. Thus, the producer of car i will solve:

$$\max_{q_{i,t}} \left\{ y_i \left(q_{i,t} - \boldsymbol{\beta}'\mathbf{z}_{i,t} - \boldsymbol{\phi}'\mathbf{u}_{i,t} - \mu_i - \zeta_{i,t} \right) \right\} \tag{21}$$

where $\mathbf{z}_{i,t}$ is a column vector of observable vehicle characteristics, $\mathbf{u}_{i,t}$ is a column vector of cost shifters, $\boldsymbol{\beta}$ and $\boldsymbol{\phi}$ are the corresponding vectors of parameters, and μ_i is the fixed effect. We also allow for a pure random term, $\zeta_{i,t}$, in the cost function. Assuming the existence of a pure-strategy interior equilibrium, the prices must then satisfy the first order conditions:

$$\begin{aligned} &\left(q_{i,t} - \boldsymbol{\beta}'\mathbf{z}_{i,t} - \boldsymbol{\phi}'\mathbf{u}_{i,t} - \mu_i - \zeta_{i,t} \right) \cdot \frac{\partial y_{i,t}}{\partial q_{i,t}} + y_{i,t} = 0 \\ &\Rightarrow \\ &q_{i,t} = \boldsymbol{\beta}'\mathbf{z}_{i,t} + \boldsymbol{\phi}'\mathbf{u}_{i,t} - \frac{y_{i,t}}{\partial y_{i,t}/\partial q_{i,t}} + \mu_i + \zeta_{i,t} \end{aligned} \tag{22}$$

for $i = 1, ..., N$.

From expressions (14) and (10) it is straight forward to calculate $\partial y_{i,t}/\partial q_{i,t}$:

$$\begin{aligned} \frac{\partial y_{i,t}}{\partial q_{i,t}} &= M \frac{\partial \psi_{i,t}}{\partial V_{i,t}} \frac{\partial V_{i,t}}{\partial q_{i,t}} = \\ &= -M\psi_{i,t}\left(1 - \psi_{i,t}\right) \lambda\rho \frac{1+\tau_t^v}{1-e^{-\rho T_i}} = \\ &= -y_{i,t}\left(1 - \frac{y_{i,t}}{M}\right) \lambda\rho \frac{1+\tau^v}{1-e^{-\rho T_i}} \end{aligned} \tag{23}$$

Thus firm i:s pricing equation may be written, for $i = 1, ..., N$:

$$q_{i,t} = \boldsymbol{\beta}'\mathbf{z}_{i,t} + \boldsymbol{\phi}'\mathbf{u}_{i,t} + \frac{1}{\lambda} \frac{1-e^{-\rho T_i}}{\rho\left(1 - \frac{y_{i,t}}{M}\right)\left(1+\tau_t^v\right)} + \mu_i + \zeta_{i,t} \tag{24}$$

This expression, together with the expression for the demand of products $i = 1, ..., N$ (20) constitutes a simultaneous equation system, since the dependent variable in the demand equation enters as an explanatory variable in the pricing equation, and vice versa. Congruently with the assumptions behind the demand equation, we will assume that all variables on the right hand side of the pricing equation except of $y_{i,t}$ are exogenously determined. In the next section, we will operationalize this model, discuss the data we use, and address some estimation issues.

3 Data and estimation issues

In this section, we consider the data used in the model and how the model is operationalized. We also describe the Swedish system for taxing automobiles and the environmental classification system, and the changes they have undergone. Finally, we address some estimation issues, most importantly, the issue of constructing instruments for the endogenously determined variables.

3.1 Data

To estimate our model, we need data on vehicle characteristics, the price of automobile fuel, on taxes, and on some time dependent variables that may affect the utility of not owning a car. We also need some measure of the consumers' discount rate. In addition, it is helpful if we can include some cost-shifters in the pricing equation. Naturally, we also need data on the number of cars sold of each car model, and on the price of these. We leave the discussion of data on taxes to Subsection 3.2. Here, we instead discuss the other groups of variables, beginning with the data on automobile sales and prices. We also defer the discussion of the environmental classification system to subsection 3.2.

3.1.1 Automobile sales and prices

Monthly data on the number of cars sold of each model were obtained from Statistics Sweden (SCB). The SCB data on car sales is highly disaggregate. The database distinguishes between over 30 000 "models". To make the data meaningful, they were aggregated over car models such as Volvo 745, Audi 100 and Toyota Corolla. Only cars produced by manufacturers that during any one of the years covered by the database sold at least 2 000 cars were included.[8] Since some of the other variables used were quarterly data, data on sales was also transformed to quarterly data.

In the data material, three size-groups of cars can be distinguished. Very small cars, which essentially mean sports cars, and very large cars, which are mainly vans, off-road vehicles etc., were excluded from the analysis. Thus, only "ordinarily" sized cars were included. In addition, some car models that were only sold during one or very few periods were excluded. After this selection, 120 models remained.

[8] 2 000 cars sold during a year correspond to around 5-10 percent of total sales of cars in Sweden.

During any one quarter, between 61 and 88 models were sold. Since it is reasonable to think of sales of cars to firms as a market separate from sales to private persons, only the latter category was included in the analysis. The data used covers the first quarter 1991 through the second quarter 1997. Thus, the total number of periods is 26.

Price information was obtained from Autograph AB, which is a private firm that assembles and sells technical data and price data on automobiles. The price variable they provide is the consumer price including VAT and taxes. Thus, this variable needs to be transformed to obtain the pre-tax price when it is used in the pricing equation. Autograph states that the price they list is "an expression of the general agent's conception of the going price level for a car with 'standard equipment'."[6] (Authors translation.) The prices are deflated using the consumer price index.

3.1.2 Vehicle characteristics

Some of the characteristics of a vehicle will enter both the demand equation and the pricing equation in a linear fashion. This is the case with the weight of the car, the size of the car, its top speed, and the horsepower of the engine. There is no obvious reason why we should assume that these variables would interact with eachother or with the other variables in any particular way. However, the automobile's gas use per kilometer interacts with the cost of gas in the demand equation. In fact, we assume that the consumer is not primarily interested in the car's gas use per se, but rather with the cost of driving it. In the pricing equation, on the other hand, we do not expect the gas price to play any role. Data on car weight, size, top speed, engine power and gas use were obtained from Autograph AB. These data are mainly from the vehicle approval records of the Swedish Motor Vehicle Inspection Company, (Svensk bilprovnings typbesiktningsprotokol) and data from the manufacturers and importers of the cars. The top speed is an estimate based on test results published in technical journals, and the size index is a sum of the "nine most important" measures of internal height, width and length.[6] As a measure of the gas price, the consumer price index for vehicle fuel, in relation to the general CPI, was used.

Another variable that will enter the pricing equation and the demand equation differently is the measure of the depreciation of the car. As such a measure, we use the "value loss" factor that the Swedish Consumer Authority uses to calculate the cost of owning an automobile.[9] It will be included in the demand equation as a proxy of $d_{i,t}$ and will thus interact with the discount factor and the life length of the car. In the pricing equation, it will enter linearly as a cost variable.

The life length of automobiles of different models, the T_i:s, will only enter the demand equation,[10] and then only in interaction with other variables. The reason for this is that it seems unreasonable to allow this variable to vary over time. (We have already indicated

[9] I am obliged to Dennis Nordberg at the Consumer Authority for giving me access to these data.

[10] Ideally, we would like to estimate the T_i:s, since the life length of a car is really the solution to an optimization problem. However, that would put too great a demand on the data.

this by omitting a time index on this variable.) Thus it cannot be included linearly in a fixed effects model. Obviously, we cannot observe T_i for a new car, since we do not know when a car bought today will be scrapped. We must thus use historical data. Swedish data on the life length of cars were unavailable, but from Statistics Norway[29] data were obtained on the average age of scrapped cars of different makes.

3.1.3 The discount rate, the size of the market, cost shifters and the utility of not owning a car

The discount rate plays an important role in our model, since it is by this rate we make the capital cost of owning a car comparable to the running costs of the car. As a measure of the discount rate, we use the real bank consumer lending rate. The nominal bank consumer lending rate is obtained from the IFS database.[20] To calculate the real interest rate, we would want a measure of the expected inflation rate. In lieu of a more precise measure, we use the actual inflation rate, calculated as the percentage change in the consumer price index from one year earlier.

The dependent variable in the demand equation is not the share of total sales of automobiles, but the share of the entire market, including consumers choosing not to own a car. Thus we need to define the size of the market. One possibility would be to define the market as the total number of households in the economy. However, that would be somewhat awkward. We have assumed that all car owners own their car until it is scrapped. Thus, an owner of a car that has not reached the end of its life will not be in the market for a new car. Rather, we would expect that the number of car owners who are in the market for a car at a given time is approximated by the size of the car stock, divided by the average life-length of a car. To arrive at the size of the market, we add to this number the number of households that are not car owners, since they may at any time decide to become car owners, if the determining factors change.

Naturally, we would like to include as many cost shifters as possible in the pricing equation, i.e. in the vector $\mathbf{u}_{i,t}$. In the present study, however, the only cost shifter we were able to include is changes in exchange rates. Data on exchange rates were obtained from the IFS database. [20] We define the exchange rate by the number of units of the national currency in the country of origin of the car model per Swedish kronor (SEK).

The utility from not owning a car is determined by the vector $\mathbf{s}_t$. Presumably, the cost of other means of transport would influence the likelihood that people are car owners. To capture this, the consumer price index of "travel and transports", divided by the general CPI was used. Also, we may assume that the general welfare of the consumers would influence the relative benefits of being and not being a car owner. Thus, we also include the gross national product as an explanatory variable in the demand equation. To obtain a stationary series, we take the first difference, and to avoid the seasonality of the GDP series, we take the difference on one year earlier.

In both the demand equation and in the pricing equation time trends and quadratic time trends are included. The reason for including the quadratic terms is that we want to allow the trends to be non-linear. In the pricing equation, we interpret the trend terms as cost shifters, i.e. as elements of the vector $\mathbf{u}_{i,t}$. In the demand equation, we interpret them as components of the vector $\mathbf{s}_t$. We view the trend as a "sink" for variations of the cost and benefit of not owning a car that are not accounted for by the other variables.

3.2 Automobile taxes and eco-labeling[11]

We now turn to a description of the Swedish system for taxation of automobiles and of the environmental classification system. We also describe the data on taxes, which we use.

In Sweden, environmental considerations have played a role in designing the tax system for automobiles at least since 1986. At that time, the registration tax for automobiles with catalytic converters was reduced, the purpose being to equalize the price of cars with and without catalytic converters. In 1991, an environmental classification system was introduced, applying to cars of model year 1993 and later. When the system was introduced, car models were registered in one of three "environmental classes", MK1, MK2 or MK3.[12] Cars belonging to environmental classes 1 or 2 had to meet certain environmental standards, more stringent than those established by law. The requirements for cars belonging to MK1 were again somewhat more stringent than for those belonging to MK2.

To give an incentive for people to buy more environmentally friendly cars, the registration tax, which was levied in relation to the weight of the car, was lowered by SEK 4 000 for cars belonging to MK1, was kept the same for cars in MK2 and was raised by SEK 2 000 for cars in MK3. When Sweden entered the European Union on January 1, 1995, the differentiation of the registration tax between MK1 and MK2 automobiles was judged not be in accordance with the acquis communautaire. As a result, the reduction of the registration tax for MK1 cars was removed. Instead, such cars got a five year exemption from the road tax. On January 1, 1996, the road tax was raised by 10 percent, thus increasing the size of the tax rebate for MK1 cars.[13]

On June 12, 1996, the registration tax was removed for MK1 and MK2 cars, and reduced to SEK 2 000 for MK3 cars. On January 1, 1997, it was abolished altogether, also for MK3 cars. A stated motive behind this latter change was that the EU commission's directive 94/12/EG, which then came into force, established minimum standards for automobiles that were equivalent to the former Swedish standards for MK2 cars. In effect, MK3 cars could thus no longer be sold. It was deemed undesirable to increase the registration tax on cars fulfilling the old requirements for MK2, and thus the tax was abolished. In connection with this, the road tax was increased again, this time by 50 percent. The shift of taxation from the registration tax to the road tax was also motivated by a wish to stimulate sales of new cars, in order to reduce the average age of

[11]The description of system for taxation of automobiles builds on SOU 1997:126. [27]

[12]"MK" is the Swedish acronym for "miljöklass", i.e. environmental class.

[13]As of June 1999, SEK 1 $\approx$ EURO 0.11 $\approx$ USD 0.12.

the vehicle stock. As pointed out above, the fact that the tax incentives have changed makes it feasible to distinguish between the eco-labeling effect and the tax incentive effect in our model.

We consider the environmental class to be a characteristic of the automobile. Thus, it is included both in the vector that enters the demand equation, i.e. the vector $\mathbf{x}_{i,t}$, and in the pricing equation, i.e. the vector $\mathbf{z}_{i,t}$. This variable takes the values 1, 2 or 3, corresponding to the car model's environmental classification, where the car is thus less harmful to the environment the lower is this value. Thus, if consumers take "green" considerations when they buy a new car, then the parameter corresponding to this variable in the demand equation should take on a negative value. It is presumably not costless for the auto producer to make more environmentally friendly cars. Thus we would expect it to have a negative sign also in the pricing equation. Data on environmental classification were obtained from the Swedish Consumer Authority (Konsumentverket).

It should be clear how the different taxes enter the demand function from the specification in (20). Data on these taxes were either obtained directly from Autograph AB, or calculated based on the information in SOU 1997:126[27]. All taxes were deflated using the consumer price index. Gasoline taxes are not included explicitly, but enter through the gas cost variable in the demand equation. In the pricing equation, the dependent variable is the price exclusive of taxes, while in the demand equation price including the registration tax and value-added-tax is used in estimation.

3.3 Estimation issues

To estimate the model, we use the least square dummy variable approach, i.e. a fixed effect specification of a panel data model. However, before we discuss this further, let us summarize the model.

The demand for a car of model i can be written as:

$$\begin{aligned}\ln \frac{y_{i,t}}{M-Y_t} &= \\ &= -\lambda \frac{\rho\left(1+\tau^v\right)}{1-e^{-\rho T_i}}\left(q_{i,t}+\tau_{i,t}^s\right)+\boldsymbol{\alpha}'\mathbf{x}_{i,t}-\lambda\left(\tau_{i,t}^r-\left(1-e^{-5\rho}\right)\tau_{i,t}^m\right)+\gamma g_{i,t}- \\ &\quad -\delta\frac{1-\left(1+\rho T_i\right)e^{-\rho T_i}}{\rho\left(1-e^{-\rho T_i}\right)}d_i-\boldsymbol{\theta}'\mathbf{s}_t+\eta_i+\varepsilon_{i,t} \qquad (20)\end{aligned}$$

where the dependent variable is thus the ratio between $y_{i,t}$, which is the number of cars of model i sold during time period t, and $M - Y_t$. Y_t is the total number of cars sold, and M is the size of the market. $M - Y_t$ is thus the number of consumers choosing the "zero-alternative". The price, exclusive of taxes, is $q_{i,t}$, and the registration tax is $\tau_{i,t}^s$. The road tax is $\tau_{i,t}^r$. If the car receives a tax exemption during the first five years, then $\tau_{i,t}^m = \tau_{i,t}^r$, otherwise, $\tau_{i,t}^m = 0$. The gas cost variable $g_{i,t}$ is defined as the consumer price index for automobile fuel divided by the general CPI and multiplied by the car's gas use per 10 kilometers. The "depreciation variable" d_i is approximated by the Swedish

Consumer Agency's "value loss factor". As a proxy for the discount factor ρ we use the real bank consumer lending rate. The life-length of automobiles T_i is treated as a known variable. We have defined the two vectors $\mathbf{x}_{i,t}$ and $\mathbf{s}_t$ by:

$$\mathbf{x}_{i,t} = \begin{pmatrix} environmenal\ class_{i,t} \\ weight_{i,t} \\ horsepower_{i,t} \\ size_{i,t} \\ speed_{i,t} \end{pmatrix}, \ \mathbf{s}_t = \begin{pmatrix} CPI_t^{transport} \\ \Delta GDP_t \\ t \\ t^2 \end{pmatrix} \tag{25}$$

The demand equation is thus linear in the parameter vector:

$$\{ \ \lambda \quad \boldsymbol{\alpha}' \quad \gamma \quad \delta \quad \boldsymbol{\theta}' \quad \eta_i \ \} \tag{26}$$

and in the error term $\varepsilon_{i,t}$. This facilitates estimation.

The price of car i is described by:

$$q_{i,t} = \boldsymbol{\beta}'\mathbf{z}_{i,t} + \boldsymbol{\phi}'\mathbf{u}_{i,t} + \frac{1}{\lambda}\frac{1-e^{-\rho T_i}}{\rho\left(1-\frac{y_{i,t}}{M}\right)\left(1+\tau_t^v\right)} + \mu_i + \zeta_{i,t} \tag{24}$$

where the dependent variable, $q_{i,t}$, is thus the price net of any taxes. The term μ_i is the "fixed effect" which may reflect cost factors not accounted for, or a model specific price mark-up. The third term on the right hand side will be a function of the demand for cars of model i, of the VAT-rate, of the discount rate, and of the life-length of the car. This term is non-linear in λ, a parameter which also appears in the demand equation. We discuss that issue below. First, however, we recall that the vectors $\mathbf{z}_{i,t}$ and $\mathbf{u}_{i,t}$ have been defined as:

$$\mathbf{z}_{i,t} = \begin{pmatrix} envionmenal\ class_{i,t} \\ weight_{i,t} \\ horsepower_{i,t} \\ size_{i,t} \\ speed_{i,t} \\ d_{i,t} \\ gasuse_{i,t} \end{pmatrix}, \mathbf{u}_{i,t} = \begin{pmatrix} e_{i,t} \\ t \\ t^2 \end{pmatrix} \tag{27}$$

where $e_{i,t}$ is the percentage change in the exchange rate of the currency of the country of origin of car model i.

Even though the model is non-linear in λ, we may consider it as a model which is linear in the parameter vector:

$$\{ \ \boldsymbol{\beta}' \quad \boldsymbol{\phi}' \quad \tfrac{1}{\lambda} \quad \mu_i \ \} \tag{28}$$

and in the error term, $\zeta_{i,t}$.

Equations (20) and (24) are both specified as fixed effect models. We could instead use the random effects specification, i.e. assume that η_i and μ_i were random variables.

However, that specification would seem unreasonable, since we are in fact using data on almost all car sales in Sweden. (See e.g. Greene [18] for a discussion.)

As the two equations are written, they contain λ in three places. The restriction that this implies on the demand equation could be incorporated by a simple transformation of the variables. However, since we are interested in the effects of variations in taxes, we do not impose this restriction. In fact, we allow the coefficients on $\tau_{i,t}^{r}$ and $(1 - e^{-5\rho})\,\tau_{i,t}^{m}$ to be different, to be able to identify if the tax rebate appears to have a discernible effect on demand. The cross-equation restriction could also be imposed in the estimation. However, acknowledging that the pricing equation rests on rather heroic assumptions, we allow this parameter also to be a free parameter.

An important issue is how instruments should be obtained for the two endogenous variables, i.e. the function of price which appears in the demand equation, and the function of the number of sold cars appearing in the pricing equation. We will simply construct all the variables appearing on the right hand side of the two equations, and then regress each of the endogenous variables on all of the thus constructed exogenous variables. The demand equation and the pricing equation are then estimated using instrumental variable estimation.

One final estimation issue with which we must deal, is the question of serial correlation among the residuals. Since it is reasonable to think that a disturbance that affects prices and sales during one period will also affect sales during the next period, we will allow for the residuals to be serially correlated by using the Cochrane-Orcutt transformation of the data. (See e.g. Greene[18].)

4 Results

We first discuss the demand equation, which is in the focus of the present study, and then comment on the results from the estimation of the pricing equation. We present the results from four different regressions: the least square estimation, the fixed effects estimation, the least square estimation allowing for serial correlation, and the fixed effect estimation allowing for serial correlation. The estimation results are presented below in Tables 1 and 2, respectively. In the last part of this section, we try to illustrate the effects of the environmental classification system on the demand for different car models.

4.1 The demand equation

As may be seen in Table 1, the results from the estimation of the demand equation are similar in the four estimations. The coefficient estimates are similar in size with the exception of the parameters on the four vehicle characteristics variables, – weight, horsepower, size and top speed – the tax credit for MK1 automobiles, and the price index for transports. The vehicle characteristics variables are highly co-linear. Thus, it is hardly surprising that these coefficients differ between different specifications of the model. Neither is it of any great consequence for the interpretation of the results. The tax credit variable is not significantly different from zero, which we will discuss further below, so we should not be surprised that we obtain diverging estimates of this coefficient with the different specifications. The price index variable, finally, is only significant in one of the regressions, and then only at the 10 percent level, so the same reasoning applies.

There is in fact indications of a rather high degree of serial correlation in the data. The estimated autocorrelation coefficient is 0.46. For this reason, we will put more trust in the autocorrelation adjusted estimates. Further, the hypothesis that the intercept terms are the same for all car models is firmly rejected. Thus, we will trust the results from the fixed effects models more than those from the least square models.

/Table 1 about here./

Let us first consider the "ordinary" vehicle characteristics – weight, horsepower, size and top speed. As pointed out above, these variables are highly co-linear. Thus, it is hard to discern the separate influence on demand by these four variables. However, that is not our objective.

The coefficient δ is supposed to capture the depreciation of automobiles, and should thus have a negative sign. This is also the case, and the coefficient is significantly different from zero at least at a 5 percent level of significance.

The variables of the vector $\mathbf{s}_t$ should be interpreted with some care. Remember that our left hand side variable is (the natural logarithm of) the market share of car i in relation to the "market share" of the outside alternative, i.e. the alternative of not owning a car. The interpretation of the consumer price index for transports and communication is straight forward – if it becomes more expensive to use other means of transport than a car, then it should become less attractive not to own a car. Thus we would expect the coefficient on this variable to be negative. Since $\boldsymbol{\theta}$ enters in equation (20) with a negative sign, the estimate should be positive, which is what we find in our estimations when we take account of autocorrelation among the residuals. However, the estimated coefficients are not significantly different from zero.

The GDP variable really should not be in our equation, since the discrete choice setting together with the assumption that the individual specific factor is independent of the price, implies that we have assumed a constant marginal utility of income.[14] In practice, however, that assumption is not reasonable. Rather, we would assume that the higher is the level of income, the larger would be the number of car owners. Thus, we would expect the GDP variable to enter with a negative sign in the conditional utility function for not owning an automobile. Consequently, it should enter with a positive sign in the demand equation, which, in fact, it does. Concerning the time trend, it is

[14] As Hanemann and Kanninen[19] points out, income, m, and price, p, must enter the utility function in a discrete choice model as $m - p$. Thus, the assumption that the individual specific term, in the conditional indirect utility function, is an additive term, implies that the price must also enter this function linearly, i.e. it implies that there are no income effects.

hard to have any a priori views on what shape it should take. However, the estimated time trend fits rather well with observed data on total car sales.

Let us now turn to the coefficients which are of most interest to us, i.e. the coefficients on taxes, the price, and on the environmental classification variable. As expected, the price of the automobile has a significantly negative effect on the demand for a given automobile model. This coefficient is similar in magnitude in all four regressions, although slightly larger in absolute value in the two fixed effect models.

The road tax coefficient has a negative sign, and is of the same order of magnitude as the price coefficient, which is what we should expect, since they are in fact different estimates of the same parameter. (Recall from equation (20) that they are both estimates of the marginal utility of income.) However, this coefficient is not significantly different from zero at ordinary significance levels in any of the regressions. The same is the case with the "tax credit" coefficient. The estimates of this coefficient shift greatly in size between the regressions and have the "wrong" sign in the two least square regressions. One reason for this result may be that the variation of these variables is too small. In the sample, the road tax varies from around SEK 400 to around SEK 3 700. In percentage terms, this is a considerable variation.[15] However, compared to the total annual cost of owning a car, the road tax is likely to be diminutive.

The gas cost coefficient is significantly different from zero in all the regressions except the fixed effect model on the untransformed data. This implies that it would be possible to make consumers buy less energy consuming cars by increasing the price of gasoline, e.g. by raising the gasoline tax.

What is of particular interest is that the environmental classification variable is negative and significantly different from zero. Taken at face value, this would imply that after account is taken of the reductions in the road tax and/or the registration tax, the demand for a car that has been classified as environmentally "friendly" will be higher. This result thus indicates that consumers care about the environmental characteristics in themselves, of the car they buy. In other words, we have established that in this market eco-labeling seems to work, in the limited sense that it positively affects the demand for eco-labeled products.[16]

4.2 The pricing equation

In the pricing model, as in the demand model, there is strong evidence of autocorrelation – the estimated autocorrelation coefficient is 0.67. Also, we must reject the hypothesis that the fixed effects are zero. Thus, we will mainly trust the results from the autocorrelation adjusted fixed effects model. We also have the same problem of multicollinearity between the vehicle characteristics variables, as we do in the demand equation. Thus,

[15] The sample mean is 1060 and the standard deviation is 506.

[16] To evaluate whether or not the system has in fact improved the environmental characteristics of the car fleet, we would also need to consider if the "right" cars have been included in the better environmental classes. However, that is not primarily an issue for an economist to deal with.

the coefficients on weight, engine power, size and top speed, differ somewhat between the different regressions. However, they are all significant at any common level of significance, and as pointed out before, the values of these coefficients are not of any major importance to us.

/Table 2 about here./

The "value loss factor", i.e. our $d_{i,t}$-variable from the demand equation, does not appear to have a significant effect on the price. The automobiles gas use per 10 kilometers driven is significant, with a negative sign in the preferred model, i.e. the fixed effects model with autocorrelation correction.[17] The interpretation of this parameter would be that if we compare two cars with the same characteristics, with the exception that one uses 0.1 liters of gas less per 10 kilometers, than does the other, then the former would cost about SEK 3 000 more to buy, exclusive of taxes.

In the pricing equation, the markup term is much smaller than we would expect from the estimates of the price coefficient in the demand equation. In two of the regressions, the estimates are not even significantly different from zero. (Recall from the two equations (20) and (24) that these coefficients should be the inverse of eachother.) It is hard to draw any conclusions from this result, other that we would like to refine our model of the firms' behavior.

The only "real" cost shift variable we have included, exchange rate changes, does not show up with a significant coefficient. In the models where we have taken account of the serial correlation among the residuals, the coefficient does at least have the expected sign, i.e. it is negative. The lack of significance of this variable is evidence of limited pass-through of the current exchange rate to prices. This is well in line with what has been found in many other markets for differentiated goods.[17] As in the demand equation, it is hard to have any a priori expectations about the time trend variables.

The variable corresponding to the environmental classification of the car has a significantly negative effect on the price of the car in the two fixed effect regressions. Thus, the price is higher for cars with better environmental characteristics. Taking the values of the parameter at face value, given other characteristics, an MK2 car would cost approximately SEK 3 000 more than an MK3 car and an MK1 car would cost another SEK 3 000 more.

4.3 Effects of the environmental classification system

In this section, we give an illustration of the qualitative effects of the environmental classifications system. Naturally, the figures presented must be taken with a grain of salt. However, it gives us some idea of the importance of this system.

[17] In the least square estimation, the *gasuse* variable is significant with a positive sign. This is probably due to the fact that many expensive cars also use lots of gas, and highlights the dangers of omitting the individual specific effects.

Consider two otherwise identical car models, one of which fulfills the requirements for environmental class 2 and one which does not. To focus on the environmental classification system, we assume that also the fixed effects and the random terms in the pricing equation and in the demand equation are identical.

There are two effects on demand for the respective car models that we need to consider. The demand will tend to be higher, ceteris paribus, for the car in environmental class 2. However, we will not have ceteris paribus, since we would also expect the price of this car model to be higher. To calculate the difference in demand between these two car models, we thus need to use both the demand equation (20) and the price equation (24). We use the coefficient estimates from the preferred model, i.e. the fixed effects model with autocorrelation correction. The difference in demand will also depend on the discount rate ρ, the life-length of the car models T_i and the VAT-rate τ^v, since these variables interact with the price in the demand equation. Using the actual values of these variables, this implies that the car that belongs to environmental class 2 would sell somewhere between 20 and 23 percent more than the otherwise identical car belonging to environmental class 3. The environmental classification system appears to make a difference.

5 Concluding remarks

The results point at some interesting conclusions about eco-labeling, tax incentives and the potential for the "greening" of the vehicle fleet. Somewhat surprisingly, it appears as if it is not the tax incentives but rather the eco-labeling in itself that affects demand. Indeed, the effect of the eco-labeling variable is almost implausibly large, while the effect of the tax incentives is hard to trace at all. We would certainly like to study the matter further before concluding that we do not need tax incentives, but may rely on eco-labeling only to get a "greener" vehicle fleet. However, we may conclude that consumer demand for less environmentally harmful products may indeed be a powerful force.

There may be several reasons why we are not able to show any effects of the tax incentives in the present study. The most obvious reason may be that the incentives have been rather small, if we compare them to the total cost of owning an automobile. Also, as should be appearant from the exposition in Section 3.2 of this paper, there have been frequent changes in the taxation of automobiles. No less than fourteen regulations affecting the road tax and the registration tax have been introduced in Sweden during the 1990s.[27] Admittedly, not all of these affect personal vehicles. Still, when the tax code changes that often, consumers may not trust that future changes will not alter the relative attractiveness of different car models. For this reason, incentive effects of taxes may be weakened.

We have thus not been able to trace any effects on demand from differences in the road tax. In contradistinction, the price of automobile fuel appears to have a significant influence on the consumers' choice between different car models. It thus seems safe to conclude that the average fuel use of automobiles is lower than it would have been, had the taxes on fuel been lower. This conclusion is also in line with what has been found in studies of gasoline demand.[22]

We have concluded that eco-labeling "works" in the sense that it seems to affect consumer demand. However, we also want an eco-labeling system to work in a broader sense. We want it to contribute to a better environment. We then need to consider some other issues. In particular, we would like to know if eco-labeled products really are better for the environment than are products not carrying such a label. Serious doubts have in fact been raised as to the appropriateness of the criteria for some eco-labeling schemes. (See e.g. Solyom and Lindfors[26].) Today, there exists a profusion of different eco-labels with different characteristics. Some are government sponsored, some are sponsored by intergovernmental organizations, such as the Nordic Council of Ministers or the European Union, some have been initiated by NGOs, and some have the form of industry standards. We have shown that eco-labeling does affect consumer demand, in one particular market. However, we need to know a lot more to be able to determine what would make an eco-labeling scheme successful, and on what markets such a scheme may produce desirable results.

References

[1] Anderson, Simon P., André de Palma and Jacques-François Thisse (1992) *Discrete choice theory of product differentiation.* Cambridge, Mass. and London: MIT Press.

[2] Atkinson, Scott E. and Robert Halvorsen (1985) "A new hedonic technique for estimating attribute demand: An application to the demand for automobile fuel efficiency." *The review of economics and statistics.* 66, s. 417-426.

[3] Bernett, P. T. Keeler and T. Hu (1995) "Oligopoly structure and the incidence of cigarette excise taxes." *Journal of Public Economics.* 57, s. 457-470.

[4] Berry, Steven T. (1994) "Estimating discrete-choice models of product differentiation." *RAND Journal of Economics.* 25(2), 242-262.

[5] Berry, Steven T., J. Levinsohn and A. Pakes (1997) "Voluntary export restraints on automobiles: Evaluating a strategic trade policy." Mimeo. Work under progress. Version: March 19, 1997. Cambridge, Mass.: NBER.

[6] *Bilfakta, våren 97.* (1997) Helsingborg: Autograph-Bilfakta AB, no 88, mars 1997.

[7] Brendemoen, Anne (1994) "Car ownership decisions in Norwegian households." Statistics Norway Discussion Papers, No. 116. Oslo: Statistics Norway, Research Department.

[8] Bresnahan, T. (1987) "Competition and collusion in the American automobile oligopoly: the 1955 price war." *Journal of Econometrics.* 17, s 201-227.

[9] Doyle, Maura P. "The effects of interest rates and taxes on new car prices." Washington, D.C.: Federal Reserve Board of Governors.

[10] Espey, Molly (1998) "Gasoline demand revisited: an international meta-analysis of elasticities." Energy Economics. 20:273-295.

[11] Falvey, Rodney E. and Alan Rogers (1986) "Fuel prices, sales taxes and passenger car market shares in New Zealand, 1963-1978." *The Economic Record.* March, s. 52-59

[12] Feenstra, R.C., J.E. Gagnon and M.M. Knetter. (1996) "Market share and exchange rate pass-through in world automobile trade." *Journal of international economics.* 40, s. 187-207.

[13] Fershtman, Chaim, Neil Gandal and Sarit Markovich (1997) "Estimating the effect of tax reform in differentiated product oligopolistic markets." The Foerder Institute for Economic Research and The Sackler Institute for Economic Studies Working Paper, No. 29-97. Tel-Aviv: The Eitan Berglas School of Economics, Tel-Aviv University.

[14] Flam, Harry and Håkan Nordström (1995) "Why do pre-tax car prices differ so much across European countries?" Seminar paper no. 591. Stockholm: Institute for International Economic Studies.

[15] Gagnon, Joseph E. and Michael M. Knetter. (1995) "Markup adjustment and exchange rate fluctuations: Evidence from panel data on automobile exports." *Journal of International Money and Finance.* 14:2, s 289-310.'

[16] Goldberg, Pinelopi K. (1995) "Product differentiation and oligopoly in international markets: the case of the US automobile industry." *Econometrica.* 63:4, s. 891-951.

[17] Goldberg, Pinelopi K. (1996) "Goods prices and exchange rates: What have we learned?" NBER Working paper 5862. Cambridge, Mass.: National Bureau of Economic Research.

[18] Greene, William H. (1993) *Econometric analysis.* 2nd edition. New York: Macmillan Publishing Company.

[19] Hanemann, W. Michael and Barbara Kanninen (1996) "The statistical analysis of discrete-response CV data." Department of Agricultural and Resource Economics, University of California at Berkley, Working paper no 798.

[20] *International Financial Statistics.* (1999) Washington, D.C.: IMF. (CD-ROM)

[21] Kirman, Alan and Nathalie Schueller. (1990) "Price leadership and discrimination in the European car market." *Journal of Industrial Economics.* 39:1, s. 69-91.

[22] Johansson, Olof and Lee Schipper (1997) "Measuring the long-run fuel demand of cars." *Journal of transport economics and policy.* September, 277-292.

[23] Levinsohn, J. "Empirics of taxes on differentiated products: The case of tariffs in the US automobile industry." i R. Baldwin, red. *Trade Policy and Empirical Analysis.* Cambridge, Mass.: NBER.

[24] Massé, Pierre (1962) *Optimal investment decisions.* Englewood Cliffs, N.J.: Prentice-Hall, Inc.

[25] Sandström, Mikael (1999) "The economics of automobile scrappage." *Manuscript.* Stockholm: Stockholm School of Economics.

[26] Solyom, Peter and Lars-Gunnar Lindfors (1998) "Evaluation of Nordic Swan eco-labelling criteria documents for 8 groups of products containing surface active agents." IVL Rapport B1307. Stockholm: Swedish Environmental Research Institute.

[27] SOU 1997:126. *Delbetänkande av trafikbeskattningsutredningen.*

[28] Train, Kenneth (1986) *Qualitative choice analysis: theory, econometrics and an application to automobile demand.* Cambridge, Mass. and London: MIT Press.

[29] *Transport and communication statistics 1995.* (1996) Oslo-Kongsvinger: Statistics Norway.

[30] Verboven, Frank. (1994) "International price discrimination in the European car market." CentER Discussion Paper no. 9451. Tilburg: Center for Economic Research.

[31] Verboven, Frank. (1996) "International price discrimination in the European car market." *RAND Journal of Economics.* 26(2), 240-268.

[32] Wetterwald, Dag G. (1994) "Car ownership decisions in Norwegian households." Statistics Norway Discussion Papers, No. 113. Oslo: Statistics Norway, Research Department.

		Least squares	Fixed effects	Least squares, Cochrane-Orcutt	Fixed effects, Cochrane-Orcutt
Observations		2049	2049	1929	1929
Parameters		15	134	15	134
R^2		0.305	0.768	0.230	0.625
Adj. R^2		0.300	0.752	0.224	0.597
Est.autocorr.			0.465		
RHO				0.465	0.465
$-\lambda$	$\frac{\rho(1+\tau^v)}{1-e^{-\rho T}}(q+\tau^s)$	-4.22E-05**	-5.95E-05***	-4.59E-05***	-5.57E-05**
		(5.72E-06)*	(1.27E-05)	(6.77E-06)	(1.34E-05)*
α_1	*env. class*	-0.540**	-0.348***	-0.443***	-0.242**
		(0.148)*	(0.105)	(0.139)	(0.113)
α_2	*weight*	-1.19E-03**	4.61E-03***	-1.34E-03**	4.17E-03**
		(5.26E-04)	(9.81E-04)	(5.92E-04)	(1.22E-03)*
α_3	*horsepower*	-6.87E-03**	-7.14E-04	-5.79E-03	-3.28E-03
		(3.02E-03)	(4.22E-03)	(3.59E-03)	(5.14E-03)
α_4	*size*	3.01E-03**	2.01E-03***	3.19E-03***	1.81E-03**
		(2.10E-04)*	(7.14E-04)	(2.63E-04)	(8.53E-04)
α_5	*speed*	1.34E-02**	1.07E-02	1.67E-02***	1.24E-02
		(4.60E-03)*	(6.63E-03)	(5.59E-03)	(8.07E-03)
$-\lambda$	τ^r	-6.60E-05	-1.39E-04	-4.46E-05	-1.00E-04
		(1.12E-04)	(8.99E-05)	(1.13E-04)	(1.04E-04)
λ	$(1-e^{-5\rho})\tau^m$	-1.14E-04	1.10E-03	-3.95E-05	8.89E-04
		(1.79E-03)	(1.16E-03)	(1.53E-03)	(1.17E-03)
γ	g	-0.945*	-0.628	-1.14**	-1.07**
		(0.521)	(0.475)	(0.530)	(0.485)
$-\delta$	$\frac{1-(1+\rho T)e^{-\rho T}}{\rho(1-e^{-\rho T})}d$	-0.192**	-0.510***	-0.167**	-0.304**
		(7.27E-02)*	(0.111)	(8.13E-02)	(0.123)
$-\theta_1$	$CPI^{transport}$	-0.918	-1.257	3.80*	2.01
		(3.17)	(1.92)	(2.09)	(1.53)
$-\theta_2$	ΔGDP	9.20E-06*	9.98E-06***	1.28E-05**	9.98E-06**
		(4.82E-06)	(3.14E-06)	(4.97E-06)	(3.69E-06)*
$-\theta_3$	t	-0.122**	-0.199***	-0.165***	-0.241**
		(2.37E-02)*	(1.69E-02)	(3.27E-02)	(2.59E-02)*
$-\theta_4$	t^2	3.93E-03**	5.44E-03***	4.57E-03***	5.71E-03**
		(9.13E-04)*	(6.12E-04)	(1.05E-03)	(8.23E-04)*
	Constant	-28.8**		-36.7***	
		(4.41)*		(1.90)	

Table 1 – The demand equation. The table displays results from four different econometric specifications of the demand equation, least squares, fixed effects, and each of these on the data transformed using the Cochrane-Orcutt method to account for autocorrelation. Standard errors are given within parentheses. One asterisk indicates that the coefficient estimate is significantly different from zero on the 10 percent level, two asterisks that it is significant at the 5 percent level, and three asterisks that it is significant at the 1 percent level.

	Least squares	Fixed effects	Least squares, Cochrane-Orcutt	Fixed effects, Cochrane-Orcutt
Observations	2049	2049	1929	1929
Parameters	12	131	12	131
R^2	0.835	0.985	0.802	0.940
Adj. R^2	0.834	0.984	0.801	0.936
Est.autocorr.		0.672		
RHO			0.672	0.672
β_1 *env. class*	5010	-3190 **	532	-2840 **
	(3340)	(1290)	(2060)	(1280)
β_2 *weight*	106 **	115 **	137 ***	153 ***
	(12.2) *	(12.2) *	(11.6)	(14.8)
β_3 *horsepower*	2040 **	1140 **	1900 ***	1150 ***
	(65.8) *	(53.6) *	(64.0)	(64.1)
β_4 *size*	-46.7 **	64.8 **	-37.6 ***	67.7 ***
	(5.59) *	(9.44) *	(6.12)	(10.9)
β_5 *speed*	-1630 **	-655 **	-1480 ***	-688 ***
	(120) *	(92.7) *	(116)	(105)
β_6 *d*	-14700	28300	-12500	7360
	(18300)	(28600)	(19800)	(32800)
β_7 *gasuse*	42300 **	-3580	-6700	-26600 ***
	(16000) *	(8600)	(13000)	(9220)
ϕ_1 *e*	6970	4030	-3390	-2220
	(16500)	(5270)	(5900)	(3380)
ϕ_2 *t*	2290 **	1580 **	2490 ***	2300 ***
	(576) *	(199) *	(736)	(440)
ϕ_3 t^2	-59.8 **	-39.8 **	-63.2 ***	-61.8 ***
	(21.1) *	(7.06) *	(22.3)	(13.2)
$\frac{1}{\lambda}\ \frac{1-e^{-\rho T}}{\rho\left(1-\frac{y}{M}\right)\left(1+\tau^{v}\right)}$	3850 **	1240 **	955	483
	(914) *	(296) *	(588)	(338)
Constant	411000 **		346000 ***	
	(56300) *		(19100)	

Table 2 – The pricing equation. The table displays results from four different econometric specifications of the pricing equation, least squares, fixed effects, and each of these on the data transformed using the Cochrane-Orcutt method to account for autocorrelation. Standard errors are given within parentheses. One asterisk indicates that the coefficient estimate is significantly different from zero on the 10 percent level, two asterisks that it is significant at the 5 percent level, and three asterisks that it is significant at the 1 percent level.

Essay 4

The economics of automobile scrappage

The economics of automobile scrappage

F. Mikael Sandström*
Stockholm School of Economics

Abstract

In recent years, economic incentives to reduce the life length of automobiles have been proposed as a means to reduce vehicle emissions. The argument is that old cars often pollute more than new ones. "Accelerated scrappage programs" have also been implemented, in some European countries and in the USA. A "chain of replacement" model is used to examine the effects of automobile taxes and of a scrappage premium on the optimal life length of cars, and on the size of the car fleet, and the predictions of the model are tested on data on the scrappage of cars in Sweden 1989-1996.

The theoretical model predicts that increased taxes on the purchase of cars should increase the life length of cars, and reduce the number of cars. A permanent scrappage premium would have the opposite effects. Changes in periodic taxes would have no effect on the life length of automobiles, but would reduce the size of the car stock. A temporary scrappage premium would lead to the scrappage of all cars worth less than the premium. These cars would immediately be replaced by new cars. The econometric analysis supports the conclusions from the model.

It is concluded that both a temporary and a permanent scrappage premium would have ambiguous effects on emissions. Thus, it is questionable if such policies reduce vehicle emissions.

*The work with this study was financed by the Government Committee on Automobile Taxation (trafikbeskattningsutredningen) and the Stockholm School of Economics. I am grateful to Per-Olov Johansson, Richard Friberg and to seminar participants at the Stockholm School of Economics and at CTS, University of Dalarna, for valuable comments.

1 Introduction

Automobiles produced today pollute the environment much less than do automobiles produced only a few years ago. Per distance driven, a gasoline powered automobile from 1996 emits less than half as much as does a 1988 automobile, in respect of all but two of the fourteen most important emission components.[12] In addition, emissions from a given car may increase with its age as pollution reduction equipment wears down, and the general condition of the car deteriorates. New cars are also generally much safer for the driver and the passengers.

Claims of such environmental and security benefits have been used as arguments for policies designed to accelerate the rate at which old automobiles are taken out of use. It should be noted that both the production of new automobiles and the disposal of old cars involve environmental hazards. Thus, it is by no means self-evident that a shortened life length of cars should be a policy goal. However, "early retirement programs", "accelerated scrappage schemes" or "accelerated vehicle retirement programs" have been advocated and in many places implemented. Such programs involve the payment of a fixed sum of money, a scrappage premium, to car owners who decide to scrap their cars under the duration of the program. This premium is usually not paid for all cars but only under certain conditions. Such conditions might include that the car should be above a given age. In many cases, the premium is tied to the purchase of a new car. This has been the case with the accelerated scrappage schemes employed in France, Greece, Italy, Spain and Ireland, but not with the Danish and Norwegian schemes. An additional motive behind schemes where the premium is tied to the purchase of a new car has been to stimulate the automobile industry. In France, this was in fact the main motive. All European accelerated vehicle retirement programs have been government financed, even though the automobile industry has sometimes provided additional incentives to consumers who buy new cars. [3],[14]

In the United States, accelerated vehicle retirement programs (AVRPs) have been promoted by the EPA as a means for state and local government and industry to meet requirements under the Clean Air Act.[7] In Delaware, an AVRP was implemented by the U.S. Generating Company to offset an increase in hydrocarbon emissions caused by transports to a new plant. [2] Under a consent decree, General Motors has agreed to implement and finance AVRPs in Southern California and in Phoenix/Maricopa County in Arizona.[9]

Systems with permanent scrappage premiums have also been used in at least two countries, Sweden and Norway. These premiums were introduced with a different motive, namely to give an incentive to car owners not to place junked cars in parking lots, forests, rivers and lakes, etc. In Sweden, such a system was introduced in 1976, and was similar to a deposit-refund system. When a car was handed over to an authorized scrappage firm, the car owner was paid a bonus. The system was financed by a fee that was paid by buyers of new cars.[6] In 1992, however, the system was altered (the details of the system are described in section 3.1), and the accelerated scrapping of old cars was introduced as a policy goal. In contradistinction to the temporary scrappage premia described in the last paragraph, this system was intended to be in continuous operation.

Thus, both temporary and permanent scrappage premiums have been employed to induce the scrapping of old cars. In addition, we may expect taxes to have an effect on the life length of automobiles. In particular, we may expect the distribution of the total burden of taxation between taxes on the sales of automobiles (e.g. a registration tax) and on the ownership and usage of automobiles (e.g. a road tax and a fuel tax) to affect their average life length.

Some attempts have been made to measure the costs and benefits of such policies, or policy proposals, or to examine the factors determining participation in a scrappage program, by the use of simulation techniques or through econometric analysis. (See e.g. [1], [2] and [8].) Alberini, Harrington and McConnell [2] also present a theoretical model of the decision to participate in a vehicle scrappage program. Another approach to the effects of policy on the life length of vehicles is presented by Innes.[10] He derives an expression for the optimal periodic tax (road tax) for cars of different "vintages" when external costs vary systematically between these, and shows that a scrappage premium that declines with the age of the car may be an alternative to a road tax which increases with the age of the car. Bohm [6] offers a thorough theoretical and empirical analysis of deposit-refund systems.

Unlike previous papers that deal with automobile scrappage, the theoretical part of this paper does not focus on the scrappage decision per se. Rather, we model how the optimal life length of automobiles and the size of the car fleet is determined by the combination of taxation and a scrappage premium. This allows us to analyze a broader range of policy issues, and also to get some insights into the likely long-term effects on emissions from accelerated scrappage programs. The model presented in Section 2 assumes an infinite horizon model with certainty, and interprets the problem as a "chain of replacement" problem. (See e.g. Massé[13].) By allowing individuals to differ in their valuation of car ownership, the consumers' decisions on whether or not to be car-owners can be studied within the model. Comparative statics for changes in taxes and a permanent scrappage premium are derived and the effects of a temporary scappage premium are analyzed. In Section 3, an econometric analysis of the scrappage rate in Sweden is presented. The Swedish scrappage premium system is described and estimation results are presented. Some policy implications of the theoretical and empirical results are then addressed in Section 4. We also discuss what factors should be included in a cost and benefit analysis of an accelerated scrappage program. In the concluding section, the main findings are summarized, and some limitations of the model are considered.

2 A theoretical model

In this section, a theoretical model is presented of how the optimal life length of an automobile is determined. The problem is described as a "chain of replacement" problem. (See e.g. [13].) Unlike in the standard formulation of this problem, we allow for an outside alternative, i.e. not owning a car, and also model the consumer's decision on whether or not to own a car. In the first subsection, the problem is defined, and the first-order condition for the optimal replacement period is derived. We then derive an expression for the number of automobiles in the economy.

In subsections 2.2 and 2.3 we analyze the effects of policy changes. The first of these subsections deals with permanent changes in taxes and the scrappage premium, while the second addresses the issue of what effects we may expect from a temporary scrappage premium, i.e. a premium that is only paid for scrapped automobiles during a brief period of time.

2.1 The chain-of-replacement problem

We will first specify a chain-of-replacement model, and derive the first order condition for the optimal life length of automobiles. We then consider factors that will determine the size of the car park.

Consider an infinitely lived individual j who derives the following instantaneous utility from owning an automobile of age t:

$$u_j(t) = \alpha(j) + s(t) - c(t) - \tau_r \tag{1}$$

where $s(t)$ is the benefit from the car, $c(t)$ is the cost of owning and running it and τ_r is a periodic tax. All of these three components are defined as benefit and cost flows when the car is t years old. They are assumed to be the same for all individuals, while $\alpha(j)$ is a taste parameter, which is constant over the age of the car, but varies over individuals. We will assume that $s(t)$ decreases and that $c(t)$ increases as the car gets older.

Assume further that the (exogenous and constant) price of a new car is:

$$p = q + \tau_s \tag{2}$$

where q is the retail price of the car, and τ_s is a registration tax, or a fixed sales tax, levied on automobiles. When the automobile is scrapped, the owner receives a scrappage premium equal to R.

Assume a simple utility function, which is linear in income. The net present value, to individual j, of the utility from owning a car that is bought when it is 0 years old, and kept until it is T years old, may then be written as:

$$B_j(T) = -(q + \tau_s) + \int_0^T e^{-\rho t}\left[\alpha(j) + s(t) - c(t) - \tau_r\right] dt + e^{-\rho T} R \tag{3}$$

where $\rho \in\,]0, 1[$ is the constant discount rate.

Since the decision problem will be identical every time the car is replaced, we can immediately conclude that the replacement time must always be the same. We also recognize that the total net present value of utility to the consumer from being a car owner will be the same, every time a replacement takes place. Thus, we may write the total present value utility from being a car owner as a function of T, as[13]:

$$\begin{aligned} U_j(T) &= B_j(T) + e^{-\rho T} U_j(T) \qquad (4) \\ &\Rightarrow \\ U_j(T) &= \frac{1}{1-e^{-\rho T}} B_j(T) \qquad (5) \end{aligned}$$

This expression defines a "chain of replacement" problem.

The consumer's maximization problem, assuming an interior solution, will thus be:

$$\max_T U_j(T) = \frac{1}{1-e^{-\rho T}} B_j(T) \qquad (6)$$

yielding the first order condition:

$$\frac{\partial U_j(T)}{\partial T} = -\rho \frac{e^{-\rho T}}{(1-e^{-\rho T})^2} B_j(T) + \\ + \frac{e^{-\rho T}}{1-e^{-\rho T}} \left[\alpha(j) + s(T) - c(T) - \tau_r - \rho R\right] = 0 \qquad (7)$$

Using (5), the first order condition may be rewritten as:

$$\rho \left[U_j(T) + R\right] = \alpha(j) + s(T) - c(T) - \tau_r \qquad (8)$$

Notice that the left hand side of this expression is the opportunity cost of waiting with the replacement of the automobile, and that the right hand side is the instantaneous utility of car ownership at time T. Thus, the optimal replacement time T^* will be set such as to equalize these two.

If we decompose the integral in (3), solve the parts containing the constants τ_r and $\alpha(j)$, and then use (5) to separate out $\left[\alpha(j) - \tau_r\right]/\rho$ from $U_j(T)$, we see that $\alpha(j)$ and τ_r cancel out in the first order condition. The optimal replacement time will thus be independent of $\alpha(j)$, and will only be affected by two of the policy variables, R and τ_s, and not by τ_r. This is not surprising since both $\alpha(j)$ and τ_r are additive in the individual's utility function and are independent of how often the individual buys a new car. As long as the consumer is a car owner, he will receive the benefit stream $\alpha(j)$ and incur the cost flow τ_s. We may thus write $T^* = T(R, \tau_s)$. All car-owners will keep their cars for the same period of time, regardless of their individual specific valuation, $\alpha(j)$.

In subsection (2.2) we will use these results to do comparative statics. However, we first turn to the issue of how the size of the car fleet is determined. To do this, we first want to determine under which conditions a consumer will be a car owner. A rational consumer will own a car if the utility from doing so is larger than the utility from not owning a car. Let us normalize by setting the utility of not owning a car to zero. Thus, consumer j will be a car owner if:

$$U_j\left[T(R, \tau_s)\right] \geq 0 \qquad (9)$$

Now, define the "individual invariant" utility from being a car owner, by:

$$U(T) = \frac{1}{1-e^{-\rho T}}\left[-(q+\tau_s) + \int_0^T e^{-\rho t}\left[s(t) - c(t) - \tau_r\right]dt + e^{-\rho T}R\right] \tag{10}$$

$$\Rightarrow$$

$$U_j(T) = U(T) + \alpha(j)/\rho \tag{11}$$

$U(T)$ is thus the part of utility derived from being a car owner that is common to all individuals. To obtain the second expression, we use the same decomposition of expression (3) as above, solve the part containing $\alpha(j)$, and are then able to separate out $\left(1-e^{-\rho T}\right)\cdot\alpha(j)/\rho$ from the resulting expression. Note that the second right hand term of expression (11) is equal to the net present value of an eternal benefit stream with a value equal to the "taste factor" in the utility function.

Using (9) and (11) we can write the condition under which individual j will be a car owner as:

$$\alpha(j) \geq -\rho U(T^*) \tag{12}$$

To go from this expression to an expression for the size of the car park, we need to make some assumptions about the distribution of the taste parameter, $\alpha(j)$. First, assume that the economy is inhabited by a continuum of individuals, and normalize their number to one. Assume further that $\alpha(j)$ is distributed over these individuals according to an integrable density function, $\psi[\alpha(j)]$, such that $\int_a^b \psi[\alpha(j)]\,d\alpha(j)$ gives the proportion of individuals with a value of $\alpha(j)$ between a and b. The number of car owners, as a function of the policy variables, can then be written:

$$n(\tau_s, \tau_r, R) = \int_A^\infty \psi[\alpha(j)]\,d\alpha(j) \text{ where } A = -\rho U(T^*) \tag{13a}$$

$$\Rightarrow$$

$$n(\tau_s, \tau_r, R) = 1 - \Psi[-\rho U(T^*)] \tag{13b}$$

where $\Psi[\alpha(j)] = \int_{-\infty}^{\alpha(j)} \psi[k]\,dk$ is the cumulative distribution of $\alpha(j)$. The expression in (13b) can then be used to analyze how policy changes will affect the size of the car fleet.

2.2 Effects of policy changes

We will now examine the effects of changes in the three policy variables, τ_s, τ_r and R, on the life length of automobiles, and on the number of cars in the economy. We derive comparative statics for the optimal life length of automobiles, and then turn to the size of the vehicle fleet.

It turns out that the road tax, τ_r, has no effect on the optimal T, and that the scrappage premium, R, is exactly equivalent to a negative registration tax, τ_s, with regard to its effect on the optimal life length of automobiles. To show this, we first recognize that since the first order condition, (7), will hold identically in optimum, we may totally differentiate it to obtain:

$$[s'(T^*) - c'(T^*)]\,dT + \rho\frac{1}{1-e^{-\rho T^*}}d\tau_s + 0\cdot d\tau_r - \rho\frac{1}{1-e^{-\rho T^*}}dR = 0 \tag{14}$$

where we have divided through by $e^{-\rho T^*}/\left(1-e^{-\rho T^*}\right)$. To obtain the first term, we have used the first order condition.

Obviously, we will have:

$$\frac{\partial T^*}{\partial \tau_r} = 0 \tag{15}$$

Thus, changes in the road tax will have no effect on the life length of automobiles. This is quite natural, since this tax is assumed to be independent of the age of the car. Thus, the consumer will need to pay the same amount regardless of how often he buys a new car.

The change of the optimal life length when the scrappage premium or the registration tax changes can be written:

$$\frac{\partial T^*}{\partial R} = -\frac{\partial T^*}{\partial \tau_s} = \rho\frac{1}{1-e^{-\rho T^*}}\frac{1}{s'(T^*) - c'(T^*)} < 0 \tag{16}$$

The fact that the expression will be negative follows from the assumptions that $s'(T^*) < 0$ and $c'(T^*) > 0$, i.e. that the benefit from owning the car declines as it gets older, while the cost increases. An increase in the permanent scrappage premium will decrease the life length of automobiles by exactly the same amount as an equally sized reduction of the registration tax. It should be noted that the effects are indeed identical - the scrappage premium is discounted by the same amount as the registration tax. The reason for this is that *given* that the consumer is a car owner, the scrappage premium will always be received at the same time as the registration tax is paid, since the consumer will always buy a new car when the old one is scrapped. Thus, with regard to the life length of automobiles, a permanent scrappage premium is exactly equivalent to a negative registration tax. As we shall see below, the same will not be true when we consider the size of the car fleet.

While the road tax has no effect on the optimal scrappage time for an automobile, it will influence the number of cars. The scrappage premium will no longer be equivalent to a negative registration tax when we consider the size of the car fleet. To establish these claims, we first recall that the number of cars in the economy will be determined by expression (13b). Differentiating this expression, and noting that in an optimum we will have $\partial U/\partial T = 0$, we get:

$$\frac{\partial n(\tau_s, \tau_r, R)}{\partial \tau_s} = -\psi\left[-\rho U(T^*)\right] \cdot \rho \frac{1}{1 - e^{-\rho T^*}} < 0 \tag{17}$$

$$\frac{\partial n(\tau_s, \tau_r, R)}{\partial R} = \psi\left[-\rho U(T^*)\right] \cdot \rho \frac{1}{1 - e^{-\rho T^*}} e^{-\rho T^*} > 0 \tag{18}$$

$$\frac{\partial n(\tau_s, \tau_r, R)}{\partial \tau_r} = -\psi\left[-\rho U(T^*)\right] < 0 \tag{19}$$

Thus, while an increase in the scrappage premium will have exactly the same effect on the life length of automobiles as an equally sized decrease of the registration tax, the latter will have a larger (positive) effect on the size of the car fleet. The reason for this is that we now need to look at the total net present value of the utility from being a car owner, and thus need to discount the registration tax and the scrappage premium differently. When we consider the optimal life length of the car, on the other hand, the effects are identical, since given that the consumer is a car owner, the registration tax will be paid and the scrappage premium received at the same time. Further, while the road tax has no effect on the life length of automobiles, it will affect the number of cars. Naturally, the derivatives will also depend on the distribution of individual valuations of car ownership. The effects of any policy change on the size of the car park will be larger, the larger is the number of people who are close to being indifferent between being car owners and not being car owners, i.e. the denser is the distribution ψ around $-\rho U(T^*)$.

2.3 A temporary scrapping premium

This far, we have only considered a permanent scrappage premium. We will now turn to the issue of temporary incentives for accelerated scrappage of automobiles, i.e. a temporary scrappage premium. Under the assumptions of the model, all cars with a market value lower than the temporary scrappage premium will immediately be scrapped. Thus, such a premium will lead to the accelerated scrappage of the oldest cars in the car park, since these will have the lowest market value. Since the premium is, by construction, a temporary measure, it can have no effect on the size of the car park. Thus, all scrapped cars will instantly be replaced by new cars. Similarly, the premium will not have any effect on the optimal life length of automobiles. Therefore, after the temporary scrappage premium has been abolished, and until the oldest cars that were not scrapped under the accelerated scrappage program reach their optimal life length, no cars will be scrapped. Also, unless the premium is introduced without warning, some car owners will delay scrapping their vehicles.

To establish these claims, we first want to determine what the market price will be for an automobile of a given age. Assume for simplicity that the permanent scrappage premium is zero, and consider the present value utility of a consumer, j, who owns a car of age $t = \bar{t}$. This utility can be written:

$$V_j(\bar{t}) = \int_{\bar{t}}^{T^*} e^{-\rho(t-\bar{t})} \left[\alpha(j) + s(t) - c(t) - \tau_r\right] dt + e^{-\rho(T^*-\bar{t})} U_j\left(T^*\right) \tag{20}$$

The consumer's reservation price will be equal to the smallest sum of money for which he would be willing to give up the car. Since the benefit from being a car owner will not change, the consumer would immediately purchase a new car after selling the old one. The lowest price at which this consumer would be willing to give up his car of age $\bar{t}$ will thus be a price $\bar{p}\left(\bar{t}\right)$, such that:

$$\bar{p}\left(\bar{t}\right) + U_j\left(T^*\right) = V_j\left(\bar{t}\right) \tag{21}$$

$$\Rightarrow$$

$$\bar{p}\left(\bar{t}\right) = \int_{\bar{t}}^{T^*} e^{-\rho(t-\bar{t})} \left[s(t) - c(t) - \tau_r\right] dt - \left(1 - e^{-\rho(T^*-\bar{t})}\right) U\left(T^*\right) \tag{22}$$

The latter expression is obtained by, once more, decomposing the integrals in $U_j\left(T^*\right)$ and $V_j\left(\bar{t}\right)$ and solving the parts involving the constant $\alpha\left(j\right)$.

Notice that this price will be independent of individual characteristics. Thus, all car owners will have the same reservation price, given the age of the car. This price will thus constitute the market price of an automobile of age $\bar{t}$. (Since all consumers have the same valuation, there will however be no trade in used cars.) The first term is equal to the net present value of the "individual invariant" utility from the car during the remainder of its life, while the second term reflects the distance in time until the car will be replaced. It is easy to verify that $\bar{p}\left(0\right) = q + \tau_s$ and that $\bar{p}\left(T^*\right) = 0$.

Suppose, then, that a temporary scrappage premium equal to $\overline{R}$ is paid to all car owners who decide to scrap their cars. Obviously, all cars with a market price lower than this amount will be scrapped. Car owners who scrap their cars will immediately buy new cars, since by assumption, the premium is only temporary and thus does not affect the inequality in (9), which decides whether or not an individual consumer will be a car owner. The size of the car stock will be unaltered. By the same reasoning, so will the optimal life length of automobiles.

Since the market price of an automobile, (22), will be strictly decreasing in the age of the vehicle,[1] we may use the inverse of the price function to define:

$$\bar{t}\left(p\right) = p^{-1}\left(\bar{t}\right) \tag{23}$$

The temporary scrappage premium will thus lead to the scrappage of all cars older than $\bar{t}\left(\overline{R}\right)$. It should be noted that all these cars would have been scrapped anyway when they reached age T^*. Thus the scrappage premium will reduce the age of the scrapped vehicles by between 0 years, for cars that would have been scrapped anyway, at the time

[1]This reader may verify this by taking the first derivative of expression (22), inserting the first order condition, (8), and then using the fact that, by assumption, $s'\left(t\right) - c'\left(t\right) < 0$.

the program is introduced, and $T^* - \bar{t}\left(\overline{R}\right)$ years. It should also be noted that in this model, no new cars would be bought during a time period after the scrappage program of $T^* - \bar{t}\left(\overline{R}\right)$ years, since all cars that would have been scrapped during that period have already been replaced by newer automobiles. Obviously, the number of automobiles that are scrapped will depend on the age distribution of the car stock.

The reasoning in this section is similar to that of Alberini, Harrington and McConnell.[2] However, they fail to address one important issue, which is that a scrappage program usually cannot be introduced as a complete surprise. Rather, it will be known to automobile owners well in advance. Thus, they may either wait with buying a new car, or buy a new car but refrain from scrapping their old car until they may receive the scrappage premium. In fact, during the latter half of 1975, before the Swedish scrappage premium was first introduced, the scrapping of cars declined, thus indicating that some car owners waited to scrap their cars until they could receive the premium.[6] A similar effect was seen in Denmark. In 1993, the year before Denmark implemented an accelerated vehicle scrappage program, the number of scrapped cars declined dramatically. In 1993, only 8 000 cars were scrapped, which may be compared with the average over the years 1985-1995 which was 93 000 cars.[2]

To illustrate why some car owners may delay scrapping their automobiles, we first note that if there is no cost involved in keeping an old car, then no cars will be scrapped between the announcement of the premium and its implementation. People who would normally have scrapped their cars would just buy a new car and keep the old one until they could receive the premium. Usually, however, some cost would be incurred to store the car and to keep it in a sufficiently good condition for it to be eligible for the scrappage program. Assume that this cost is a constant periodic cost, m. In addition, scrappage programs usually require that the road tax should have been paid for the car. Thus, the cost of keeping the automobile for a period r, after it has reached age T^*, will be:

$$\int_{T^*}^{T^*+r} e^{-\rho[t-T^*]} \left(m+\tau_r\right) dt = \frac{1-e^{-\rho r}}{\rho}\left(m+\tau_r\right) \tag{24}$$

If this amount is smaller than the present value of the scrappage premium, then it will pay to wait. Thus, if it is known that an accelerated vehicle retirement program will be implemented at a given date in the future, a car owner who has a car that reaches its optimal life length r years before that date will delay scrapping it if:

$$e^{-\rho r}\overline{R} \geq \frac{1-e^{-\rho r}}{\rho}\left(m+\tau_r\right) \tag{25}$$

$$\Rightarrow$$

$$r \leq r^*\left(\overline{R}, m+\tau_r\right) = -\frac{1}{\rho}\ln\left(\frac{m+\tau_r}{\rho\overline{R}+m+\tau_r}\right) \tag{26}$$

[2]The number of scrapped cars is calculated as the difference between the number of registered cars, and the change in the size of the car stock during a given year. Data were obtained from the 1989, 1995 and 1996 editions of *World Road Statistics*[15].

where $r^*\left(\overline{R}, m + \tau_r\right)$ is thus the longest time any car owner will delay the scrapping of an automobile in order to obtain the scrappage premium.

Naturally, the number of car owners who are going to wait to scrap their car will be larger, the larger is r^*. It is easy to verify that the number of delayed scrappings will be increasing in the scrappage premium, $\overline{R}$, and decreasing in the "storage cost", m, and in the road tax, τ_r. If the owner derives some benefits from owning the car, even after it is older than T^*, then this will also increase the number of delayed scrappings. The total number of delayed scrappings will, naturally, also depend on the age distribution of the car fleet.

A temporary scrappage premium will thus have two effects. Firstly, some automobiles will be scrapped earlier then they would have been in the absence of the premium, and secondly, some scrappings will be delayed by car owners who would anyway have scrapped their cars in the period between the announcement of the scrappage program and its implementation. Since both of these effects depend on the age distribution of the car stock, we cannot say which effect will dominate. A temporary scrappage premium will have no effect on the optimal life length of automobiles, or on the size of the car stock.

3 An empirical illustration

To empirically test the implications of the model, a simple econometric analysis was performed on Swedish data of the scrappage rate, i.e. the ratio between the number of scrapped cars and the total number of cars. A main goal was to see if changes in the cost of buying and of owning and running a car, respectively, and changes in the scrappage premium had the predicted effects on the scrappage rate. The data used were monthly data from the years 1989 through 1996. In the first subsection, the Swedish scrappage premium system is described. Then, in subsection 3.2, the data and the econometric model are discussed. In the last subsection, the estimation results are presented and analyzed.

What we are really interested in is how the life length of automobiles and the size of the car stock changes when policy variables change. These two variables may be difficult to study, as in real life we may expect them to adjust slowly. However, the transition from one equilibrium to another must affect the scrappage rate. Thus, the scrappage rate can be viewed as a "shortcut" to the study of these two variables.

3.1 The Swedish scrappage premium

The original purpose of the scrappage premium, when it was introduced in Sweden in 1976, was to give car owners an incentive to dispose of scrapped automobiles properly. Rusting hulks in woods, in parking lots, etc. were perceived as a serious problem. The system appears to have achieved this policy objective reasonably well in the beginning of its operation, but seems to have been less of a success later on. (See Bohm [6].)

Originally, the premium was SEK 300, but this amount was increased to SEK 500 in 1988. From 1992, a higher premium, SEK 1 500 was paid for cars that had passed a vehicle inspection within 14 months before they were scrapped. The argument for introducing this higher premium was to reduce pollution by accelerating the scrappage of old automobiles.[4] In 1994, the condition for receiving the higher premium was made more stringent, by requiring that the car should have passed a vehicle inspection within 9 months before scrappage.[5] In 1998 the higher scrappage premium was abolished.

Even though the nominal level of the scrappage premium has only been changed a few times, the real value has varied considerably, due to inflation. In 1997 SEK, the scrappage premium in January 1980 was around SEK 800.[3] When it was raised in April 1988, the real value had fallen to around SEK 440. After it was raised, the premium was SEK 725 in 1997 SEK. Thus the increase in the premium was not enough to fully compensate for inflation since the beginning of the decade.

When the scrappage premium was differentiated, in January 1994, the real value of the lower premium had fallen to SEK 550, still in 1997 SEK. The real value of the higher premium was then SEK 1 660. The lowest value since then, during the period studied, was reached in April 1996 when the real value of the higher scrappage premium, i.e. for cars that had passed vehicle inspection within 14 months, was SEK 1 480, and the lower premium, for other cars, was SEK 493. This variation makes it possible to study the effects of the scrappage premium on the scrappage rate.

3.2 The data and the model

The theoretical model presented in Section 2 predicts that the scrappage premium should have a negative effect on the life length of automobiles, and a positive effect on the size of the car stock. The registration tax should have the opposite effect. The road tax should have no effect on the life length of automobiles, but it should tend to reduce the number of automobiles.

As pointed out above, both the life length of cars and the size of the car stock may change slowly. However, such changes must work through the scrappage frequency. Thus the dependent variable is the ratio between the number of cars that are scrapped and the total car stock. The data used were obtained from Statistics Sweden. The regression was run on monthly data. Since the scrappage premium will increase the turn-over of the car stock, we would expect it to increase the scrappage frequency. The registration tax should have the opposite effect. Since the road tax should have a negative effect on the size of the car stock, it can be expected to increase the scrappage frequency.

Both the registration tax and the road tax vary between different car models. Also, factors other than taxes which change the cost of owning and running an automobile, or the price of buying a car, may affect the life length of cars and the size of the car stock. Thus, instead of using the taxes per se, indices of the cost of buying and of owning and

[3]USD 1 $\approx$ SEK 8, as of February 1999. The Swedish inflation rate has been virtually zero since 1997.

running cars were constructed from the consumer price indices for these two groups of goods. The indices were obtained from Statistics Sweden.

Three different variables had to be used for the scrappage premium: one for the "base" premium, paid for any properly scrapped automobile, one for the scrappage premium paid to automobiles that had passed vehicle inspection within 9 months before scrappage, and one for the premium for cars that had passed inspection within 14 months before scrappage. These data were obtained from "Bilismen i Sverige 1996".[5]

It is reasonable to allow for leads and lags of the scrappage premium. Firstly, we may expect the anticipation of a change in the scrappage premium to induce automobile owners to scrap their cars earlier (when the premium falls) or later (when the premium rises). Secondly, we may expect changes in the premium to have effects on scrappage with a lag, partly as a result of inexactitudes in the data - there may be a lag in reporting. Also, even though the model would predict that reactions to changes in the scrappage premium should be instantaneous, we may also expect a lag due to the consumers' behavior. The theory predicts that an increase in the scrappage premium would lead to the instantaneous scrappage of all cars with a market value lower than the premium. A decrease in the premium would eliminate all scrappage until the oldest automobiles have reached the new optimal life length. However, in practice, we would expect adjustment to be gradual. One factor that is not dealt with in the model, which may induce lags in the adjustment process, is that it takes some time to buy a new car. Also, in real life some people buy used cars, so we would expect the adjustment to an increase in the scrappage premium to involve a "chain reaction". People at the bottom of the "food chain" would scrap their cars, and buy another used car from someone who would in turn buy a slightly less used car, and somewhere at the end of this chain, someone would buy a new car. This would further slow down the adjustment process.

The model is specified to be linear:

$$SCRAP_k = \beta_0 + \beta_1 PBUY_k + \beta_2 PUSE_k + \\ + \sum_{i=1}^{3} \sum_{j=-L_1}^{L_2} \gamma_{i,j} PREM_{i,k+j} + \varepsilon_k \quad (27)$$

where $SCRAP_k$ is the ratio between the number of scrapped cars and the number of automobiles in use during month k. $PBUY_k$ and $PUSE_k$ are, respectively, the price indices for the costs of buying and of owning and running a car. L_1 is the number of lags in the model, and L_2 is the number of leads, and i is used to index the three different scrappage premiums - the premium for cars that have passed vehicle inspection within 9 months before scrapping ($i = 3$), for cars that have passed vehicle inspection within 14 months before scrapping ($i = 2$), and for other cars ($i = 1$). $PREM_{i,k+j}$ is the scrappage premium of type i during month $k + j$, in 1997 SEK. The β:s and the γ:s, finally, are parameters to be estimated, and ε_k is an error term.

3.3 Estimation results

The model was estimated using ordinary least squares. Parameter estimates are presented in Table 1. Different numbers of lags and leads were included, and specifications with less than three leads and lags were rejected (F-tests) at least at the 1 percent level of significance. However, larger numbers of leads and lags did not significantly improve the explanatory power. The adjusted R^2 of the preferred model specification is 0.90. Breusch's and Godfrey's Lagrange multiplier test was used to test for autocorrelation. (See e.g. Greene [11].) This is a test of the null hypothesis of no autocorrelation among the residuals against the composite alternative hypothesis that the residuals follow a moving average or an autoregressive form of order P or less. The test was run setting $P = 5$, to allow for two more periods of autocorrelation beyond the lags included in the model. The null hypothesis of no autocorrelation could not be rejected.[4]

The estimation results support the predictions from the theoretical model. An increase in the cost of buying an automobile decreases the scrappage rate, while an increase in the cost of owning and using cars increases the scrappage rate. Both these effects are significantly different from zero at least at the 5 percent level of significance. (1 percent for the $PBUY_k$-variable.) The joint significance of the parameters of each of the scappage premium variables, including all leads and lags, as well as the joint significance of pairs of these groups of parameters, were tested using F-tests. Specifications excluding one or two of the scrappage premium variables (with all leads and lags) were rejected.

/ Table one about here. /

The scrappage rate is thus positively related to the cost of owning a car. An increase in the scrappage rate may have two explanations. Either it is due to a reduction of the life length of automobiles, or it is due to a reduction in the size of the automobile park. There is, however, no theoretical reason to believe that the cost of owning and running cars should have any major impact on the life length of cars. Thus, our results support the conclusion that if it becomes more expensive to own and use an automobile, e.g. because of a higher road tax, fewer people will own cars.

The cost of buying a car is, on the contrary, negatively related to the scrappage rate. This is perhaps not as obvious, since an increase in the price of cars should also reduce the size of the car stock. However, this reduction can only be accomplished through a reduction in the sale of new cars, since the increased cost does not affect the car owner until he exchanges his old car for a new one. Since buying a new car is more expensive, he will be induced to wait longer. This will cause the scrappage rate to decline. Thus, while we may expect a higher registration tax to reduce the size of the car park, we may also expect car owners to keep their cars longer.

The scrappage premium has a positive effect on the scrappage rate. Since there is no reason to believe that an increase in the scrappage premium, without an accompanying

[4]The χ^2-statistic produced was 9.201, yielding a probability value from the χ^2-distribution with 5 d.f. of 0.899.

increase in some tax or fee, would decrease the size of the car stock, it is reasonable to draw the conclusion that it will in fact reduce the life length of automobiles. The fact that only three lags are needed to achieve a good fit indicates that the adjustment to changes is rather rapid. This is what we would expect since, if the scrappage premium is raised, and your automobile is worth less than the premium, then there is no point in waiting to scrap it later. The results also indicate that the stringency of the conditions under which the scrappage premium is paid does matter, since the exclusion of any of the three different groups of scrappage premium variables is rejected.

4 Policy implications

The stated objective of most accelerated vehicle retirement programs, and also of the Swedish permanent scrappage premium, is to reduce vehicle emissions. Thus, we would like to translate the findings in the previous sections into net changes in such emissions. This is no simple task. However, we may draw some conclusions, at least with regard to which factors need to be taken into account. It turns out that both in the case of a temporary scrappage premium and in the case of a permanent premium, the effects on emissions will be ambiguous. In the last paragraphs of this section, we will consider which factors would need to enter a social cost and benefit analysis of an accelerated scrappage program.

Let us first consider a permanent scrappage premium. A higher scrappage premium will reduce the optimal life length of automobiles, and this will tend to reduce emissions. However, the size of the car fleet will be increased, and this will tend to increase emissions. It is not a foregone conclusion which effect will dominate. Rather, whether the net effect will be positive or negative will depend on how the emission characteristics vary over the life of the automobile, and on the pace of technological progress.[5] Thus, if technological development is rapid, or if emission properties deteriorate rapidly towards the end of a car's life, then it may be justifiable to have policies designed to reduce the life length of automobiles. Instead of using a scrappage premium, however, we may adjust the ratio between the registration tax and the road tax to influence the life length of automobiles. Given the total burden of taxation on cars, if a larger share of the tax revenues is obtained through a registration tax, the car owners will wait longer until they buy a new car. Another policy alternative may be to differentiate the road tax according to the age of the car. Naturally, other measures than taxes and subsidies may also be used to purge the car fleet of highly polluting old automobiles. Such measures might include compulsory vehicle inspection, or more stringent technical requirements for automobiles.

It should be noted that the larger number of cars resulting from the higher scrappage premium could be counterbalanced by increasing the road tax, τ_r, thus resulting

[5] Naturally, it will also depend on the functions determining the optimal life length, (8), and the size of the car park, (13b).

in unambiguously lower pollution in the economy with a higher scrappage premium. Obviously, the "deposit-refund motive", i.e. the use of a scrappage premium to avoid uncontrolled scrappage of cars in forests and lakes, etc., is a separate issue. This issue is thoroughly analyzed by Bohm[6], and will not be addressed in this paper.

Let us now discuss an "accelerated scrappage program" under which a temporary scrappage premium is paid, and assume for the moment that it is introduced completely unexpectedly. Thus, we need not consider delayed scrappings. Our findings suggest that the oldest cars in the car park would then be scrapped as soon as, or at least shortly after, the temporary scrappage premium is introduced. These cars would then instantly, or almost instantly, be replaced by new cars. The reduction in emissions will thus be the difference between the emissions from the scrapped cars and from the new cars. Note, however, that if the emission properties of the automobiles deteriorate with the age of the cars, but are constant over time for cars of the same age, then total emissions over a car's life span will be the same. Thus, the size of the benefits would depend on how future emissions are discounted. If, on the other hand, emission properties of automobiles are improved over time, also given the car's age, then the emissions reduction will be permanent. However, after the scrappage premium has been abolished, and until the oldest cars that were not scrapped under the accelerated scrappage program reach their optimal life length, and are scrapped, no new cars will be sold, and no old cars scrapped. (Remember that the optimal life length of automobiles as well as the size of the car stock are unaffected by a temporary scrappage premium.) This will tend to counteract the initial reduction in emissions, as automobile emissions will increase, or at least not decrease during this period. Had the temporary scrappage premium not been introduced, on the other hand, automobile emissions would have been reduced during this period due to technological improvements in new cars. The net effect on emissions seen over a longer period of time will thus be uncertain, and will depend not only on the pace of technological development, but also on whether the pace of technological development is relatively constant, or if it takes place by leaps. If technology develops by leaps, then the gain from getting rid of old polluting cars, after such a leap has taken place, may be large. If no major changes in technology take place in the period shortly after the accelerated scrappage program is abolished, then the counteracting increase in pollution may be small. The introduction of catalythic converters may be an example of such a "leap" in technology. Naturally, the same reasoning can be applied to technological advances that improve the safety of automobiles.

As we have discussed above, it is usually not realistic to think that a scrappage premium can be introduced unexpectedly. Suppose instead that the temporary scrappage premium is not introduced unexpectedly, but more realistically, that it is anticipated some time ahead. We then need to consider car owners who delay scrapping their cars. If these cars are just kept in a parking lot, and not used, they will have no effect on emissions. If, however, they are used, perhaps because the automobile owner waits to buy a new car until the old one is scrapped, or as a second car, then this will have an off-setting effect on any emissions reduction. The number of delayed scrappings may perhaps be reduced by introducing more stringent requirements for an automobile to be eligible to receive the premium. However, that may induce car owners to take costly action to make the car eligible, e.g. by repairing it, thus increasing the social cost of delayed scrappings. From the inequality in (26) we know, however, that the number of delayed scrappings will decrease if the road tax is increased. Thus, if the road tax is increased when the scrappage premium becomes known, or some time before it is implemented, then this problem may be mitigated.

Making a social cost and benefit analysis of an accelerated vehicle scrappage program or of a permanent scrappage premium must obviously involve an assessment of the net effect on emissions. Such an assessment would require assumptions about how technological development will affect specific emissions, as well as of how the age composition of the vehicle stock would evolve with and without a scrappage premium. When a scrappage premium is introduced, we need to take account of delayed scrappings. Also, the cost incurred by a car owner (excluding taxes) of keeping the car until it is scrapped must be included in a cost and benefit analysis. In addition, the payment of a scrappage premium to car owners who would anyway have scrapped their cars will constitute a redistribution of resources, which may be unwarranted. If the accelerated scrappage program is publicly funded, and if the social cost of public funds is larger than unity, this redistribution will also give rise to a net social cost.

Finally, it must once more be pointed out that the additional environmental costs that arise due to the increased number of scrapped cars, and increased automobile production, must be included in the analysis. It is not certain that even a "successful" accelerated vehicle retirement program will improve the environment.

5 Concluding remarks

Any model involves simplifications. However, I believe that the qualitative results of the model would be robust to alternative specifications, with two possible exceptions. Firstly, we have assumed that the price of automobiles is constant and exogenous, i.e. that supply is infinitely elastic. Secondly, the market for scrapped cars is not included in the analysis at all. Let us first consider the supply of automobiles. The model in this paper predicts that a temporary scrappage premium would lead to a large temporary increase in the demand for new cars, followed by a sharp reduction in demand, before normal conditions return. Such swings in demand may affect prices, if demand is not perfectly elastic, and also industry profits, if some form of imperfect competition prevails. Whether profits will increase or decrease depends on market conditions. However, the fact that the automobile industry in several countries has lobbied in favor of introducing "accelerated vehicle scrappage programs" may indicate that the industry's representatives, at least, expect the effect to be positive.[6]

[6] It is easy to show that if the price is unchanged, and the profit margin is not affected, then industry profits will increase in the model used in this paper. Since the automobile stock will be unchanged, we

Interaction with the market for scrapped cars may also affect our results. Many parts of a discarded automobile have a market value. However, a sharp increase in the number of automobiles scrapped is likely to affect the price of used parts and of scrap metal. This is in fact one reason why the initially successful Swedish deposit-refund system failed to achieve its objective after a few years of operation. The price of scrap metal fell from SEK 130 per ton in 1975 to SEK -20 per ton in 1978, i.e. the scrappage firms had to pay to get rid of scrap metal.[6] Thus, effects on this market are likely to have important consequences for the success of a scrappage program. An exhaustive investigation into the effects of an accelerated vehicle retirement program would thus need to consider the supply of automobiles and the demand for scrapped cars.

Is a scrappage premium, temporary or permanent, a good way of reducing pollution then? Firstly, it should be concluded that, if initially the size of the car stock is socially optimal (or at least not below its optimal level), an increase in the permanent scrappage premium should be accompanied by an increase in the road tax. Otherwise, it will increase the number of cars in the economy. Given that taxes are adjusted so as to keep the car stock constant, the net effect on the environmental harm from automobiles will depend on how much less each car will emit, on average, over its life, how fast technological development improves the emission characteristics of automobiles, and finally on how much environmental harm is caused by the increase in production and scrappage of automobiles. As Innes [10] points out, a well designed scrappage premium may have its place in a second best setting. Since old cars pollute more than new ones, it may be desirable to tax them more heavily. Innes claims that a scrappage premium which declines with the age of the car may be an alternative to a road tax that increases with the age of the car.

It is not self evident that a temporary scrappage premium does in the long run reduce emissions. Even disregarding adverse environmental effects due to more rapid scrappage and larger production of cars, we cannot be sure that emissions will be reduced. The immediate effect, when old cars are scrapped, and new ones put in their place, is likely to be a reduction of emissions. However, the cars scrapped are most likely those that would anyway soon have been scrapped. Thus, it is questionable if the gain will be that large. In addition, some time after an accelerated vehicle retirement program, we may expect the car stock to be on average *older* than it would have been, had the program not been implemented. The reason for this is that the cars that are scrapped during the program are replaced sooner than they would otherwise have been. When the time is reached when those cars would have been replaced in the absence of the program, the vehicle stock will thus be older. Depending on how technological development affects automobile emissions, an accelerated vehicle retirement program may thus well cause emissions to *rise*.

will in effect move some car sales closer in time by introducing a temporary scrappage premium, thus increasing the net present value of profits.

References

[1] Alberini, Anna, Winston Harrington and Virginia McConnell (1996) "Estimation of an emissions supply function from accelerated vehicle retirement programs." *Review of Economics and Statistics.* 78:2:251-65.

[2] Alberini, Anna, Winston Harrington and Virginia McConnell (1995) "Determinants of participation in accelerated vehicle-retirement programs." *RAND Journal of Economics*, 26:1:93-112.

[3] Berglund, Carl Magnus and Pontus Matstoms (1997) "Effekter av ekonomiska styrmedel på bilparkens ålderssammansättning." (mimeo.) Linköping: Statens väg- och transportforskningsinstitut.

[4] "Bilen, miljön och säkerheten." (1997) SOU 1997:126. Delbetänkande av trafikbeskattningsutredningen. Stockholm: Finansdepartementet.

[5] "Bilismen i Sverige 1996." (1996) Stockholm: Bilindustriföreningen - AB Bilstatistik.

[6] Bohm, Peter (1981) *Deposit-refund systems.* Baltimore and London: John Hopkins University Press for Resources for the Future.

[7] "Guidance for the implementation of accelerated retirement of vehicles programs." (1993) Environmental Protection Agency, Office of Mobile Sources, February 1993.

[8] Hahn, Robert W. (1995) "An economic analysis of scrappage." *RAND Journal of Economics.* 26:2:222-242.

[9] Herman, Steven A. (1996) "Landmark GM settlement protects public, promotes innovative pollution prevention projects." Environmental Protection Agency, Office of Enforcement and Compliance Assistance. (December 1996 - January 1997) <http://es.epa.gov/oeca/naag/december.html> [Accessed 23 April 1999].

[10] Innes, Robert. "Regulating automobile pollution under certainty, competition, and imperfect information." *Journal of Environmental Economics and Management.* 31:219-239.

[11] Judge, George G. et al. (1985) *The theory and practice of econometrics.* 2nd ed. New York: John Wiley and Sons.

[12] Kommunikationsforskningsberedningen (1997) (Swedish Transport and Communications Research Board) Meddelande 97:22.

[13] Massé, Pierre (1962) *Optimal investment decisions.* Englewood Cliffs, N.J.: Prentice-Hall, Inc.

[14] "New cars for old." (1995) London, UK: Society of Motor Manufacturers and Traders.

[15] *World road statistics.* (1989, 1995 and 1996) Geneva: International Road Federation.

Coefficient	Variable	Estimate	St. dev.	t-value
β_0	Constant	-2.24E-03	2.28E-03	-0.982
β_1	P_BUY_t	-5.68E-03***	1.78E-03	-3.19
β_2	P_USE_t	3.07E-03**	1.37E-03	2.233
$\gamma_{1,-3}$	$SPREM(1)_{t-3}$	5.81E-06	1.15E-05	0.504
$\gamma_{1,-2}$	$SPREM(1)_{t-2}$	-3.56E-05**	1.63E-05	-2.18
$\gamma_{1,-1}$	$SPREM(1)_{t-1}$	4.40E-05***	1.61E-05	2.731
$\gamma_{1,0}$	$SPREM(1)_t$	-3.71E-05**	1.64E-05	-2.259
$\gamma_{1,1}$	$SPREM(1)_{t+1}$	1.53E-05	1.61E-05	0.951
$\gamma_{1,2}$	$SPREM(1)_{t+2}$	1.95E-05	1.61E-05	1.212
$\gamma_{1,3}$	$SPREM(1)_{t+3}$	1.67E-06	1.14E-05	0.146
$\gamma_{2,-3}$	$SPREM(2)_{t-3}$	-2.52E-07	3.61E-07	-0.7
$\gamma_{2,-2}$	$SPREM(2)_{t-2}$	-4.23E-07	4.94E-07	-0.858
$\gamma_{2,-1}$	$SPREM(2)_{t-1}$	3.16E-07	4.94E-07	0.639
$\gamma_{2,0}$	$SPREM(2)_t$	8.90E-07*	4.94E-07	1.801
$\gamma_{2,1}$	$SPREM(2)_{t+1}$	1.56E-08	4.95E-07	0.032
$\gamma_{2,2}$	$SPREM(2)_{t+2}$	-1.05E-07	4.96E-07	-0.212
$\gamma_{2,3}$	$SPREM(2)_{t+3}$	1.61E-07	3.85E-07	0.42
$\gamma_{3,-3}$	$SPREM(3)_{t-3}$	-9.77E-07*	5.06E-07	-1.931
$\gamma_{3,-2}$	$SPREM(3)_{t-2}$	6.24E-07	6.78E-07	0.921
$\gamma_{3,-1}$	$SPREM(3)_{t-1}$	-4.79E-06***	6.80E-07	-7.053
$\gamma_{3,0}$	$SPREM(3)_t$	6.80E-06***	6.79E-07	10.004
$\gamma_{3,1}$	$SPREM(3)_{t+1}$	-1.04E-06	6.83E-07	-1.527
$\gamma_{3,2}$	$SPREM(3)_{t+2}$	-9.99E-07	6.91E-07	-1.446
$\gamma_{3,3}$	$SPREM(3)_{t+3}$	1.05E-06*	5.41E-07	1.943

Table 1 – Estimation results. The table displays the coefficients from an OLS estimation of a model of the scrappage rate. The dependent variable is thus the number of cars scrapped per month divided by the total car stock. The second index on the coefficients for the scrappage premium indicates the number of lags (minus) or leads. Estimates with one asterisk are significantly different from zero at the 10 percent level, two asterisks indicates significance at the 5 percent level, and three asterisks at the 1 percent level of significance.

EFI
The Economic Research Institute

Reports since 1995

Published in the language indicated by the title

1999

Andersson, P., Experto Credite: Three Papers on Experienced Decision Makers.
Ekman, G., Från text till batong – Om poliser, busar och svennar.
Gunnarsson, J., Portfolio-Based Segmentation and Consumer Behavior: Empirical Evidence and Methodological Issues.
Hamrefors, S., Spontaneous Environmental Scanning.
Japanese Production Management in Sunrise or Sunset. Christer Karlsson (editor).
Lindé, J., Essays on the Effects of Fiscal and Monetary Policy.
Ljunggren, U., Indikatorer i grundskolan i Stockholms stad före stadsdels-nämndsreformen – en kartläggning
Mägi, A., Store Loyalty? An Empirical Study of Grocery Shopping.
Nilsson, K., Ledtider för ledningsinformation.
Rognes, J., Telecommuting – Organisational Impact of Home Based – Telecommuting. 1999.
Skalin, J., Modelling Macroeconomic Time Series with Smooth Transition Autoregressions.
Spagnolo, G., Essays on Managerial Incentives and Product-Market Competition.
Strauss, T., Governance and Structural Adjustment Programs: Effects on Investment, Growth and Income Distribution.
Svedberg Nilsson, K., Effektiva företag? En studie av hur privatiserade organisationer konstrueras.
Söderström, U., Monetary Policy under Uncertainty.

1998

Andersson, M., On Testing and Forecasting in Fractionally Integrated Time Series Models.
Berg-Suurwee, U., Styrning av kultur- och fritidsförvaltning innan stadsdelsnämnds-reformen
Berg-Suurwee, U., Nyckeltal avseende kultur- och fritidsförvaltning innan stadsdels-nämndsreformen.
Bergström, F., Essays on the Political Economy of Industrial Policy.
Bild, M., Valuation of Takeovers.
Charpentier, C., Samuelson, L.A., Effekter av en sjukvårdsreform – en analys av Stockholmsmodellen.
Eriksson-Skoog, G., The Soft Budget Constraint: The Emergence, Persistence and Logic of an Institution. The Case of Tanzania 1967-1992.
Gredenhoff, M., Bootstrap Inference in Time Series Econometrics.
Ioannidis, D., I nationens tjänst? Strategisk handling i politisk miljö – en nationell teleoperatörs interorganisatoriska, strategiska utveckling.
Johansson, S., Savings Investment, and Economic Reforms in Developing Countries.
Levin, J., Essays in Company Valuation.

Ljunggren, U., Styrning av grundskolan i Stockholms stad innan stadsdelsnämnds-reformen.
Mattsson, S., Från stat till marknad – effekter på nätverksrelationer vid en bolagise-ringsreform.
Nyberg, A., Innovation in Distribution Channels – An Evolutionary Approach.
Olsson, P., Studies in Company Valuation.
Reneby, J., Pricing Corporate Securities.
Roszbach, K., Essays on Banking Credit and Interest Rates.
Runsten, M., The Association Between Accounting Information and Stock Prices. Model development and empirical tests based on Swedish Data.
Segendorff, B., Essays on Bargaining and Delegation.
Sjöberg, L., **Bagozzi, R.**, **Ingvar, D.H.**, Will and Economic Behavior.
Sjögren, A., Perspectives on Human Capital: Economic Growth, Occupational Choice and Intergenerational Mobility.
Studier i kostnadsintäktsanalys, red Jennergren, P.
Söderholm, J., Målstyrning av decentraliserade organisationer. Styrning mot finansiella och icke-finansiella mål.
Thorburn, K., Cash Auction Bankruptcy and Corporate Restructuring
Wijkström, F., Different Faces of Civil Society.
Zethraeus, N., Essays on Economic Evaluation in Health Care. Evaluation of Hormone Replacement Therapy and Uncertainty in Economic Evaluations.

1997

Alexius, A., Essays on Exchange Rates, Prices and Interest Rates.
Andersson, B., Essays on the Swedish Electricity Market.
Berggren, N., Essays in Constitutional Economics.
Ericsson, J., Credit Risk in Corporate Securities and Derivatives. Valuation and Optimal Capital Structure Choice.
Charpentier, C., Budgeteringens roller, aktörer och effekter. En studie av budget-processerna i en offentlig organisation.
De Geer, H., **Silfverberg, G.**, Citizens'Trust and Authorities' Choices, A Report from The Fourth International Conference on Ethics in the Public Service, Stockholm June 15-18, 1994.
Friberg, R., Prices, Profits and Exchange Rates.
Från optionsprissättning till konkurslagstiftning. red. Bergström, C., Björk, T.
Hagerud, G.E., A New Non-Linear GARCH Model.
Haksar, A., Environmental Effects of Economywide Policies: Case Studies of Costa Rica and Sri Lanka.
He, C., Statistical Properties of Garch Processes.
Holmgren, M., Datorbaserat kontrollrum inom processindustrin; erfarenheter i ett tidsperspektiv.
Jennergren, P., Tutorial on the McKinsey Model for Valuation of Companies.
Lagerlöf, J., Essays on Political Economy, Information, and Welfare.
Lange, F., **Wahlund, R.**, Planerade och oplanerade köp - Konsumenternas planering och köp av dagligvaror.
Löthgren, M., Essays on Efficiency and Productivity; Contributions on Bootstrap, DEA and Stochastic Frontier Models.
Nilsson, H.E., Using Conceptual Data Modelling in Management Accounting: A Case Study.
Sjöberg, L., **Ramsberg, J.**, En analys av en samhällsekonomisk bedömning av ändrade säkerhetsföreskrifter rörande heta arbeten.
Säfvenblad, P., Price Formation in Multi-Asset Securities Markets.

Sällström, S., On the Dynamics of Price Quality.
Södergren, B., På väg mot en horisontell organisation? Erfarenheter från näringslivet av decentralisering och därefter.
Tambour, M., Essays on Performance Measurement in Health Care.
Thorén, B., Berg-Surwee, U., Områdesarbete i Östra Hökarängen - ett försök att studera effekter av decentralisering.
Zhang Gang., Chinese Rural Enterprises Between Plan and Market.
Åhlström, P., Sequences in the Profess of Adopting Lean Production.
Åkesson, G., Företagsledning i strategiskt vakuum. Om aktörer och förändringsprocesser.
Åsbrink, S., Nonlinearities and Regime Shifts in Financial Time Series.

1996

Advancing your Business. People and Information Systems in Concert.
red. Lundeberg, M., Sundgren, B.
Att föra verksamheten framåt. Människor och informationssystem i samverkan.
red. Lundeberg, M., Sundgren, B.
Andersson, P., Concurrence, Transition and Evolution - Perspectives of Industrial Marketing Change Processes.
Andersson, P., The Emergence and Change of Pharmacia Biotech 1959-1995. The Power of the Slow Flow and the Drama of Great Events.
Asplund, M., Essays in Industrial Economics.
Delmar, F., Entrepreneurial Behavior & Business Performance.
Edlund, L., The Marriage Market: How Do You Compare?
Gunnarsson, J., Three Studies of Financial Behavior. Research Report.
Hedborg, A., Studies of Framing, Judgment and Choice.
Holgersson, C., Höök, P., Ledarutveckling för kvinnor - Uppföljning av en satsning på Volvo - Research Report.
Holmberg, C., Stores and Consumers - Two Perspectives on Food Purchasing.
Håkansson, P., Wahlund, R., Varumärken. Från teori till praktik.
Karlsson, A., The Family Business as an Heirloom. Research Report.
Linghag, S., Man är handelsstudent. Research Report.
Molin, J., Essays on Corporate Finance and Governance.
Mägi, A., The French Food Retailing Industry - A Descriptive Study.
Mölleryd, B., Så byggdes en världsindustri - Entreprenörskapets betydelse för svensk mobiltelefoni. Research Report.
Nielsen, S., Omkostningskalkulation for avancerede produktions-omgivelser - en sammenligning af stokastiske og deterministiske omkost-ningskalkulationsmodeller.
Normark, P., Danielsson, L., Larsson, A., Lundblad, P., Kooperativa nyckeltal. Research Report.
Sandin, R., Heterogeneity in Oligopoly: Theories and Tests.
Sandén-Håkansson, U., Från kampanjmål till mediemix - en studie av samarbete mellan annonsörer, reklambyråer och mediebyråer. Research Report.
Stein, J., Söderlund, M., Framgång i arbetet, strategier för att utföra arbetsuppgifterna, arbetsuppgifternas karaktär och utbildningskvalitet. En empirisk studie av civilekonomer. Research Report.
Strömberg, P., Thorburn, K., An Empirical Investigation of Swedish Corporations in Liquidation Bankruptcy. Research Report.
Söderlund, M., Och ge oss den nöjda kunden. En studie av kundtillfredsställelse och dess orsaker och effekter. Research Report.
Thodenius, B., Användningen av ledningsinformationssystem i Sverige: Lägesbild 1995. Research Report.

Ulfsdotter, U., Internationalisering för expansion eller hemmamarknadsförsvar? De nordiska marknaderna för fruktyoghurt 1982-1994.
Westelius, A., A Study of Patterns of Communication in Management Accounting and Control Projects.
Wijkström, F., Den svenska ideella sektorn och pengarna. Research Report.
Örtendahl, M., Health and Time - A Problem of Discounting.

1995

Becker, T., Essays on Stochastic Fiscal Policy, Public Debt and Private Consumption.
Blomberg, J., Ordning och kaos i projektsamarbete - en socialfenomenologisk upplösning av en organisationsteoretisk paradox.
Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L., Styrelsearbete i koncerner
Brännström, T., Bias Approximation and Reduction in Vector Autoregressive Models. Research Report.
Ekonomisk politik i omvandling. red. Jonung, L.
Gunnarsson, J., Wahlund, R., Hushållens finansiella strategier. En explorativ studie.
Höök, P., Chefsutveckling ur könsperspektiv - Mentorskap och nätverk på Vattenfall - Research Report.
Levin, J., Olsson, P., Looking Beyond the Horizon and Other Issues in Company Valuation. Research Report.
Mägi, A., Customer Statisfaction in a Store Performance Framework. Research Report.
Persson, P-G., Modeling the Impact of Sales Promotion on Store Profits.
Roman, L., Institutions in Transition. A Study of Vietnamese Banking.
Sandberg, J., Statistisk metod - dess vetenskapliga hemvist, grundläggande principer och möjligheter inom samhällsvetenskapen. Research Report.
Sandberg, J., How Do We Justify Knowledge Produced by Interpretative Approaches? Research Report.
Schuster, W., Redovisning av konvertibla skuldebrev och konvertibla vinstandelsbevis - klassificering och värdering.
Söderberg, K., Farmartjänsten - Ny kooperation inom lantbruket. Research Report.
Söderqvist, T., Benefit Estimation in the Case of Nonmarket Goods. Four Essays on Reductions of Health Risks Due to Residential Radon Radiation.
Thorén, B., Användning av information vid ekonomisk styrning - månadsrapporter och andra informationskällor.