Essays on Environmental Policy-Making in Developing Countries: Applications to Costa Rica.

Francisco Alpízar R.

Department of Economics
Göteborg University
Box 640, SE405-30 Göteborg
SWEDEN
e-mail: Francisco.Alpizar@economics.gu.se

Contents

Abstracts

This thesis consists of five papers dealing with fairly heterogeneous issues, based on the problems or topics analyzed, but also based on the methodologies used to approach them. The overriding motives are the design of environmental policies in the context of a typical developing country (where Costa Rica is used as a representative of such countries), and the study and application of techniques that can provide the necessary information for policy-making.

Paper 1: "Using Choice Experiments for Valuing the Environment". (with F.Carlsson and P. Martinsson)

This paper provides the latest research developments in the method of choice experiments applied to valuation of non-market goods. Choice experiments, along with the, by now, well-known contingent valuation method, are very important tools for valuing non-market goods and the results are used in both cost-benefit analyses and litigations related to damage assessments. The paper should provide the reader with both the means to carry out a choice experiment and to conduct a detailed critical analysis of its performance in order to give informed advice about the results. A discussion of the underlying economic model of choice experiments is incorporated, as well as a presentation of econometric models consistent with economic theory. Furthermore, a detailed discussion on the development of a choice experiment is provided, which in particular focuses on the design of the experiment and tests of validity. Finally, a discussion on different ways to calculate welfare effects is presented.

Paper 2: "Policy Implications and Analysis of the Determinants of Travel Mode Choice: An Application of Choice Experiments to Metropolitan Costa Rica". (with F. Carlsson)

In this paper we study a group of policies aimed at discouraging the use of private transportation during peak hours, both directly and indirectly, by increasing the attractiveness of the only available substitute, the bus. This is done using a choice experiment constructed to find the answer to the following basic question: Given fixed house-to-work structures and no working hour flexibility, by how much is the choice of travel mode for commuters to work sensitive to changes in travel time, changes in costs for each mode and other service attributes? This information is then used to identify the most suitable combination of policies dealing with air pollution and congestion in the typical developing country context of metropolitan Costa Rica. We also provide estimates of the value of travel time as a measure of the potential benefits gained from reduced congestion.

Paper 3: "How much do we care about absolute versus relative income and consumption?" (with F. Carlsson and O. Johansson).

We find, using survey-experimental methods, that most individuals are concerned with both relative income and relative consumption of particular goods. The degree of concern varies in the expected direction depending on the properties of the good. However, contrary to what has been suggested in the previous literature, we find that relative consumption is also important for vacation and insurance, which are typically seen as non-positional goods. Further, absolute consumption is also found to be important for cars and housing, which are widely regarded as highly positional. Implications for Pareto-efficient taxation are illustrated using the results from the experiment.

Paper 4: "Collective versus Random Fining: An Experimental Study on Controlling Non-Point Pollution". (with Prof. Till Requate)

This paper presents an experimental study of collective vis-à-vis random fining as means to regulate non-point pollution. Using samples from both managers from Costa-Rican coffee mills, and Costa-Rican students we confirm the hypothesis that the two games are equivalent and, in the majority of cases, lead to efficient outcomes through Nash play. However, we reject the hypothesis that managers and students behave equally. Off the equilibrium managers tend to over-abate, whereas students tend to under-abate. This result suggests the importance of considering subject pool differences in the evaluation of environmental policies by means of experiments, particularly if those policies involve certain forms of management decisions.

Paper 5: “The Pricing of Protected Areas in Nature-Based Tourism: A Local Perspective

This paper seeks to provide the theoretical underpinnings for the optimal pricing of protected areas used in recreation, from the perspective of a local park agency interested in maximizing welfare subject to a budget constraint. The theoretical model incorporates several relevant issues, including third degree price discrimination based on visitor type, public versus private benefits and costs, economy-wide impact of changes in entrance fees, congestion costs and investment in quality. The model also includes a distributional dimension that is very relevant for the case of international nature-based tourism in poorer, less developed countries. Based on this foundation and using visitation data from the Costa Rican National Park System, the paper provides an estimation of the optimal entrance fees and associated revenues under the assumption, due to the lack of information, of no congestion or spill over effects. Still, the theoretical model allows the analysis of those effects taking the assumed case as a baseline.

Preface

First of all, I would like to thank Thomas Sterner and Gunnar Köhlin for their dedication to something that has changed my life and that of many others. If your objective is to “make the world a better place”, in Thomas and Gunnar you will find two perfect examples of how to do things correctly. Their kindness, generosity, honorability and professionalism have been an example to me since the beginning. With all my heart I hope that the future allows them to continue the task of improving the conditions in developing countries through education and research. The Swedish International Development Agency has been an important ally in this process, and I whish to thank them for the financial support of my studies.

Second, in one way or another, this thesis has been marked by my dear friend and colleague, Fredrik Carlsson. All chapters, including the ones were he was not coauthor, have been the scene of fierce battles between Fredrik and myself. These battles ranged from punctuation to economic theory passing trough movies and music, and took place in cafes and seminar rooms indistinctively. Olof Johansson frequently took part in these battles, bringing his enormous knowledge and sniper-sharp vision to the battlefield. From these discussions I learned the fun in doing research, and for that I will forever be indebted to Fredrik and Olof. The combination of dedication, patience, enthusiasm, friendship and sharp criticism made them the ideal supervisory team.

I would like to thank my other co-authors, Peter Martinsson and Till Requate, for their dedication. It has been a pleasure and an honor to work with you. Several others have kindly and enthusiastically participated in improving the chapters of this thesis, providing ideas, comments and, above all, their precious time. My gratitude goes out to Gardner Brown, Karl-Göran Mäler, Henk Folmer, Thomas Sterner, Gunnar Köhlin, Sharon Kühlman, John List and Katrin Millock.

I strongly believe that achieving a PhD is a matter of mental health as much as of mental capacity. Fredrik, Åsa, Henrik, Anna, Mattias, Susanna, Johan, Peter, Sharon and Victor opened their lives to make mine happier and mentally safer. Through lunches, beers, parties and squash games, these friends and several others, too many to mention them all,

have made my years in Sweden a delightful experience. In addition, the Environmental Economics Unit has provided a stronghold to develop as a person and as a researcher. To all of you, my deepest gratitude. Special thanks to Katarina and Elizabeth for their help and guidance.

A great deal of the person I have become is thanks to my family and friends. Tata siempre decía que no hay que soltarle el rabo a la chancha aunque te cage la mano, y esa frase ha marcado mi camino desde la infancia. Mis abuelos son y serán siempre ejemplos de trabajo duro, perseverancia y determinación. Quiero agradecer especialmente a mi madre, la mamá gallina, por su amor y apoyo incondicional. Este logro comenzó en las tardes de dictados y caligrafía y por tanto es también tu logro. Gracias a Felo y a Macho, mis hermanos, y a mi tío Bernard por haberme ayudado cuando lo necesitaba. Gracias a ellos tuve desde oficina hasta asistentes de investigación en Costa Rica. Mi agradecimiento es también para mi padre, por haber tenido la claridad para ver que en la educación estaba mi futuro y su mejor herencia. I would also like to thank Natacha and don Eduardo, my parents-in-law, for their support and friendship. In addition, don Eduardo was an economist of exceptional caliber and professionalism, and served me as an example of how things can go right as an economist in Latin America. I whish he was still among us - he would have been one of the first to receive this thesis.

There is a saying in Spanish, "dime con quien andas y te dire quién eres". I have always felt proud and honored of having the friends of the "panel" with me. They have not only fiercely believed in me, but have "taken care" of Gaby when I was not there, and even received me with cheers and posters at the airport. I have known most of the "panelistas" since I was a kid, we have stayed together through good and bad times, and after many years of almost daily contact via internet I have the happy feeling that we will always be together, no matter how far apart. Knowing this made my days of loneliness bearable, and I thank you all for that.

Finally, much of the fun in my life comes from my lover, my friend, my partner and wife, all in one, Gaby. She has been involved in this thesis in many ways. She was patient and loving enough to stick to me through all those years that we lived apart, and when the time came, she was strong and supportive to the point that she postponed a brilliant career

to help me get one. She has read several of the papers in this thesis and provided lapidarian criticism where it was required ("you economists think that human behavior can be studied without ever reading Freud"). Most importantly, Gaby has the ability of pulling me out of whatever I am doing by simply closing the door behind me when I get home. The joy of what you do depends on it not becoming a permanent burden. I enjoy what I do, a great deal thanks to her. For this and many other reasons, her name should be next to mine in the cover of this booklet.

Introduction

This thesis contains five separate self-contained chapters, dealing with a heterogeneous selection of environmental problems, and using three different broad methodologies. All five papers revolve around two basic issues: *i*) the design and implementation of environmental policies with an explicit consideration of the particular conditions present in most less developed countries; and *ii*) the study and application of methods capable of providing the necessary information for the success of environmental policy-making. I have used my home country, Costa Rica, as the source of relevant environmental issues to study in this thesis, and accordingly, all data for the empirical applications were obtained there. Still, I have made an effort to generalize my results whenever possible, and I believe that all the chapters, except possibly Paper 2, contain conclusions that can be generalized to most developing countries. Even Paper 2, from a methodological point of view, contains interesting results for future applications of choice experiments in other countries.

Environmental Policy in Developing Countries

Any environmental program must be designed and implemented in a way that makes it compatible with its socio-economic and political environment. Less developed countries exhibit certain charactersitics that add complexity to the selection of instruments for environmental policy-making. In my opinion, there are four elements that need particular consideration.

1. In designing a structure of economic incentives one must always bear in mind the distributional impacts associated with it. Although relevant in all settings, this fact acquires special significance in light of the poverty problem which exists in all less developed countries.
2. The commitment to environmental solutions is in many cases rather weak. Environmental problems are not always high on the priority list of most citizens of developing countries, and consequently they usually have low priority on the political agenda. As a result, policies are likely to be abandoned if they become a

threat to short term private economic prosperity, or they may be simply "adjusted" in order to fulfill the wishes of powerful lobby groups.

3. Both market-based and command-and-control instruments require strong institutions, adequate legislation, and effective monitoring and enforcement. In practice, many developing countries show institutional weaknesses, such as inexperienced and underpaid staff, unclear jurisdiction and corruption. Another problem related to implementation is the lack of resources to conduct frequent monitoring and even to buy the technology required to enforce certain policies.
4. There is generally very limited information on baseline environmental quality, on firm environmental behavior, and on consumer's preferences for environmental goods and services, which makes it especially difficult to design and evaluate the success of environmental policies. This is why the study of suitable techniques for obtaining the information necessary for the successful design of environmental policy instruments is crucial.

From the first two elements we typically have that in order to be successful, environmental policy-makers should not only avoid policies that hurt the poor, but they must also consider that governmental authorities will be very hesitant to implement regulation that could jeopardize short-run production. Paper 2 (*"Policy Implications and Analysis of the Determinants of Travel Mode Choice: An Application of Choice Experiments to Metropolitan Costa Rica"*), Paper 3 (*"How Much Do We Care about Absolute Versus Relative Income and Consumption?"*) and Paper 5 ("*The Pricing of Protected Areas in Nature-Based Tourism: A Local Perspective"*) deal explicitly with these two issues.

Paper 2 studies the commuter's mode-choice decision in metropolitan Costa Rica (car versus bus) with the aim of designing policies to discourage the use of private transportation, while at the same time providing a suitable alternative for getting to work. Such a combination of policies seeks to reduce congestion and possibly pollution, while addressing the potential opposition to, and distributional impacts of, more expensive private transport.

Paper 3 arose initially as a byproduct of Paper 2. In preparing for the study of the mode-choice decision, it was necessary to discuss the car-ownership decision. It was argued that it was simpler to affect the choice of mode for commuting to work, since the decision to buy a car might be motivated by factors not directly related to transportation. In particular, cars can play an important role in defining status. This is an interesting topic in itself, so for Paper 3 my coauthors and I conducted an experimental survey to determine by how much a unit of consumption (of cars, houses, insurance or days of vacation) or income contributed to an individual's utility, either directly via an absolute increase in consumption, or indirectly, via comparison with other members of society. The students were presented with different imagined societies described both by their own income or consumption and by the averages in the society. Subjects were asked to choose repeatedly in a way that allows the calculation of their degree of 'positionality', or loosely speaking concern for status. This was done for various goods ranging from insurance to automobiles.

Paper 5 develops a theoretical model for optimal pricing of recreation in protected areas (parks, natural reserves, etc.) in the presence of a cost-recovery or minimum-profit constraint. The model also incorporates a distributional dimension by exploring the impact of different welfare weights on the consumer surplus of foreign visitors to the park. Increasing the financial sustainability of protected areas is important for several reasons, especially since it sends a signal to the government that conservation is not necessarily a financial burden.

Tourism plays an important role in increasing the profitability of the national park. In addition, in recent years several private reserves have been established, with entrance fees as one of their main source of funding. I studied the use of price discrimination between foreign and national visitors in the context of a park agency interested in maximizing national welfare from visitation to a system of protected areas, while subjected to the restriction that at least recreational cost should be recovered. Furthermore, the analysis was extended to include the possibility that governmental authorities might impose a minimum profit condition. The optimal entrance fees for recreation calculated would increase the financial sustainability of the Costa Rican system of national parks, thereby reducing the pressure to exploit these areas in alternative ways.

More generally, the environmental authorities in Costa Rica correctly believe that by increasing the profitability of conservation it would be easier to reduce deforestation on privately owned lands, increase watershed protection, promote the preservation of important ecosystems, etc. Universities, municipalities, and NGOs are thus dedicated to exploring, and sometimes to implementing innovative systems for payment for environmental services like watershed protection, scenic beauty, and, most importantly, carbon sequestration.

The third and fourth elements mentioned above as requiring particular attention in the context of developing countries relate to the application of environmental policies under institutional weaknesses and with very limited environmental information. Paper 4 (*"Collective versus Random Fining: An Experimental Study on Controlling Non-Point Pollution"*) reports on an economic experiment that studied two alternative mechanisms aimed at achieving a desired pollution reduction when only ambient pollution is observable, albeit not the discharges of any individual polluter. Ambient monitoring is likely to be cheaper and less prone to corruption than individual firm monitoring. The subjects in this experiment played two non-cooperative games where the pay-offs in each game were determined by the mechanisms for pollution reduction. For this game, my coauthor and I interviewed two different types of subject pools. The experiments were conducted using, first, the usual convenience-sample of students; in our case, these students were from the University of Costa Rica. The second subject pool was a sample of real managers of polluting firms, specifically coffee mill managers. The experiments were not particularly addressed to this type of industry, but rather we view our sample of managers as representative sample of this type of individuals in any industry. This setting allowed us to compare the behavior of subjects from the convenience sample with that of a more appropriate set of subjects. Significant differences were identified and analyzed.

Given the typical lack of appropriate information it is natural to also look for *stated*-preference or experimental methods, and hence not only rely on actual behavior or revealed-preference methods. In this thesis I have used two promising methods, namely economic experiments in Papers 3 and 4, and survey based methods, in particular choice experiments, in Paper 2.

Paper 1 (*"Using Choice Experiments for Valuing the Environment"*) attempts to provide the latest research developments in the method of choice experiments applied to valuation of non-market goods. Choice experiments, along with the by now well-known contingent valuation method, are very important tools for valuing non-market goods. In a choice experiment, individuals are given a hypothetical setting and asked to choose their preferred alternative among several alternatives in a choice set, and they are usually asked to perform a sequence of such choices. Each alternative is described by a number of attributes or characteristics, including a monetary value. Thus, when individuals make their choice, they implicitly make trade-offs between the levels of the attributes in the different alternatives presented in a choice set. The results are often used in both cost-benefit analyses and policy design. By explicitly incorporating the trade-offs of a given policy into the design of the experiment, choice experiments, in particular, can be constructed to obtain the preferences of individuals for that policy. The method of choice experiments therefore provides a powerful tool in the development of management strategies based in the trade-off between private short-term profit and the long-term social welfare obtained from a sustainable flow of environmental services. The transport application in Paper 2 exploits the potential of choice experiments with regard to various kinds of policy changes, for which there is otherwise no or very limited experience and no available data. Papers 3 and 4 are applications based on experimental data.

The choice of applications

The first three papers originated in my interest in the problems of congestion and pollution associated with transportation in metropolitan Costa Rica. As described in Paper 2, congestion and pollution are serious problems in the capital of Costa Rica, just as they are in most metropolitan areas in less developed countries, with resulting health problems, stress, deterioration of buildings, and time looses, among other things. Given the fact that the deterioration of environmental conditions in urban areas affects a large share of the population, and the urgent need for solutions, I believe this problem is possibly the most critical environmental issue for Costa Rica, and therefore was an obvious starting point for this thesis.

Another critical problem is deterioration of watersheds due to the inflow of pollution from agro-industrial and housing sources. The situation is desperate with regards to the flora and fauna of rivers and marine ecosystems close to river deltas, and is even starting to threat the provision of drinking water. Any solution has to start with the basics, namely the construction of properly working sewage systems and appropriate treatment plants. For example, although Costa Rica's legal system now prohibits the disposal of any pollutant directly into a river system, the government is well aware that municipalities routinely deposit their sewage directly into the river. The situation with agro-industrial firms is similar. The basic problems are the inability to monitor compliance with the law and corruption. One sector that has been put particularly under pressure to improve their use of water is the coffee-milling industry, so managers of these plants are well aware that they need to reduce pollution in order to avoid more drastic consequences.[1] Still, the government is not able to monitor individual coffee mills, and the fact that generally several coffee mills share the same river basin means that this could be considered as a non-point-pollution problem. For this reason, the managers of coffee mills are a perfect subject pool for exploring the efficiency of mechanisms for reducing pollution in such circumstances. The mechanisms studied in Paper 4 are only two potential ways of avoiding frequent monitoring of pollution and eliminating corruptible inspectors. Our results indicate that, after a suitable learning period, firms can understand and adapt their behavior to elaborate regulatory contracts leading to efficient outcomes. This result is particularly relevant for the Costa Rican setting, where skepticism to any regulatory contract is widespread, especially so to any elaborate one.

Finally, the last paper contains an application of optimal pricing of recreation in the Costa Rican system of national parks. The case of Costa Rica is widely cited as an ecotourism success story, but, as in any other typical developing country, the Costa Rican system of national parks is under constant threat from the lack of funds to manage the resources properly, and the pressure to exploit them in alternative commercial activities.

[1] In Costa Rica, coffee berries are peeled and then cleaned using river water, resulting in heavily polluted, stinky effluents.

In the last 15 years, Costa Rica has experienced a large increase in international tourism. Most of this increase is due to Costa Rica's system of national parks, which encompasses approximately 24% of the national territory and a vast array of different types of marine and land ecosystems. According to a survey of the local tourism board (ICT, 2000), 71% of all holiday visitors to Costa Rica visit at least one national park, and most of the activities of foreign visitors are highly related to natural attractions in the country. Despite the importance of nationally protected areas for tourism in Costa Rica, the increase in tourism activities has not caused an increase in the government budget dedicated to managing and protecting this resource. On the contrary, this budget, together with international support, has been steadily decreasing. In 1994, the government decided to recover the costs of managing the parks for recreation, and started to price discriminate based on nationality. Since then, there have been several changes in price to foreign visitors, but little or no formal criteria have been used in the determination of price changes. An improvement in the criteria for setting the entrance fee is important not only for Costa Rica, but for the rest of the countries in the region, and potentially for other regions of the world attempting to develop an ecotourism strategy for their protected areas.

Using Choice Experiments for Non-Market Valuation

Francisco Alpízar

Fredrik Carlsson

Peter Martinsson

Abstract

This paper provides the latest research developments in the method of choice experiments applied to valuation of non-market goods. Choice experiments, along with the, by now, well-known contingent valuation method, are very important tools for valuing non-market goods and the results are used in both cost-benefit analyses and litigations related to damage assessments. The paper should provide the reader with both the means to carry out a choice experiment and to conduct a detailed critical analysis of its performance in order to give informed advice about the results. A discussion of the underlying economic model of choice experiments is incorporated, as well as a presentation of econometric models consistent with economic theory. Furthermore, a detailed discussion on the development of a choice experiment is provided, which in particular focuses on the design of the experiment and tests of validity. Finally, a discussion on different ways to calculate welfare effects is presented.

Key words: Choice experiments; non-market goods; stated preference methods; valuation.

1. Introduction

The methods of valuation of non-marketed goods have become crucial when determining the costs and benefits of public projects. Non-market valuation exercises have been conducted in many different areas, ranging from health and environmental applications to transport and public infrastructure projects. In the case of a good that is not traded in a market, an economic value of that good obviously cannot be directly obtained from the market. Markets fail to exist for some goods either because these goods simply do not exist yet, or because they are public goods, for which exclusion is not possible. Nevertheless, if one wants to compare different programs by using cost-benefit analysis, the change in the quality or quantity of the non-market goods should be expressed in monetary terms. Another crucial application of valuation techniques is the determination of damages associated with a certain event. Under the Comprehensive Environmental Response, Compensation and Liability Act of 1980 in the US, and after the events that followed the Exxon Valdez oil spill in 1989, the methods of valuation have become a central part of litigation for environmental and health related damages in the United States and in several other countries.

Over the years, the research on valuation of non-market goods has developed into two branches: revealed preference methods and stated preference methods. The first branch, the revealed preference method, infers the value of a non-market good by studying actual (revealed) behaviour on a closely related market. The two most-well-known revealed preference methods are the hedonic pricing method and the travel cost method (see Braden and Kolstad, 1991). In general, the revealed preference approach has the advantage of being based on actual choices made by individuals. However, there are also a number of drawbacks; most notably that the valuation is conditioned on current and previous levels of the non-market good and the impossibility of measuring non-use values, i.e. the value of the non-market good not related to usage such as existence value, altruistic value and bequest value. Thus, research in the area of valuation of non-market goods has therefore seen an increased interest in another branch, the stated preference method, during the last 20 years.

Stated preference method assesses the value of non-market goods by using individuals' stated behaviour in a hypothetical setting. The method includes a number of different approaches such as conjoint analysis, contingent valuation method (CVM) and choice experiments. In most applications, CVM has been the most commonly used approach. In particular, closed-ended CVM surveys have been used, in which respondents are asked whether or not they would be willing to pay a certain amount of money for realizing the level of the non-market good described or, more precisely, the change in the level of the good (see Bateman and Willis, 1999 for a review). The idea of CVM was first suggested by Ciriacy-Wantrup (1947), and the first study ever done was in 1961 by Davis (1963). Since then, CVM surveys have become one of the most commonly used methods for valuation of non-market goods, although its use has been questioned (see e.g. Diamond and Hausman, 1994 and Hanemann, 1994, for a critical assessment). At the same time as CVM was developed, other types of stated preference techniques, such as choice experiments, evolved in both marketing and transport economics (see Louviere 1993 and Polak and Jones 1993 for overviews).

In a choice experiment, individuals are given a hypothetical setting and asked to choose their preferred alternative among several alternatives in a choice set, and they are usually asked to perform a sequence of such choices. Each alternative is described by a number of attributes or characteristics. A monetary value is included as one of the attributes, along with other attributes of importance, when describing the profile of the alternative presented (see figure 1). Thus, when individuals make their choice, they implicitly make trade-offs between the levels of the attributes in the different alternatives presented in a choice set.

Figure 1: This is an example of a choice set containing two profiles of a given alternative (a park). Each profile is described in terms of 4 attributes, including the entrance fee. Each attribute has two or more levels. A choice experiment contains a sequence of such choice sets.

	Park A	**Park B**
Available facilities	*Visitor center*	*Information office*
Extension of walking tracks	*2 kms*	*10 kms*
Condition of tracks	*Rustic tracks*	*Stoned tracks*
Entrance fee	*8 US$*	*10 US$*

Which of the two options would you prefer for a one day visit?

□ Park A □ Park B

The purpose of this paper is to give a detailed description of the steps involved in a choice experiment and to discuss the use of this method for valuing non-market goods. Choice experiments are becoming ever more frequently applied to the valuation of non-market goods. This method gives the value of a certain good by separately evaluating the preferences of individuals for the relevant attributes that characterize that good, and in doing so it also provides a large amount of information that can be used in determining the preferred design of the good. In fact, choice experiments originated in the fields of transport and marketing, where it was mainly used to study the trade offs between the characteristics of transport projects and private goods, respectively. Choice experiments have a long tradition in those fields, and they have only recently been applied to non-market goods in environmental and health economics. We believe that applications of this technique will become more frequent in other areas of economics as well. Only recently has the aim of damage assessment in litigation shifted from monetary compensation to resource compensation. Therefore identification and evaluation of the different attributes of a damaged good is required in order to design the preferred restoration project (Adamowicz et. al., 1998b; Layton and Brown, 1998). Choice experiments are especially well suited for this purpose, and one could expect this method to be a central part of future litigation processes involving non-market goods.

The first study to apply choice experiments to non-market valuation was Adamowicz et al. (1994). Since then there has been an increasing number of studies, see e.g. Adamowicz et al. (1998a), Boxall et al. (1996), Layton and Brown (2000) for applications to environment, and e.g. Ryan and Hughes (1997) and Vick and Scott (1998) for applications to health. There are several reasons for the increased interest in choice experiments in addition to those mentioned above: (i) reduction of some of the potential biases of CVM, (ii) more information is elicited from each respondent compared to CVM and (iii) the possibility of testing for internal consistency.

In a choice experiment, as well as in a CVM survey, the economic model is intrinsically linked to the statistical model. The economic model is the basis of the analysis, and as such, affects the design of the survey and the analysis of the data. In this sense, we argue that the realization of a choice experiment is best viewed as an integrated and cyclical process that starts with an economic model describing the issue to analyse. This model is then continually revised as new information is received from the experimental design, the statistical model, focus groups and pilot studies, etc. In this paper, we pay special attention to the link between the microeconomic and the statistical foundations of a choice experiment, when it comes to designing the choice experiment, estimating the econometric model as well as calculating welfare measures. Furthermore, we address the issue of internal and external validity of a choice experiment, and provide a discussion of the possibility of misrepresentation of preferences by strategic responses. The literature on choice experiments has been reviewed by other authors, e.g.

Adamowicz et al., 1998b; Hanley et al., 1998; Louviere et al., 2000. This paper contributes to providing a thorough description of each of the steps needed when performing a choice experiment on a non-market good, with special attention to the latest research results in design and estimation.

The rest of the paper is organized as follows: Section 2 discusses the underlying economic theory of choice experiments. In Section 3, econometric models are discussed and linked to the section on economic theory. Section 4 concentrates on the design of a choice experiment, given the theoretical and empirical models presented in the two previous sections. Respondent behaviour and potential biases are discussed in Section 5. Section 6 presents different techniques to apply when estimating welfare effects. Finally, Section 7 concludes this paper.

2. The Economic Model

The basis for most microeconomic models of consumer behavior is the maximization of a utility function subject to a budget constraint. Choice experiments were inspired by the Lancasterian microeconomic approach (Lancaster, 1966), in which individuals derive utility from the characteristics of the goods rather than directly from the goods themselves. As a result, a change in prices can cause a discrete switch from one bundle of goods to another that will provide the most cost-efficient combination of attributes. In order to explain the underlying theory of choice experiments, we need to link the Lancasterian theory of value with models of consumer demand for discrete choices (Hanemann, 1984 and 1999).

In many situations, an individual's decisions can be partitioned into two parts: (i) which good to choose and (ii) how much to consume of the chosen good. Hanemann (1984) calls this a discrete/continuous choice. An example of this choice structure is the case of a tourist deciding to visit a national park. The decision can be partitioned into which park to visit, and how long to stay. In order to obtain a value of a certain park, both stages of the decision-making process are crucial to the analysis and should be modelled in a consistent manner.

In general, choice experiments applied to non-marketed goods assume a specific continuous dimension as part of the framework, in which a discrete choice takes place. By referring to the example above, one could ask for a discrete choice (which type of park do you prefer to visit?) given a one week (day, month) trip. In this case, the decision context is constructed so that it isolates the discrete choice, therefore allowing the individual to make a purely discrete choice (Hanemann, 1999). A CVM survey assumes the same specific continuous dimension since the objective is to obtain the value of a certain predefined program that includes a given continuous decision. Finally, note that many non-marketed goods are actually public in nature, especially in the sense that the same quantity of the good is available for all agents. In such cases, each individual can only choose one of the offered alternatives, given its cost.

The economic model presented in this section deals only with such purely discrete choices. For more information on the discrete and continuous choice see Hanemann (1984). Formally, each individual solves the following maximization problem:

$$
\begin{aligned}
& Max_{c,x} U\left[c_1(A_1),\ldots,c_N(A_N); z\right] \\
& s.t. \qquad i.\ y = \sum_{i=1}^{N} p_i c_i(A_i) + z \\
& \qquad\qquad ii.\ c_i c_j = 0\,,\ \forall i \neq j \\
& \qquad\qquad iii.\ z \geq 0\,,\ c_i(A_i) \geq 0 \text{ for at least one } i
\end{aligned}
\tag{1}
$$

where, $U[..]$ is a quasiconcave utility function; $c_i(A_i)$ is alternative combination i (profile i) as a function of its generic and alternative specific attributes, the vector A_i; p_i is the price of each profile; z is a composite bundle of ordinary goods with its price normalized to 1 and y is income. A number of properties follow from the specification of the maximization problem:

1. The $c_i's$ are profiles defined for all the relevant alternatives. For example, one such profile could be a visit to a national park in a rainforest, with 50 kms of marked walking tracks through the park and a visitor center. Additionally, the choice of any profile is for a fixed, and given, amount of it, e.g. a day or a unit. There are N such profiles, where N is in principle given by all relevant profiles. However, in practice, N will be determined depending on the type of design used to construct the profiles, the number of attributes, and the attribute levels included in the choice experiment. Consequently, with the

selection of attributes and attribute levels for a choice experiment we are already limiting or defining the utility function.

2. The price variable in the budget restriction must be related to the complete profile of the alternative, including the given continuous dimension, for example price per day or per visit.

3. Restriction *ii* defines the number of alternatives that can be chosen. In general, in a choice experiment we are interested in obtaining a single choice. For example, in the case of perfect substitutes, there will be a corner solution with only one profile chosen.[1] Alternatively, the choice experiment can specify the need for a single choice. If the alternatives refer to different public goods or environmental amenities, one can specify that only one will be available. Even if the alternatives refer to private goods such as a specific treatment program, the researcher can specify that only one of them can be chosen.

4. In a purely discrete choice, the selection of a particular profile $c_j(A_j)$, which is provided in an exogenously fixed quantity, implies that, for a given income, the amount of ordinary goods *z* that can be purchased is also fixed. Combining this with the restriction that only a single profile, c_j, can be chosen results in:

$$z = y - p_j c_j \tag{2}$$

5. Restriction *iii* specifies that the individual will choose a non-negative quantity of the composite good and the goods being studied. If we believe that the good is essential to the individual or that an environmental program has to be implemented, then we have to force the respondent to make a choice ($c_i > 0$ for at least one i).

To solve the maximization problem we follow a two-step process. First we assume a discrete choice, profile *j* is chosen, i.e. $c_j = c_j^{fixed}, c_i = 0 \quad \forall i \neq j$, where c_j^{fixed} is the fixed continuous measure of the given profile. We further assume weak complementarity, i.e. the attributes of the other non-selected profiles do not affect the utility function of profile *j* (Mäler, 1974; Hanemann, 1984). Formally we write:

---

[1] In the case of perfect substitutes, it is the form of the utility function rather than restriction *ii* that ensures a single choice.

$$\text{if } c_i = 0 \text{, then } \frac{\partial U}{\partial A_i} = 0 \text{, } \forall i \neq j \text{.} \tag{3}$$

Using (2) and (3) we can write the conditional utility function, given $c_j = c_j^{fixed}$ as:

$$\mathrm{U_j = V_j[c_j(A_j), p_j, y, z] = V_j(A_j, y - p_j c_j)}. \tag{4}$$

In the next step we go back to the unconditional indirect utility function:

$$\mathrm{V[A, p, y] = \max[V_1\ (A_1, y - p_1 c_1), ..., V_N(A_N, y - p_N c_N)]}, \tag{5}$$

where the function $V[..]$ captures the discrete choice, given an exogenous and fixed quantitative assumption regarding the continuous choice. Thus, it follows that the individual chooses the profile j if and only if:

$$\mathrm{V_j(A_j, y - p_j c_j) > V_i(A_i, y - p_i c_i)},\ \forall i \neq j \tag{6}$$

Equations (5) and (6) complete the economic model for purely discrete choices. These two equations are the basis for the econometric model and the estimation of welfare effects that are discussed in the following sections.

Note that the economic model underlying a CVM study can be seen as a special case of the model above, where there are only two profiles. One profile is the "before the project" description of the good, and the other is the "after the project" description of the same good. Thus a certain respondent will say yes to a bid if $V_i^1[c_i(A_i^1), y - bid] > V_i^0[c_i(A_i^0), y]$, where A_i^t entirely describes the good, including its continuous dimension.

Until now we have presented and discussed a deterministic model of consumer behaviour. The next step is to make such a model operational. There are two main issues involved: one is the assumption regarding the functional form of the utility function and the other is to introduce a component into the utility function to capture unobservable behaviour. In principle, these issues are linked, since the form of the utility function determines the relation between the probability distribution of the disturbances and the probability distribution of the indirect utility function.

3. The Econometric Model

Stated behaviour surveys sometimes reveal preference structures that may seem inconsistent with the deterministic model. It is assumed that these inconsistencies stem from observational deficiencies arising from unobservable components such as characteristics of the individual or non-included attributes of the alternatives in the experiment, measurement error and/or heterogeneity of preferences (Hanemann and Kanninen, 1999). In order to allow for these effects, the Random Utility approach (McFadden, 1974) is used to link the deterministic model with a statistical model of human behaviour. A random disturbance with a specified probability distribution, ε, is introduced into the model, and an individual will choose profile *j* if and only if:

$$V_j(A_j, y - p_j c_j, \varepsilon_j) > V_i(A_i, y - p_i c_i, \varepsilon_i);\ \forall i \neq j \qquad (7)$$

In terms of probabilities, we write:

$$P\{choose\ j\} = P\{V_j(A_j, y - p_j c_j, \varepsilon_j) > V_i(A_i, y - p_i c_i, \varepsilon_i); \forall i \neq j\} \qquad (8)$$

The exact specification of the econometric model depends on how the random elements, ε, enter the conditional indirect utility function and the distributional assumption. Let us divide the task into two parts: (i) specification of the utility function, and (ii) specification of the probability distribution of the error term.

3.1 Specification of the Utility Function

The most common assumption is that the error term enters the utility function as an additive term. This assumption, although restrictive, greatly simplifies the computation of the results and the estimation of welfare measures. In section 3.2 we present a random parameter model, which is an example of a model with the stochastic component entering the utility function via the slope coefficients, i.e. non-additively (Hanemann, 1999).

Under an additive formulation the probability of choosing alternative *j* can be written as:

$$P\{\text{choose j}\} = P\{V_j(A_j, y - p_j c_j) + \varepsilon_j > V_i(A_i, y - p_i c_i) + \varepsilon_i; \forall i \neq j\} \qquad (9)$$

In order to specify a utility function, we need to specify the functional form for $V(...)$ and to select the relevant attributes (A_i) that determine the utility derived from each alternative. These attributes should then be included in the choice experiment.

When choosing the functional form, there is a trade-off between the benefits of assuming a less restrictive formulation and the complications that arise from doing so. This is especially relevant for the way income enters the utility function. A simpler functional form (e.g. linear in income) makes estimation of the parameters and calculation of welfare effects easier, but the estimates are based on restrictive assumptions (Ben-Akiva and Lerman, 1985). Most often researchers have been inclined to use a simpler linear in the parameters utility function. Since the need for simple functional forms is linked to the estimation of welfare measures, we will postpone the discussion to section 6, where we investigate in more detail the implications of the chosen functional form on the calculation of exact welfare estimates.

Regarding the selection of attributes it is important to be aware that the collected data come from a specific design based on a priori assumptions regarding estimable interaction effects between attributes. Once the experiment has been conducted we are restricted to testing for only those effects that were considered in the design. This shows the importance of focus groups and pilot studies when constructing the experiment.

3.2 Specification of the Probability Distribution of the Error Term

The most common model used in applied work has been the Multinomial Logit (MNL) model. This model relies on restrictive assumptions, and its popularity rests on its simplicity of estimation. We begin by introducing the MNL model and discussing its limitations, and then we introduce less restrictive models. Suppose that the choice experiment consists of M choice sets, where each choice set, S_m, consists of K_m

alternatives, such that $S_m = \{A_{1m},...., A_{Km}\}$, where A_i is a vector of attributes. We can then write the choice probability for alternative *j* from a choice set S_m as

$$P\{j \mid S_m\} = P\{V_j(A_{jm}, y - p_j c_j) + \varepsilon_j > V_i(A_{im}, y - p_i c_i) + \varepsilon_i; \forall i \in S_m\} = \quad (10)$$

$$= P\{V_j(...) + \varepsilon_j - V_i(...) > \varepsilon_i; \forall i \in S_m\}.$$

We can then express this choice probability in terms of the joint cumulative density function of the error term as:

$$P(j \mid S_m) = CDF_{\varepsilon|S_m}(V_j + \varepsilon_j - V_1, V_j + \varepsilon_j - V_2, \ldots, V_j + \varepsilon_j - V_n). \quad (10')$$

The MNL model assumes that the random components are independently and identically distributed with an extreme value type I distribution (Gumbel). This distribution is characterized by a scale parameter μ and location parameter δ.[2] The scale parameter is related to the variance of the distribution such that $\text{var}_\varepsilon = \pi^2 / 6\mu^2$. If we assume that the random components are extreme value distributed, the choice probability in (10) can be written as:

$$\mathrm{P(j \mid S_m, \beta)} = \frac{\exp(\mu V_j)}{\sum_{i \in S_m} \exp(\mu V_i)}. \quad (11)$$

In principle, the size of the scale parameter is irrelevant when it comes to the choice probability of a certain alternative (Ben-Akiva and Lerman, 1985), but by looking at equation (11) it is clear that the true parameters are confounded with the scale parameter. Moreover, it is not possible to identify this parameter from the data. For example, if the scale is doubled, the estimated parameters in the linear specification will adjust to double their previous values.[3] The presence of a scale parameter raises several

[2] In practice, the distribution chosen is the standard Gumbel distribution with $\mu = 1$ and $\delta = 0$.

[3] In a linear specification, $\beta^{estimated} = \mu\beta^{true}$, and $\beta^{estimated}$ will adjust to changes in μ. The issue of the scale parameter is not specific to multinomial models and Gumbel distributions. For the case of probit

issues for the analysis of the estimations. First consider the variance of the error term: $\text{var}_\varepsilon = \pi^2 / 6\mu^2$. An increase in the scale reduces the variance; therefore high fit models have larger scales. The two extreme cases are $\mu \to 0$ where, in a binary model, the choice probabilities become ½, and $\mu \to \infty$ where the model becomes completely deterministic (Ben-Akiva and Lerman, 1985). Second, the impact of the scale parameter on the estimated coefficients imposes restrictions on their interpretation. All parameters within an estimated model have the same scale and therefore it is valid to compare their signs and relative sizes. On the other hand, it is not possible to directly compare parameters from different models as the scale parameter and the true parameters are confounded. Nevertheless, it is possible to compare estimated parameters from two different data sets, or to combine data sets (for example stated and revealed preference data). Swait and Louviere (1993) show how to estimate the ratio of scale parameters for two different data sets. This procedure can then be used to compare different models or to pool data from different sources (see e.g. Adamowicz et al., 1994; Ben-Akiva and Morikawa, 1990).

There are two problems with the MNL specification: (i) the alternatives are independent and (ii) there is a limitation in modelling variation in taste among respondents. The first problem arises because of the IID assumption (constant variance), which results in the independence of irrelevant alternatives (IIA) property. This property states that the ratio of choice probabilities between two alternatives in a choice set is unaffected by changes in that choice set. If this assumption is violated the MNL should not be used. One type of model that relaxes the homoskedasticity assumption of the MNL model is the nested MNL model. In this model the alternatives are placed in subgroups, and the variance is allowed to differ between the subgroups but it is assumed to be the same within each group. An alternative specification is to assume that error terms are independently, but non-identically, distributed type I extreme value, with scale parameter μ_i (Bhat, 1995). This would allow for different cross elasticities among all pairs of alternatives, i.e. relaxing the IIA restriction. Furthermore, we could also model heterogeneity in the covariance among nested alternatives (Bhat, 1997).

---

models, the scale parameter of the normal distribution is $1/\sigma$. Everything we say here about the scale parameter of the Gumbel distribution applies to nested MNL and probit models as well.

The second problem arises when there is taste variation among respondents due to observed and/or unobserved heterogeneity. Observed heterogeneity can be incorporated into the systematic part of the model by allowing for interaction between socio-economic characteristics and attributes of the alternatives or constant terms. However, the MNL model can also be generalized to a so-called mixed MNL model in order to further account for unobserved heterogeneity. In order to illustrate this type of model, let us write the utility function of alternative *j* for individual *q* as:

$$U_{jq} = \beta x_{jq} + \varepsilon_{jq} = \overline{\beta} x_{jq} + \tilde{\beta}_q x_{jq} + \varepsilon_{jq}. \tag{12}$$

Thus, each individual's coefficient vector β is the sum of the population mean $\overline{\beta}$ and individual deviation $\tilde{\beta}_q$. The stochastic part of utility, $\tilde{\beta}_q x_{jq} + \varepsilon_{jq}$, is correlated among alternatives, which means that the model does not exhibit the IIA property. If the error terms are IID standard normal we have a random parameter multinomial probit model. If instead the error terms are IID type I extreme value, we have a random parameter logit model.

Let tastes, β, vary in the population with a distribution with density $f(\beta \mid \theta)$, where θ is a vector of the true parameters of the taste distribution. The unconditional probability of alternative *j* for individual *q* can then be expressed as the integral of the conditional probability in (11) over all values of β:

$$P_q(j \mid \theta) = \int P_q(j \mid \beta) f(\beta \mid \theta) d\beta = \int \frac{\exp(\mu \beta x_{jq})}{\sum_{i=1}^{K_m} \exp(\mu \beta x_{iq})} f(\beta \mid \theta) d\beta. \tag{13}$$

In general the integrals in equation (13) cannot be evaluated analytically, and we have to rely on simulation methods for the probabilities (see e.g. Brownstone and Train, 1999).

When estimating these types of models we have to assume a distribution for each of the random coefficients. It may seem natural to assume a normal distribution. However, for many of the attributes it may be reasonable to expect that all respondents have the same sign for their coefficients. In this case it may be more sensible to assume a log-normal distribution. For example, if we assume that the price coefficient is log-normally distributed, we ensure that all individuals have a non-positive price coefficient.

In most choice experiments, respondents make repeated choices, and we assume that the preferences are stable over the experiment. Consequently, the utility coefficients are allowed to vary among individuals but they are constant among the choice situations for each individual (Revelt and Train, 1998; Train, 1998). It is also possible to let the coefficients for the individual vary over time; in this case among the choice situations in the survey. This type of specification would be valid if we suspect fatigue or learning effects in the survey.

McFadden and Train (2000) show that under some mild regularity conditions any discrete choice model derived from random utility maximization has choice probabilities that can be approximated by a mixed MNL model. This is an interesting result because mixed MNL models can then be used to approximate difficult parametric random utility models, such as the multinomial probit model, by taking the distributions underlying these models as the parameter distributions.

4. Design of a Choice Experiment

There are four steps involved in the design of a choice experiment: (i) definition of attributes, attribute levels and customisation, (ii) experimental design, (iii) experimental context and questionnaire development and (iv) choice of sample and sampling strategy. These four steps should be seen as an integrated process with feedback. The development of the final design involves repeatedly conducting the steps described here, and incorporating new information as it comes along. In this section, we focus on the experimental design and the context of the experiment, and only briefly discuss the other issues.

4.1 Definition of Attributes and Levels

The first step in the development of a choice experiment is to conduct a series of focus group studies aimed at selecting the relevant attributes. A starting point involves studying the attributes and attribute levels used in previous studies and their importance in the choice decisions. Additionally, the selection of attributes should be guided by the attributes that are expected to affect respondents' choices, as well as those attributes that are policy relevant. This information forms the base for which attributes and relevant attribute levels to include in the first round of focus group studies.

The task in a focus group is to determine the number of attributes and attribute levels, and the actual values of the attributes. As a first step, the focus group studies should provide information about credible minimum and maximum attribute levels. Additionally, it is important to identify any possible interaction effect between the attributes. If we want to calculate welfare measures, it is necessary to include a monetary attribute such as a price or a cost. In such a case, the focus group studies will indicate the best way to present a monetary attribute. Credibility plays a crucial role and the researcher must ensure that the attributes selected and their levels can be combined in a credible manner. Hence, proper restrictions may have to be imposed (see e.g. Layton and Brown, 1998).

Customisation is an issue in the selection of attributes and their levels. It is an attempt to make the choice alternatives more realistic by relating them to actual levels. If possible an alternative with the attribute levels describing today's situation should be included which would then relate the other alternatives to the current situation. An alternative is to directly relate some of the attributes to the actual level. For example, the levels for visibility could be set 15% higher and 15% lower than today's level (Bradley, 1988).

The focus group sessions should shed some light on the best way to introduce and explain the task of making a succession of choices from a series of choice sets. As Layton and Brown (1998) explain, choosing repeatedly is not necessarily a behavior that could be regarded as obvious for all goods. When it comes to recreation, for example, it is clear that choosing a site in a choice set does not preclude choosing another site given different circumstances. However, in the case of public goods, such repeated choices might require further justification in the experiment.

A general problem with applying a choice experiment to an environmental good or to an improvement in health status is that respondents are not necessarily familiar with the attributes presented. Furthermore, the complexity of a choice experiment in terms of the number of choice sets and/or the number of attributes in each choice set may affect the quality of the responses; this will be discussed in Section 4.3. Basically, there is a trade-off between the complexity of the choice experiment and the quality of the responses. The complexity of a choice experiment can be investigated by using verbal protocols, i.e. by asking the individual to read the survey out loud and/or to think aloud when responding; this approach has been used in CVM surveys (e.g. Schkade and Payne, 1993), thereby identifying sections that attract the readers' attention and testing the understanding of the experiment

4.2 Experimental Design

Experimental design is concerned with how to create the choice sets in an efficient way, i.e. how to combine attribute levels into profiles of alternatives and profiles into choice sets. The standard approach in marketing, transport and health economics has been to use so-called orthogonal designs, where the variations of the attributes of the alternatives are uncorrelated in all choice sets. Recently, there has been a development of optimal experimental designs for choice experiments based on multinomial logit models. These optimal design techniques are important tools in the development of a choice experiment, but there are other more practical aspects to consider. We briefly introduce optimal design techniques for choice experiments and conclude by discussing some of the limitations of statistical optimality in empirical applications.

A design is developed in two steps: (i) obtaining the optimal combinations of attributes and attribute levels to be included in the experiment and (ii) combining those profiles into choice sets. A starting point is a full factorial design, which is a design that contains all possible combinations of the attribute levels that characterize the different alternatives. A full factorial design is, in general, very large and not tractable in a choice experiment. Therefore we need to choose a subset of all possible combinations, while following some criteria for optimality and then construct the choice sets. In choice experiments, design techniques used for linear models have been popular. Orthogonality in particular has often been used as the principle part of an efficient design. More recently researchers in marketing have developed design techniques based on the D-optimal criteria for non-linear models in a choice experiment context. D-optimality is related to the covariance matrix of the *K*-parameters, defined as

$$D - efficiency = [|\Omega|^{1/K}]^{-1} . \tag{14}$$

Huber and Zwerina (1996) identify four principles for an efficient design of a choice experiment based on a non-linear model: (i) orthogonality, (ii) level balance, (iii) minimal overlap and (iv) utility balance. Level balance requires that the levels of each attribute occur with equal frequency in the design. A design has minimal overlap when an attribute level does not repeat itself in a choice set. Finally, utility balance requires that the utility of each alternative within a choice set is equal. The last property is important since the larger the difference in utility between the alternatives the less information is extracted from that specific choice set. At the same time, this principle is difficult to satisfy since it requires prior knowledge about the true distribution of the parameters. The theory of optimal design for choice experiments is related to optimal design of the bid vector in a CVM survey. The optimal design in a CVM survey depends on the assumption regarding the distribution of WTP (see e.g. Duffield and Patterson, 1991; Kanninen, 1993).

Several design strategies explore some or all of the requirements for an efficient design of a choice experiment. Kuhfeld et al. (1994) use a computerized search algorithm to minimize the D-error in order to construct an efficient, but not necessarily orthogonal, linear design. However, these designs do not rely on any prior information about the utility parameters and hence do not satisfy utility balance. Zwerina et al. (1996) adapt the search algorithm of Kuhfeld et al. (1994) to the four principles for efficient choice designs as described in Huber and Zwerina (1996).[4] In order to illustrate their design approach it is necessary to return to the MNL model. McFadden (1974) showed that the maximum likelihood estimator for the conditional logit model is consistent and asymptotically normally distributed with the mean equal to β and a covariance matrix given by:

$$\Omega = (\mathbf{Z}'\mathbf{P}\mathbf{Z})^{-1} = [\sum_{n=1}^{N}\sum_{j=1}^{J_m}\mathbf{z}'_{jn} P_{jn}\mathbf{z}_{jn}]^{-1} , \tag{15}$$

$$\text{where } \mathbf{z}_{jn} = \mathbf{x}_{jn} - \sum_{i=1}^{J_n}\mathbf{x}_{in}P_{in} .$$

[4] The SAS code is available at ftp://ftp.sas.com/techsup/download/technote/ts643/.

This covariance matrix, which is the main component in the D-optimal criteria, depends on the true parameters in the utility function, since the choice probabilities, P_{in}, depend on these parameters.[5] Consequently, an optimal design of a choice experiment depends, as in the case of the optimal design of bid values in a CV survey, on the value of the true parameters of the utility function. Adapting the approach of Zwerina et al. (1996) consequently requires prior information about the parameters. Carlsson and Martinsson (2000) discuss strategies for obtaining this information, which includes results from other studies, expert judgments, pilot studies and sequential designs strategies. Kanninen (1993) discusses a sequential design approach for closed-ended CVM surveys and she finds that this approach improves the efficiency of the design. A similar strategy can be used in designing choice experiments. The response data from the pilot studies and the actual choice experiment can be used to estimate the value of the parameters. The design can then be update during the experiment depending on the results of the estimated parameters. The results from these estimations may not only require a new design, but changes in the attribute levels as well. There are other simpler design strategies that do not directly require information about the parameters. However, in all cases, some information about the shape of the utility function is needed in order to make sure that the individuals will make trade-offs between attributes. The only choice experiment in environmental valuation that has adopted a D-optimal design strategy is Carlsson and Martinsson (2001). In a health economic application by Johnson et al. (2000) a design partly based on D-optimal criteria is applied.

Kanninen (2001) presents a more general approach to optimal design than Zwerina et al. (1996). In her design, the selection of the number of attribute levels is also a part of the optimal design problem. Kanninen (2001) shows that in a D-optimal design each attribute should only have two levels, even in the case of a multinomial choice experiment, and that the levels should be set at the two extreme points of the distribution of each attribute.[6] Furthermore, Kanninen (2001) shows that for a given number of attributes and alternatives, the D-optimal design results in certain response probabilities. This means that updating the optimal design is simpler than updating the design presented in Zwerina et al. (1996). In order to achieve the desired response

[5] This is an important difference from the design of linear models where the covariance matrix is proportional to the information matrix, i.e. $\Omega = (\mathbf{X}'\mathbf{X})^{-1}\sigma^2$.

probabilities the observed response probabilities from previous applications have to be calculated, and a balancing attribute is then included. This type of updating was adopted by Steffens et al. (2000) in a choice experiment on bird watching. They found that the updating improved the efficiency of the estimates.

There are several problems with these more advanced design strategies due to their complexity, and it is not clear whether the advantages of being more statistically efficient outweigh the problems. The first problem is obtaining information about the parameter values. Although some information about the coefficients is required for other design strategies as well, more elaborate designs based on utility balance are more sensitive to the quality of information used, and incorrect information on the parameters may bias the final estimates. Empirically, utility balance makes the choice harder for the respondents, since they have to choose from alternatives that are very close in terms of utility. This might result in a random choice. The second problem is that the designs presented here are based on a conditional logit model where, for example, homogeneous preferences are assumed. Violation of this assumption may bias the estimates. The third problem is the credibility of different combinations of attributes. If the correlation between attributes is ignored, the choice sets may not be credible to the respondent (Johnson et al., 2000, and Layton and Brown, 1998). In this case it may be optimal to remove such combinations although it would be statistically efficient to include them.

4.3 Experimental Context, Test of Validity and Questionnaire Development

In the previous section, we addressed optimal design of a choice experiment from a statistical perspective. However, in empirical applications there may be other issues to consider in order to extract the maximum amount of information from the respondents.

Task complexity is determined by factors such as the number of choice sets presented to the individual, the number of alternatives in each choice set, the number of attributes describing those alternatives and the correlation between attributes for each alternative (Swait and Adamowicz, 1996). Most authors find that task complexity affects the decisions (Adamowicz et. al., 1998a; Bradley, 1988). Mazotta and Opaluch (1995) and Swait and Adamowicz (1996) analyze task complexity by assuming it affects the variance term of the model. The results of both papers indicate that task

[6] The design is derived under the assumption that all attributes are quantitative variables.

complexity does in fact affect the variance, i.e. an increased complexity increases the noise associated with the choices. Task complexity can also arise when the amount of effort demanded when choosing the preferred alternative in a choice set may be so high that it exceeds the ability of the respondents to select their preferred option. The number of attributes in a choice experiment is studied by Mazotta and Opaluch (1995) and they find that including more than 4 to 5 attributes in a choice set may lead to a severe detriment to the quality of the data collected due to the task complexity.

In complex cases, respondents may simply answer carelessly or use some simplified lexicographic decision rule. This could also arise if the levels of the attributes are not sufficiently differentiated to ensure trade-offs. Another possibility is 'yea' saying or 'nay' saying, where the respondent, for example, always opt for the most environmentally friendly alternative. Finally, lexicographic orderings may be an indication of strategic behaviour of the respondent. In practice, it is difficult to separate these cases from preferences that are genuinely lexicographic, in which case the respondents have a ranking of the attributes, but the choice of an alternative is based solely on the level of their most important attribute. Genuine lexicographic preferences in a choice experiment are not a problem, although they provide us with little information in the analysis compared to the other respondents. However, if a respondent chooses to use a lexicographic strategy because of its simplicity, systematic errors are introduced, which may bias the results. One strategy for distinguishing between different types of lexicographic behaviour is to use debriefing questions, where respondents are asked to give reasons why they, for example, focused on only one or two of the attributes in the choice experiment. However, in a thoroughly pre-tested choice experiment using focus groups and pre tests, these problems should have been detected and corrected.

An issue related to task complexity in is the stability of preferences. In choice experiments the utility function of each individual is assumed to be stable throughout the experiment. The complexity of the exercise might cause violations of this assumption, arising from learning and fatigue effects. Johnson et al. (2000) test for stability by comparing responses to the same choice sets included both at the beginning and at the end of the experiment. They find a strong indication of instability of preferences. However, there is a potential problem of confounding effects of the sequencing of the choice sets and the stability of the preferences. An alternative approach, without the confounding effect, is applied in Carlsson and Martinsson (2001) in a choice experiment on donations to environmental projects. In their exercise, half of the respondents receive the choice sets in the order {A,B} and the other half in the order {B,A}. A test for stability is then performed by comparing the preferences obtained for the choices in subset A, when it was given in the sequence {A,B}, with the preferences obtained when the choices in subset A were given in the sequence {B,A}. This can then be formally tested in a likelihood ratio test between the pooled model of the choices in subset A and the separate groups. A similar test can be performed for subset B. By using this method Carlsson and Martinsson (2001) find only a minor problem with instability of preferences. Layton and Brown (2000) conduct a similar test of stability in a choice experiment on policies for mitigating impacts of global climate change; they did not reject the hypothesis of stable preferences. Bryan et al. (2000) compare responses in the same way, but with the objective of testing for reliability, and find that 57 percent of the respondents did not change their responses when given the same choice set in a two-part choice experiment. Furthermore, in an identical follow-up experiment two weeks after the original experiment, 54 percent of the respondents made the same choices on at least eleven out of twelve choice situations.

Another issue to consider in the development of the questionnaire is whether or not to include a base case scenario or an opt-out alternative. This is particularly important if the purpose of the experiment is to calculate welfare measures. If we do not allow individuals to opt for a status quo alternative, this may distort the welfare measure for non-marginal changes. This decision should, however, be guided by whether or not the current situation and/or non-participation is a relevant alternative. A non-participation decision can be econometrically analysed by e.g. a nested logit model with participants and non-participants in different branches (see e.g. Blamey et al., 2000). A simpler alternative is to model non-participation as an alternative where the levels of the attributes are set to the current attribute levels. Another issue is whether to present the alternatives in the choice sets in a generic (alternatives A, B, C) or alternative specific form (national park, protected area, beach). Blamey et al. (2000) discuss advantages of these two approaches and compare them in an empirical study. An advantage of using alternative specific labels is familiarity with the context and hence the cognitive burden is reduced. However, the risk is that the respondent may not consider trade-offs between attributes. This approach is preferred when the emphasis is on valuation of the labelled alternatives. An advantage of the generic model is that the respondent is less inclined to only consider the label and thereby focus more on the attributes. Therefore, this approach is preferred when the emphasis is on the marginal rates of substitution between attributes.

In the random utility model, unobservable effects are modelled by an error term and, in general, we assume that respondents have rational, stable, transitive and monotonic preferences. Also, we assume they do not have any problems in completing a choice experiment, and that there are no systematic errors, such as respondents getting tired or changing their preferences as they acquire experience with the experiment, i.e. learning effects. Internal tests of validity are designed to check these standard assumptions. These tests can be directly incorporated into the design of an experiment. There have been several validity tests of choice experiments in the marketing and transport literature, for example Ben-Akiva et al. (1992) and Leigh et al. (1984). The evidence from a large proportion of studies is that choice experiments generally pass these tests of validity. However, it is not obvious that these results carry over into choice experiments done in an environmental or health economic context. The reason is that these non-market goods in many respects differ from, for example, transportation, which is a good that most respondents are familiar with. It is therefore of importance to test the validity of choice experiments in the context of valuation of general non-marketed goods. Since there are few applications of choice experiments in valuation, few tests of internal validity have been performed.

In order to test for transitive preferences, we have to construct such a test. For example, in the case of a pair-wise choice experiment we have to include three specific choice sets: (1) Alt. 1 versus Alt. 2, (2) Alt. 2 vs. Alt. 3, and (3) Alt. 1 vs. Alt. 3. For example if the respondent chooses Alt. 1 in the first choice set and Alt. 2 in the second choice set, then Alt. 1 must be chosen in the third choice if the respondent has transitive preferences. Carlsson and Martinsson (2001) conduct tests of transitivity and they do not find any strong indications of violations. Internal tests of monotonicity can also be implemented in a choice experiment and in a sense tests of monotonicity are already built-in in a choice experiment as the level of an attribute changes in an experiment. Comparing the expected sign to the actual sign and significance of the coefficient can be seen as a weak test monotonicity. Johnson et al. (2000) discuss a simple test of dominated pair, which simply tests if a respondent chooses a dominated alternative.

4.4 Sample and Sampling Strategy

The choice of survey population obviously depends on the objective of the survey. Given the survey population, a sampling strategy has to be determined. Possible strategies include a simple random sample, a stratified random sample or a choice-based sample. A simple random sample is generally a reasonable choice. One reason for choosing a more specific sampling method may be the existence of a relatively small but important sub-group which is of particular interest to the study. Another reason may be to increase the precision of the estimates for a particular sub-group. In practice the selection of sample strategy and sample size is also largely dependent on the budget available for the survey.

Louviere et al. (2000) provide a formula to calculate the minimum sample size. The size of the sample, n, is determined by the desired level of accuracy of the estimated probabilities, $\hat{p}$. Let p be a true proportion of the relevant population, a is the percentage of deviation between $\hat{p}$ and p that can be accepted and α is the confidence level of the estimations such that: $\Pr(|\hat{p}-p| \leq ap) \geq \alpha$ for a given n. Given this, the minimum sample size is defined as:

$$n \geq \frac{1-p}{pa^2}\Phi^{-1}\left(\frac{1+\alpha}{2}\right). \qquad (16)$$

Note that n refers to the size of the sample and not the number of observations. Since each individual makes r succession of choices in a choice experiment, the number of observations will be much larger (a sample of 500 individuals answering 8 choice sets each will result in 4000 observations). One of the advantages of choice experiments is that the amount of information extracted from a given sample size is much larger than, for example, using referendum based methods and, hence, the efficiency of the estimates is improved. The formula above is only valid for a simple random sample and with independency between the choices. For a more detailed look at this issue see e.g. Ben-Akiva and Lerman (1985). In a health economic context, the availability of potential respondents can in certain cases be limited and hence the equation above can be used to solve for a, i.e. the percentage deviation between $\hat{p}$ and p that we must accept given the sample size used.

5. Elicitation of preferences in choice experiments

There has been an extensive discussion about the possibility of eliciting preferences for non-market goods in hypothetical surveys. While the discussion has focused on CVM (see e.g. Diamond and Hausman, 1994 and Hanemann, 1994) most of the results are valid for choice experiments as well. We believe that there are particular problems with measuring so-called non-use values in hypothetical surveys. We do not take the position that non-use values should not be measured, but rather that there are some inherent problems with measuring these values. The reason for this is that non-use values are largely motivated by "purchase of moral satisfaction" (Kahneman and Knetsch, 1992) and "warm glow" (Andreoni, 1989), and that they often involve an "important perceived ethical dimension" (Johansson-Stenman and Svedsäter, 2001). We are not questioning these values per se; on the contrary, they may even be important shares of total value. The problem is that the cost of acquiring a "warm glow" or a satisfaction of acting ethical is much lower in a hypothetical survey situation than in an actual situation. This leaves us in a difficult position, since stated preference methods are essentially the only methods available for measuring non-use values. However, there are reasons to believe that choice experiments may be less prone to trigger this type of behaviour than CVM surveys. The reason for this is that in a choice experiment individuals have to make trade-offs between several attributes, several of which may contain non-use values.

Another issue involves incentives for truthfully revealing preferences in hypothetical surveys. Carson et al. (1999) argue that given a consequential survey a binary discrete choice is incentive compatible for the cases of (i) a new public goods with coercive payments, (ii) the choice between two public goods and (iii) a change in an existing private or quasi-public good. A consequential survey is defined as one that is perceived by the respondent as something that may potentially influence agency decisions, as well as one where the respondent cares about the outcome. The problem arises when the individual faces not one but a sequence of binary choices. Let us assume we are dealing with a public good, i.e. everybody will enjoy the same quantity and composition of the good after the government has decided its provisions. The respondents could then perceive the sequence of binary choices as a voting agenda, and, if they expect one of their less preferred outcomes to be chosen, they would have an incentive to misrepresent their true preferences. The same type of problem arises with multinomial choices. If only one alternative is to be chosen, the multinomial choice is reduced to a binary choice between the two alternatives that the respondent believes are most likely to be chosen, even if these two alternatives are not the most preferred ones. The problem with these incentives is that the preference profile constructed from the survey is not a reflection of the true preferences, but rather a reflection of strategic behaviour. The choice experiment would then be flawed and any welfare estimate would not be reliable. This issue clearly demands attention from researchers, although we believe that the importance of these results should not be overemphasized.

It is in general more difficult to behave strategically in a choice experiment, when compared to a CVM survey. In a CVM survey the respondent "only" has to consider a single change in a project involving a certain payment. A typical choice experiment consists of two to four alternatives, where each alternative is described by at least three or four attributes. The selection of all attributes is done under the premise that they are relevant determinants of choice behaviours of individuals and the levels are set such that they imply meaningful changes in utility. Furthermore, there is, generally, no clearly identifiable agenda in a sequence of choices, where almost all levels of attributes change from one choice set to another. Thus, it is more difficult for a respondent to behave strategically in a choice experiment. First they need to create an expectation regarding the values of each of the alternatives in the choice set. Based on this expectation they need to calculate the decision weights for each pair-wise decision. Of particular importance is the fact that most choice experiments, as well as CVM surveys, deal with situations that are not familiar to the respondent. The fact that there are no markets for some of the evaluated goods means that there is limited, if any, information about the preferences of other individuals. There are seldom any opinion polls, prices or other types of information that the respondent can use. Thus in general the respondent is in an unfamiliar situation and with limited prior information on the preferences of others.

The assumption that each respondent has perfect information regarding the preferences of other respondents is unrealistic and the question is how uncertainty affects the incentives for truthful revelation. Here we illustrate this with the model of Gutowski and Georges (1993). Each respondent has a subjective value of each of three alternatives, a_1, a_2 and a_3. A respondent with the preference ordering $a_1 \succ a_2 \succ a_3$, where the subjective value of the most preferred alternative, $v(a_1)$, is equal to one, and the subjective value of the least preferred alternative, $v(a_3)$, is equal to zero. The subjective value of a_2, $v(a_2)$, is uniformly distributed between zero and one. Any particular respondent does not have perfect information regarding other respondents' preference orderings, but is assumed to form subjective beliefs regarding the chances of various scenarios. These are represented by decision weights that measure the extent to which each of the pair wise competitions are incorporated into a respondent's choice among admissible strategies. There are three possible pair-wise competitions, and consequently three decision weights, w_{12}, w_{13}, and w_{23}, where $w_{12} + w_{13} + w_{23} = 1$. The decision weight w_{ij} is the weight associated with the competition between a_i and a_j, and it reflects the expected probability from the respondents perspective, that the outcome of the survey is defined only by the competition between a_i and a_j. We assume that the value of a strategy is the weighted average of the possible outcomes of that strategy. Finally, we assume that the respondent is only interested in the survey if the response is critical in determining the alternative. Let us now analyze the incentives for a respondent with the preference ordering $a_1 \succ a_2 \succ a_3$. Gutowski and Georges (1993) show that the only admissible strategies are to choose a_1 or a_2, i.e. it can never be optimal to choose the least preferred alternative. Consequently, with three alternatives the respondent has to make a choice between the most preferred or the second most preferred alternative. Setting the value distributions of choosing a_1 and a_2 equal, we can find the critical value of $v(a_2)$ at which the respondent is indifferent between a strategic and a non-strategic behaviour, defined as

$$v(\overline{a}_2) = \frac{2w_{12} + w_{13}}{2w_{12} + w_{23}}. \tag{17}$$

If the true subjective value of outcome a_2 is larger than the critical value then the respondent acts strategically and chooses alternative a_2, although alternative a_1 is the most preferred alternative. A number of interesting conclusions can be drawn from this expression: (i) A respondent will always choose truthfully if $w_{13} \geq w_{23}$, since $v(a_2) < 1$. This means that if the perceived competition between a_1 and a_3 is larger than that between a_2 and a_3, the respondent will choose truthfully. Furthermore, this implies that in the case of equal decision weights the respondent will always choose truthfully. The latter case would perhaps be likely when the respondent does not have much information regarding other individuals' preferences, and therefore puts equal decision weights on all pair-wise competitions. (ii) The probability of acting truthfully is decreasing in $v(a_2)$. This means that if the utility of the two alternatives is close, then there is a higher probability of strategic behaviour. (iii) A respondent will in general only choose strategically if w_{23} is considerably larger than both w_{12} and w_{13}. Four straightforward and important conclusions can be drawn from above. First, introducing imperfect information does not ensure that the degree of strategic behaviour is reduced. It may well be the case that respondents form such expectations so that they act strategically even if they would not have done so with perfect information. Second, using a generic (no labels) presentation of the alternatives instead of an alternative specific (labels) form probably reduces the risk of strategic behaviour, since it increases the complexity of forming expectations regarding other respondents' preferences. Third, it is generally advisable to explicitly introduce uncertainty into the choice experiment. This can be done by saying that there is uncertainty regarding individuals' preferences for the alternatives and the attributes. We believe that this strategy should be used in general with choice experiments. It is important to convince the respondent of the importance of he/she carefully answering the questionnaire, and that his/her choice can affect the outcome. Fourth, it is not clear whether differences in utility between alternatives in a choice set should be small or large. If the utility difference is small, then it is more difficult for the respondents to form expectations regarding how other respondents will choose, thereby making it more difficult to act strategically. At the same time, if the alternatives are close in utility the cost of acting strategically and being wrong is not that high compared to choosing sincerely, thereby increasing the probability of choosing strategically.

The empirical counterpart to the above discussion is tests of external validity, i.e. comparisons of actual and hypothetical behaviour. In transport economics, validity tests are either comparative studies with both hypothetical choice/ranking data and revealed preference data (e.g. Benjamin and Sen, 1982), or comparisons of predicted market shares from hypothetical choice/ranking studies with observed market shares (e.g. Wardman, 1988). The evidence from a large proportion of studies is that choice experiments generally pass external tests of validity. However, as we have discussed it is not obvious that these results carry over to hypothetical experiments on non-market goods. Carson et al. (1996) perform a meta-analysis, comparing results of CVM studies with revealed preference studies, and they find that the CVM estimates are slightly lower than their revealed preferences counterparts. However, several other experimental tests of the validity of CVM show that individuals overstate their WTP in hypothetical settings (see e.g. Cummings et al., 1995; Frykblom, 1997). We are only aware of three external validity tests for environmental goods. Carlsson and Martinsson (2001) conduct a classroom experiment consisting of both a hypothetical and an actual choice experiment, and they cannot reject the hypothesis of external validity. Johansson-Stenman and Svedsäter (2001) conduct a similar experiment as Carlsson and Martinsson, but allows for between-subjects tests. They find a significant difference between actual and hypothetical behaviour, arguing that the difference in results is due to their between-subjects test. Cameron et al. (1999) compare six different hypothetical choice formats with actual purchase behavior. They assume an underlying indirect utility function, which allows the data from the choice formats to be used independently or pooled with heteroskedasticity across the formats. They cannot reject the hypothesis of the same indirect utility function across the question formats: actual behavior, closed-ended CV (phone survey), closed-ended CV (mail survey) and a pair-wise choice experiment.

6. Welfare Effects

The main purpose of a choice experiment is to estimate the welfare effects of changes in the attributes. In order to obtain these, researchers have generally assumed a simple functional form of the utility function by imposing a constant marginal utility of income. We focus on purely discrete choices; this means that in some cases welfare measures have to be interpreted with care in some cases. For example in the case of a site choice experiment, the welfare measures are per trip or per week, depending on what has been defined in the survey.

Let us assume the following utility function:

$$u = h(A) + \gamma(Q,z)z + \varepsilon, \tag{18}$$

where the function $h(A)$ captures the effect of the different attributes on utility, Q is a vector of personal characteristics and z is a composite bundle. This is a flexible specification of the marginal utility of income as it may vary by both the level of income and the personal characteristics of the individual. However, let us begin with the common case of constant marginal utility of income and independence of personal characteristics. For such a utility function, the ordinary and compensated demand functions coincide. Given this functional form and the assumption of weak complementarity, we can write the conditional indirect utility function for the purely discrete choice as:

$$V_j(A_j, p_j, y, \varepsilon) = h_j(A_j) + \bar{\gamma}(y - p_j c_j) + \varepsilon. \tag{19}$$

Furthermore, we can write the probability that alternative j is preferred as:

$$\begin{aligned} P\{j\} = P\{h_j(A_j) + \bar{\gamma}(y - p_j c_j) + \varepsilon_j > h_i(A_i) + \bar{\gamma}(y - p_i c_i) + \varepsilon_i ; \forall i \neq j\} = \\ P\{h_j(A_j) - \bar{\gamma} p_j c_j + \varepsilon_j > h_i(A_i) - \bar{\gamma} p_i c_i + \varepsilon_i ; \forall i \neq j\}. \end{aligned} \tag{20}$$

Equation (20) shows that income does not affect the probability of choosing a certain alternative under the current assumptions and hence the welfare measures will have no income effects. Thus, we can express the unconditional indirect utility function as:

$$v(A,p,y,s) = \bar{\gamma} y + \max\left[h_1(A_1) - p_1 c_1 + \varepsilon_1, \ldots, h_N(A_N) - p_N c_N + \varepsilon_N\right]. \qquad (21)$$

The Compensating Variation (CV) is obtained by solving the equality: $V(A^0, p^0, y) = V(A^1, p^1, y - CV)$. Using the functional form in equation (21), we have:

$$\begin{aligned} &\bar{\gamma} y + \max\left[h_1(A_1^0) - p_1^0 c_1 + \varepsilon_1, \ldots, h_N(A_N^0) - p_N^0 c_N + \varepsilon_N\right] = \\ &\bar{\gamma}(y - CV) + \max\left[h_1(A_1^1) - p_1^1 c_1 + \varepsilon_1, \ldots, h_n(A_N^1) - p_N^1 c_N + \varepsilon_N\right]. \end{aligned} \qquad (22)$$

We can solve for CV and this results in:

$$\begin{aligned} CV = \frac{1}{\bar{\gamma}}\{ &\max\left[h_1(A_1^1) - p_1^1 c_1 + \varepsilon_1, \ldots, h_n(A_N^1) - p_N^1 c_N + \varepsilon_N\right] - \\ &\max\left[h_1(A_1^0) - p_1^0 c_1 + \varepsilon_1, \ldots, h_N(A_N^0) - p_N^0 c_N + \varepsilon_N\right]\}. \end{aligned} \qquad (23)$$

If the error terms are extreme value distributed, i.e. the MNL model, the expected CV for a change in attributes is (Hanemann, 1999):

$$E(CV) = \frac{1}{\mu\bar{\gamma}}\left\{\ln\sum_{i\in S}\exp(\mu V_{i1}) - \ln\sum_{i\in S}\exp(\mu V_{i0})\right\}, \qquad (24)$$

where μV_{i0} and μV_{i1} represent the estimated indirect utility before and after the change, $\mu\bar{\gamma}$ is the confounded estimate of the scale parameter and the marginal utility of money and S is the choice set.[7] With a linear utility function and only one attribute changing, the CV for a discrete choice is given by:

$$CV = \frac{1}{\gamma}\ln\left\{\frac{e^{V_{11}}}{e^{V_{10}}}\right\} = \frac{1}{\gamma}(V^1 - V^0) = \frac{\beta_k}{\gamma}(A_k^1 - A_k^0). \qquad (25)$$

[7] Note that this welfare measure is independent of the scale and, in practice, the scale parameter is set to equal one.

By looking at the expression in equation (25) it is easily seen that for a linear utility function, the marginal rate of substitution between two attributes is simply the ratio of their coefficients, and that the marginal willingness to pay for a change in attribute is given by

$$MWTP_i = -\frac{\beta_i}{\gamma}. \tag{26}$$

For policy purposes it is of interest, and often necessary, to obtain the distribution of the welfare effects. This can be done either by bootstrapping or by the Krinsky-Robb method (Krinsky and Robb, 1986). With bootstrapping a number of new data sets are generated by resampling, with replacement, of the estimated residuals. The utility across alternatives, along with the parameter point estimates, is calculated in order to create the dependent variable. For each of these new data sets the model is re-estimated and welfare measures are calculated. The Krinky-Robb method is based on a number of random draws from the asymptotic normal distribution of the parameter estimates and the welfare measure is then calculated for each of these draws. The Krinsky-Robb method is less computationally burdensome than bootstrapping, but its success critically depends on how closely the distribution of errors and asymptotically normal distribution coincide. For example Kling (1991) and Chen and Cosslett (1998) find that the two procedures give quite similar standard deviations.

The assumptions underlying the closed form solution of the welfare measures were (i) additive disturbances, (ii) an extreme value distribution[8] and (iii) constant marginal utility of income. Let us relax the assumption of constant marginal utility of income and no effect of personal characteristics. The CV is in generally found by solving the equality: $V(A^0, p^0, y) = V(A^1, p^1, y - CV)$. The problem is how to obtain an estimate of CV, when income enters the utility function nonlinearly. In such a case the marginal utility of income is not constant and there is no closed-form solution to calculate the welfare effects. McFadden (1995) suggests either estimating the welfare effects by simulation or by calculating theoretical bounds on the welfare effects. Morey et al

[8] A closed form solution of the welfare measure does in fact exist for the GEV distribution in which the extreme value distribution is a member, see McFadden (1995).

(1993) suggest an approach using a representative consumer approach, whereas Morey and Rossman (2000) impose piecewise constant marginal utility of income in the econometric model. The simulation approach is conducted in the following steps for a choice experiment consisting of *K* alternatives, and with *T* choice situations for each of the individuals. First, at iteration *t*, *K* randomly draws from a pre-specified distribution, e.g. an extreme value distribution, is performed. This results in the vector ε^t. Then, a numerical routine is applied to search for the CV^t, defined as:

$$V(A^0, p^0, y, s) \equiv E\left[v(A^0, p^0, y, \varepsilon^t)\right] = E\left[v(A^1, p^1, y - CV^t, \varepsilon^t)\right] \equiv V(A^1, p^1, y - CV \quad (27)$$

.

This procedure is repeated T times. Second, for each individual, the expected CV is approximated by

$$E(CV^t) = \frac{1}{T}\sum_{t=1}^{T} CV^t \;. \quad (28)$$

If the sample of *N* individuals represents a random sample of the population under study then the expected CV for the population is

$$E(CV) = \frac{1}{NT}\sum_{n=1}^{N}\sum_{t=1}^{T} CV^t \;. \quad (29)$$

The approach is easy in the case of an extreme value distribution, but more difficult for a GEV or a multivariate normal distribution. For a more detailed discussion on the simulation approach see e.g. McFadden (1995) or Morey (1999). The representative consumer approach, describe by Morey et al. (1993), uses a utility function of a representative individual. The result is that the repeated draws in McFadden's simulation approach can be skipped and a numerical routine can be directly applied in order to search for $E(CV)$. McFadden (1995) finds that this approach results in biased estimations of CV and that the percentage of bias increases with the size of the welfare change. The benefit of McFadden's theoretical bounds approach is that it makes the computations less difficult by imposing bounds on the welfare effects from a change. The piecewise constant marginal utility approach by Morey and Rossman (2000) is easy to apply since the welfare effects can be calculated directly from the estimates. Furthermore, Morey and Rossman also present how to calculate the welfare effects when the CV results in a change from one income level with a specific constant marginal utility of income to another income level with a different constant marginal utility of income.

7. Conclusions

This paper has discussed valuation of non-market goods when using choice experiments. The advantages of choice experiments are that values for each attribute as well as marginal rate of substitution between non-monetary attributes can be obtained. Moreover, rigorous tests of internal validity can be performed. The success of a choice experiment depends on the design of the experiment which, as stressed several times in the paper, is a dynamic process involving definition of attributes, attribute levels and customisation, context of the experiment, experimental design and questionnaire development. Important tasks in future research include improving the knowledge about how respondents solve a choice experiment exercise and if preferences are consistent over the course of the experiments. Furthermore, the choice sets created by the chosen experimental design strategy have an important impact on the results. This paper describes the D-optimal approach. One of the problems with this approach is the criterion of utility balance. As we mention, it is not clear that utility balance necessarily improves the results and further studies are needed on this issue.

If a stated choice preference method has to be used to value a non-market good, either a closed-ended CVM survey or a choice experiment can be applied. As a rule of thumb we would recommend that practitioners apply a closed-ended CVM survey if the interest is purely in valuing a certain environmental change. In other cases, a choice experiment may be more suitable since it produces more information. In the future, however, more research is needed on both methods, and particularly on their abilities to elicit true preferences.

Acknowledgements

We have received valuable comments from Gardner Brown, Henrik Hammar, G.S. Haripriya, Gunnar Köhlin, David Layton, Karl-Göran Mäler, Olof Johansson-Stenman and Thomas Sterner. Financial support from the Swedish International Development Agency, the Bank of Sweden Tercentenary Foundation and the Swedish Transport and Communications Research Board is gratefully acknowledged.

References

Adamowicz, W., P. Boxall, M. Williams and J. Louviere (1998a) Stated preferences approaches to measuring passive use values. *American Journal of Agricultural Economics* 80, 64-75.

Adamowicz, W., J. Louviere and J. Swait (1998b) Introduction to attribute-based stated choice methods. Report to NOAA Resource Valuation Branch, Damage Assessment Centre.

Adamowicz, W., J. Louviere and M. Williams (1994) Combining revealed and stated preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management* 26, 271-292.

Andreoni, J. (1989) Giving with impure altruism: Applications to charity and Ricardian equivalence. *Journal of Political Economy* 97, 1447-1458.

Bateman, I. and K. Willis (1999) *Valuing Environmental Preferences*. Oxford University Press.

Batsell, R. and J. Louviere. (1991) Experimental Analysis of Choice. *Marketing Letters* 2, 99-214.

Ben-Akiva, M. and S. Lerman (1985) *Discrete Choice Analysis. Theory and Applications to Travel Demand*. MIT Press.

Ben-Akiva, M. and T. Morikawa (1990) Estimation of travel demand models from multiple data sources. In Koshi, M. (ed.), *Transportation and Traffic Theory*, New York: Elsevier.

Ben-Akiva, M., T. Morikawa and F. Shiroishi (1992) Analysis of the reliability of preference ranking data. *Journal of Business Research* 24, 149-164.

Benjamin, J. and L. Sen (1982) Comparison of the predictive ability of four multiattribute approaches to attitudinal measurement. *Transportation Research Record* 890, 1-6.

Bhat, C. (1995) A heteroskedastic extreme value model of intercity travel mode choice. *Transportation Research B* 29, 471-483.

Bhat C. (1997) Covariance heterogeneity in nested logit models: Econometric structure and application to intercity travel. *Transportation Research B* 31, 11-21.

Blamey, R., J. Bennett, J. Louviere, M. Morrison and J. Rolfe (2000) A test of policy labels in environmental choice modeling studies. *Ecological Economics* 32, 269-286.

Boxall, P., W. Adamowicz, J. Swait, M. Williams and J. Louviere (1996) A comparison of stated preference methods for environmental valuation. *Ecological Economics* 18, 243-253.

Braden, J. and C. Kolstad (1991) *Measuring the Demand for an Environmental Improvement*. North-Holland, Amsterdam.

Bradley, M. (1988) Realism and Adaptation in Designing Hypothetical Travel Choice Concepts. *Journal of Transport Economics and Policy* 22, 121-137.

Brownstone, D. and K. Train (1999) Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics* 89, 109-129.

Bryan, S., L. Gold, R. Sheldon and M. Buxton (2000) Preference measurement using conjoint methods: An empirical investigation of reliability. *Health Economics* 9, 385-395.

Cameron, T., G. Poe, R. Ethier, and W. Schulze (1999) Alternative nonmarket value-elicitation methods: Are revealed and stated preferences the same? Working Paper Department of Economics, University of California.

Carlsson, F. and P. Martinsson (2001) Do hypothetical and actual marginal willingness to pay differ in choice Experiments? – Application to the valuation of the environment. *Journal of Environmental Economics and Management* 41, 179-192.

Carlsson, F. and P. Martinsson (2000) Design strategies for choice experiments in health economics. Working Paper, Department of Economics, Lund University.

Carson, R., N. Flores, K. Martin and J. Wright (1996) Contingent valuation and revealed preference methodologies. Comparing the estimates for quasi-public goods. *Land Economics* 72, 80-99.

Carson, R., R. Groves, and M. Machina (1999) Incentive and informational properties of preference questions. Paper presented at EAERE Ninth Annual Conference, Oslo.

Chen, H. and S. Cosslett (1998) Environmental quality preference and benefit estimation in multinomial probit models: A simulation approach. *American Journal of Agricultural Economics* 80, 512-520.

Ciriacy-Wantrup, S. (1947) Capital returns from soil-conservation practices. *Journal of Farm Economics* 29, 1181-1196.

Cummings, R., G. Harrison and E. Rutstrom (1995) Home-grown values and hypothetical surveys: Is the dichotomous choice approach incentive compatible. *American Economic Review* 85, 260-266.

Davis, R. (1963) *The Value of Outdoor Recreation: An Economic Study of the Maine Woods*. Ph.D. dissertation, Department of Economics, Harvard University.

Duffield, W. and D. Patterson (1991) Inference and optimal design for a welfare measure in dichotomous choice contingent valuation. *Land Economics* 67, 225-239.

Diamond, P. and J. Hausman (1994) Contingent valuation: Is some number better than no number? *Journal of Economic Perspectives* 8, 45-64.

Frykblom, P. (1997) Hypothetical question modes and real willingness to pay. *Journal of Environmental Economics and Management* 34, 275-287.

Gutowski, W. and J. Georges (1993) Optimal Sophisticated Voting Strategies in Single Ballot Elections Involving Three Candidates. *Public Choice* 77, 225-247

Hanemann, M. (1984) Discrete/Continuous Models of Consumer Demand. *Econometrica*. 52, 541-561.

Hanemann, M. (1994) Valuing the environment through contingent valuation. *Journal of Economic Perspectives* **8**, 19-43.

Hanemann, M. (1999) Welfare analysis with discrete choice models. In Herriges, J. and C. Kling (eds.) *Valuing Recreation and the Environment*, Edward Elgar.

Hanemann, M. and B. Kanninen (1999) The statistical analysis of discrete-response CV data. In Bateman, I. and K. Willies (eds.) *Valuing Environmental Preferences*, Oxford University Press.

Hanley, N., R. Wright and W. Adamowicz (1998) Using choice experiments to value the environment. *Environmental and Resource Economics* 11, 413-428.

Huber, J. and K. Zwerina (1996) The importance of utility balance in efficient choice designs. *Journal of Marketing Research* 33, 307-317.

Johansson-Stenman, O. and H. Svedsäter (2001) Choice experiments and self image: Hypothetical and actual willingness to pay. Working Paper Department of Economics, Gothenburg University.

Johnson, R., R. Ruby and W. Desvousges (2000) Willingess to pay for improved respiratory and cardiovascular health: A multiple-format, stated-preference approach. *Health Economics* 9, 295-317.

Johnson, R., K. Mattews and M. Bingham (2000) Evaluating welfare-theoretic consistency in multiple-response, stated-preference survey. TER Working Paper T-0003.

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

Kanninen, B. (1993) Optimal experimental design for double-bounded dichotomous choice contingent valuation. *Land Economics* 69, 138-146.

Kanninen, B. (2001) Optimal design for multinomial choice experiments. Unpublished Working Paper.

Kling, C. (1991) Estimating the precision of welfare measures. *Journal of Environmental Economics and Management* 21, 244-259.

Krinsky, I. and R. Robb (1986) On approximating the statistical properties of elasticities. *Review of Economics and Statistics* 68, 715-719.

Kuhfeld, W., R. Tobias and M. Garrat (1994) Efficient experimental design with marketing research applications. *Journal of Marketing Research* 31, 545-557.

Lancaster, K. (1966) A New Approach to Consumer Theory. *Journal of Political Economy* 74, 132-157.

Layton, D. and G. Brown (1998) Application of stated preference methods to a public good: Issues for discussion. Paper presented at the NOAA Workshop on the Application of Stated Preference Methods to Resource Compensation, Washington, DC, June 1-2, 1998.

Layton, D. and G. Brown (2000) Heterogenous preferences regarding global climate change. *Review of Economics and Statistics* 82, 616-624.

Leigh, T., D. MacKay and J. Summers (1984) Reliability and validity of conjoint analysis and self-explicated weights: A comparison. *Journal of Marketing Research* 21, 456-462.

Louviere, J. (1993) Conjoint analysis. In Bagozzi, R. (ed.), *Advanced Methods in Marketing Research*, Cambridge: Blackwell Business.

Louviere, J., D. Hensher and J. Swait (2000) *Stated Choice Methods. Analysis and Application*. Cambridge: Cambridge University Press.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behavior. In Zarembka, P. (ed.) Frontiers in Econometrics, Academic Press, New York.

McFadden, D. (1995) Computing Willingness-to-pay in Random Utility Models. Working Paper, Department of Economics, University of California, Berkeley.

McFadden, D. and K. Train (2000) Mixed MNL models for discrete response. *Journal of Applied Econometrics* 15. 447-470.

Mazotta, M. and J. Opaluch (1995) Decision making when choices are complex: A test of Heiners hypothesis. *Land Economics* 71, 500-515.

Morey, E. (1999) TWO RUMs uncloaked: Nested-logit models of site choice and nested-logit models of participation and site choice. In Herriges, J. and C. Kling (eds.) *Valuing Recreation and the Environment*, Edward Elgar.

Morey, E. and K. Rossman (2000) The compensating variation for a change in states when there is only one alternative in each state and where utility is assumed a linear spline function of income. Working paper at the University of Colorado.

Morey, E., R. Rowe and M. Watson (1993) A repeated nested-logit model of Atlantic Salmon Fishing. *American Journal of Agricultural Economics* 75, 578-592.

Mäler, K-G. (1974) *Environmental Economics: A Theoretical Inquiry*. Resources for the Future.

Polak, J. and P. Jones (1997) Using stated-preference methods to examine travelers preferences and responses. In Stopher, P. and M. Lee-Gosselin (eds.), *Understanding Travel Behavior in an Era of Change*, Oxford: Pergamon.

Ryan, M. and J. Hughes (1997) using conjoint analysis to assess women's preferences for miscarrige management. *Health Economics* 6, 261-273.

Revelt, D. and K. Train (1998) Mixed logit with repeated choices: Households' choices of appliance efficiency level. *Review of Economics and Statistics* 80, 647-657.

Schkade, D.A. and J.W. Payne (1993), Where do the number come from?: How people respond to contingent valuation questions. In: Hausman J.A. (ed.), *Contingent valuation: a critical assessment* Amsterdam, North-Holland.

Steffens, K., F. Lupi, B. Kanninen and J. Hoen (2000) Implementing an optimal experimental design for binary choice experiments: An application to Bird Watching in Michigan. In S. Polasky (ed.) *Benefits and Costs of Resource Policies Affecting Public and Private Land*. Western Regional Research Publication.

Swait, J. and W. Adamowicz (1996) The effect of choice environment and task demands on consumer behavior. Discriminating between contribution and confusion. Working Paper, Department of Rural Economy, University of Alberta.

Swait, J. and J. Louviere. (1993) The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. *Journal of Marketing Research* 30, 305-314.

Train, K. (1998) Recreation demand models with taste differences over people. *Land Economics* 74, 230-239.

Vick, S. and A. Scott (1998) Agency in health care: Examining patients' preferences for attributes of the doctor-patient relationship. *Journal of Health Economics* 17, 587-605.

Wardman M. (1988) A comparison of revealed and stated preference models of travel behaviour. *Journal of Transport Economics and Policy* 22, 71-91.

Zwerina, K., J. Huber and W. Kuhfeld (1996) A general method for constructing efficient choice designs. Working Paper, Fuqua School of Business, Duke University.

Policy Implications and Analysis of the Determinants of Travel Mode Choice: An Application of Choice Experiments to Metropolitan Costa Rica

Francisco Alpizar
Fredrik Carlsson

Abstract
In this paper we study a group of policies aimed at discouraging the use of private transportation during peak hours, both directly and indirectly, by increasing the attractiveness of the only available substitute, the bus. This is done using a choice experiment constructed to find the answer to the following basic question: Given fixed house-to-work structures and no working hour flexibility, by how much is the choice of travel mode for commuters to work sensitive to changes in travel time, changes in costs for each mode and other service attributes? This information is then used to identify the most suitable combination of policies dealing with air pollution and congestion in the typical developing country context of metropolitan Costa Rica. We also provide estimates of the value of travel time as a measure of the potential benefits gained from reduced congestion.

1. Introduction

The last 20 years have been characterized by a dramatic increase in the urban populations of most developing countries. Even small countries like the ones in Central America now have metropolitan areas that surpass one million inhabitants. This increase, in combination with a lack of urban planning and an inefficient transportation system, causes problems of decreased air quality and traffic congestion in urban environments. Pollution is associated with a wide variety of health problems, deterioration of buildings, acid rain and global warming. Traffic congestion further complicates matters since it not only imposes high costs in terms of lost time and high stress, but also increases emissions by decreasing the speed of travel.

The main objective of this study is to contribute to the design of policies dealing with the problems of congestion and air pollution in the urban context of a typical developing country. We study the determinants of the choice of transport mode for work trips in the short run, where we treat the number of trips as exogenously given. This is done with a choice experiment conducted on car commuters. These commuters are responsible for the worst congestion and pollution during peak hours, and at the same time are the ones who can most easily switch to public modes of transportation. Our basic argument is that a commuting mode switch from car to bus can increase the efficiency of the transport system in two basic ways: i- it makes a more efficient use of the available infrastructure by transporting a higher number of individuals per unit of physical area, potentially reducing congestion and travel time, and ii- it makes a more efficient use of the environment by creating less emissions per commuter and, if congestion is reduced and travel speed increases, less emissions in general.

This argument can be subjected to many qualifications; particularly, it strongly depends on the average number of passengers per bus and the environmental performance of buses, compared to cars. For instance, if the mode switch creates the need for more buses and the bus company introduces a 20-year-old bus that will run almost empty most of the time, then most likely the mode switch will worsen the pollution problem. Our view is that there are many dimensions of a sustainable strategy to manage transport, and imposing minimum requirements on the environmental performance of buses should be a first step. Additional steps are required including the provision of incentives to manage the demand for private transportation. We explore a combination of policies aimed at increasing the cost of private transportation, specifically increased fuel and parking costs, and improving public transportation; in this case reduced travel time, subsidized fares and improved quality of the service. The joint implementation of these policies aims at reducing congestion and pollution, especially during peak hours, by restraining demand for private transportation while providing a suitable substitute. Currently the Costa Rican Transport Ministry is involved in a program to improve and restructure the public transportation system, emphasizing the type and quality of service provided. One of the aims of this program is to achieve a reduction in the level of congestion and pollution originating from urban transportation. Therefore, our survey is also a test of the ability of such a system to detract customers from private transportation. In addition, our survey is constructed to identify which aspects of public and private transportation are truly relevant determinants of the commuting mode choice decision, and this information can be used to prioritize the efforts of the Ministry. To our knowledge this is the first mode choice study conducted in Costa Rica.

2. Environmental Regulation of Transport in a Developing Country Context

Both market-based and command and control policies require strong and stable institutions, adequate legislation and effective monitoring and enforcement. Many developing countries have weak institutions and unclear legislation, and in many cases there is a weak political commitment to environmental goals, which means that resistance to environmental measures is more likely to be successful. In addition, uncertainty about the permanence of environmental regulations causes an inefficient adaptation to the long run incentives of any given policy (O'Connor, 1999; Eskeland and Jimenez, 1992). In this context it may be particularly important to revise the perspective of environmental regulations by emphasizing their short run effects, and identifying ways of reducing the potential resistance against certain types of policies (see e.g. Calfee and Winston, 1998; Harrington et al. 1998).

The regulation of the urban transport system, intended to reduce congestion and pollution, requires a combination of policies. On one hand, regulating private transportation requires a number of policies that can, more or less, mimic the properties of the first-best solution.[1] On the other hand, public transportation is often regulated in terms of fees, routes, number of departures and other characteristics. Decision makers need information regarding which of these characteristics are most relevant to travelers. This is particularly relevant if there is a wish to shift travelers from private to public transportation, and if there is a need to provide travelers with an acceptable alternative mode, as a way to increase the acceptance of the environmental regulations.

In this paper we focus on policies aimed at reducing the amount of private trips, by promoting a change from private to public modes of transportation, for a given structure of car ownership and house-to-work locations. These policies include the complementary application of instruments to reduce the number of private trips by increasing the cost per journey, and by improving the system of public transport.

An increase in the cost per trip for the private mode is expected to have two effects in the short run. First, there will be a reduction in the overall number of trips and second there will be a substitution towards public transport.[2] In the case of trips to work, we expect the first effect to be small since most trips are likely to be compulsory. A substitution from private to public transportation is expected to have two further effects: (i) a reduction in congestion levels, which has an unequivocal direct effect on emissions and (ii) if the new passengers could be accommodated by the existing bus journeys, then passengers per bus trip would increase, and emissions per passenger will be lowered. For the particular case of Costa Rica, buses run during peak hours with an average load of approximately two thirds of their capacity, so a marginal mode switch is not expected to raise the number of bus journeys required to cope with the increase in demand (Ministry of Public Infrastructure and Transport, 1999). In cities where buses run with approximately full loads, the number of bus journeys might increase due to increased demand. In such a case, the net benefit of a mode switch might not necessarily be positive and will depend on the performance of the bus fleet.[3]

Policies aimed at decreasing private transportation by means of increased costs include fuel taxes, parking fees and road tolls in city centers. The last two are particularly relevant for tackling congestion, since they directly raise the cost of entering the city (Small, 1992; Willson and Shoup, 1990). Fuel taxes have been successful in reducing fuel demand in many countries, although public perception about them is negative (Sterner, 1994; Thomson, 1998). One way to reduce the political opposition to such measures is to provide a suitable system of public transportation. The

---

[1] See for example De Borger et al. (1996), Eskeland and Devarajan (1996) and Fullerton and West (2001).

[2] In the long run, drivers can buy smaller and more fuel efficient cars, or even move closer to work in response to more expensive travel.

[3] Metropolitan Costa Rica, as many other cities worldwide, has embarked in a process of renewing the bus fleet. Newer buses have certainly improved their environmental performance, and scrapping old buses will be one of the first steps of a strategy to reduce pollution.

design and quality of the public transportation system is thus an important factor itself in decreasing the use of private transportation, by providing a suitable substitute to the private mode.

3. Urban transport in Costa Rica

Almost half of the population and most of the economic and governmental activity in Costa Rica is located in the metropolitan areas of San José, Alajuela, Cartago and Heredia. The city of San José, the capital, shows signs of a collapse in its transport infrastructure,[4] particularly during the peak hours of the morning (7 to 9 am) and the evening (5 to 8pm). The roads linking the other cities to San José have also reached severe congestion levels. Most inhabitants of metropolitan Costa Rica are exposed daily to very high pollution levels, with road transportation contributing more than 80% of the total air pollution in metropolitan San José (Alfaro, 1999). A survey conducted by the Ministry of Health identified respiratory and pulmonary illnesses as the most common causes for visiting the public health hospitals (Ministry of Health, 1997).

Government reactions have been mostly limited to the provision of new infrastructure. Unfortunately, the capacity of the additional infrastructure is overflowed sometimes even before its construction is completed, in an extreme example of the so-called "law of highway congestion". In addition, this focus on infrastructure provision has further encouraged car use (Small, 1992).

In recent years there have been some attempts to tackle the problem of air pollution, mostly by using command and control measures to reduce the amount of emissions per kilometer. As of 1995, all new cars are required to have a working catalytic converter, and lead was eliminated from fuels in 1996. Also in 1996, the government introduced emissions standards for all vehicles. These standards were, however, not very strict and there is little evidence that this policy has had any effect on emissions (Pujol, 1996; Jiménez, 1997). The tax structure for importing and owning a car has had perverse effects on the generation of pollution (Echeverría and Solórzano, 2000). Since 1986, older cars face an import tariff of 30%, whereas new cars paid a 100% tariff. The

[4] Although approximately 70% of all trips are made using the bus mode (mostly low and middle-low income individuals), the number of commuters in private modes surpasses the capacity of the available

impact of this policy was dramatic. By the end of 1997, 72% of the car fleet for private use was more than 10 years old (RECOPE, 1999). In the last years the government has tried to reduce this distortion despite opposition from the affected parties. In addition, the yearly road tax is calculated based on the value of the car, such that newer cars, which are more fuel-efficient and generate fewer emissions, are subject to higher taxes than older vehicles.

Car ownership and fuel consumption have been increasing at high rates. Between 1989 and 1999, the car fleet grew at an average rate of 7.6% per year,[5] with gasoline and diesel use increasing at an average rate of 10.4% and 7% per year, respectively. In the same period, per capita GDP increased, on average, 2.2% per year (Proyecto Estado de la Nación, 1999). The main factors behind this development are most likely the declining real price of fuel, the perverse tax structure for car ownership and the lack of an adequate system of public transportation. As of 1999, real prices on all fuels were slightly lower than in 1988, with an average yearly growth of 0.02% in that period.

On the other hand, the public transportation system is underdeveloped. The government has paid little attention to its quality and service. Many times companies operating similar routes find themselves competing for passengers along the same street, with most routes leading all the way to the center of the capital (i.e. a radial system). This has resulted in deteriorating levels of service and congestion deadlocks in the city center at peak hours. Fares have been determined by the regulatory authority based on basic operating costs, paying little attention to the quality and type of service provided. Now, the Ministry of Transportation is trying to organize the sector by restructuring the number of routes into a trunk and feeding system. Furthermore, companies face higher standards regarding the vintage of buses. The new program also provides economic incentives to bus companies that comply with the new regulations. Bus fares will be linked to an evaluation of the service provided, as measured by various factors. If the project becomes a reality, it will imply better buses, better quality of service, less congestion due to fewer buses entering the city center and potentially

---

infrastructure. Furthermore, there are institutional and geographical limitations to the expansion of the road network.

shorter travel times. As mentioned in the introduction, our survey then is also a test of the ability of such a system to attract new customers away from private transport.

4. Mode Choice Experiments

In order to evaluate the potential impact of different policies on the substitution between private and public transportation, information regarding traveler preferences for the attributes of the transport system is needed. Since some of the attributes of interest do not exist today, it is not possible to rely only on revealed preference data for that purpose. Therefore, we conduct a mode choice experiment evaluating traveler preferences for different attributes of both private and public transportation. The basic idea behind a mode choice experiment is to create a hypothetical situation and elicit the preferences of commuters for different attributes, through their choice of mode of transport in each of those situations.

In this paper we apply a general type of model called Random Parameter Models, where taste variation among individuals is explicitly treated (see e.g. Bhat, 2000; Train, 1998). We assume a linear latent indirect utility function *U,* consisting of a systematic and a stochastic part:

$$U_{iqt} = \alpha_{iq} + \gamma_i s_q + \beta_{iq} x_{iqt} + \varepsilon_{iqt}, \qquad (1)$$

where α_{iq} captures an intrinsic preference of individual *q* for alternative *i*, $\gamma_i s_q$ captures systematic preference heterogeneity as a function of socio-demographic characteristics (s_q), x_{iq} is the vector of *K* attributes (including costs) for alternative *i* and ε_{iq} is a stochastic component that reflects observational deficiencies arising from unobservable components, measurement error and/or heterogeneity of preferences. In choice experiments, the respondents face a sequence of such decisions, where each decision set (indexed by *t*) contains different profiles of the alternatives. The probability that individual *q* chooses alternative *i* in a choice situation *t*, conditional on β_{iq}, is given by:

[5] The car fleet doubled from 1978 to 1988 and doubled again from 1988 to 1998. Approximately 75% of the fleet is for private use (RECOPE, 1999). According to a recent local newspaper survey, 47% of all

$$P_q(it \mid \beta_q) = P\left\{(\alpha_{iq} + \gamma_i s_q + \beta_{iq} x_{iqt} + \varepsilon_{iqt}) > (\alpha_{jq} + \gamma_j s_q + \beta_{jq} x_{jqt} + \varepsilon_{jqt}) \forall j \in \mathbf{A}_t\right\} \quad (2)$$

The vector of coefficients β_{iq} is assumed to vary in the population, with probability density given by $f(\beta \mid \theta)$, where θ is a vector of the true parameters of the taste distribution. In this simple form, the utility coefficients vary across individuals, but are constant across the choice situations for each individual. This reflects an underlying assumption of stable preference structures for all individuals (Train, 1998). An important element of these Random Parameter Models is the assumption regarding the distribution of each of the random coefficients. We use two alternative formulations. The first assumption is a normal distribution, i.e. that the coefficient for the *k*-attribute is given by $\beta_k \sim N[b_k, w_k]$. However, a coefficient might then be negative for some individuals and positive for others. For most of the variables in this mode choice experiment it is more reasonable to expect that all respondents have the same coefficient sign. A more reasonable assumption would be to assume a log-normal distribution. In this case, the log-normal coefficients have the following form:

$$\beta_{kq} = \pm \exp(b_k + v_{kq}), \quad (3)$$

where the sign of coefficient β_k is determined by the researcher according to expectations, b_k is constant and the same for all individuals, and v_{kq} is normally distributed across individuals with mean and variance equal to 0 and σ_k^2, respectively. The coefficient has the following properties: a) median = $\exp(b_k)$, b) mean = $\exp(b_k + \sigma_k^2/2)$; and c) standard deviation = $\exp(b_k + \sigma_k^2/2)(\exp(\sigma_k^2) - 1)^{0.5}$. For a more detailed treatment of preference heterogeneity, see for example Bhat (1998, 2000).

---

households own at least one vehicle. This data refer to the whole country, and the numbers are probably much higher in the metropolitan area.

5. The Mode Choice Survey

The population used in the survey is individuals with work that have access to a car, and that are living and working in the metropolitan area of San Jose. The reason is that we want to ensure that the respondents can actually make a choice, in the short run, between private and public transportation. This sampling strategy excludes the low-income segments of society from our study, since they cannot afford to own a car. In addition, as pointed out by a referee, changes in the costs of driving a car and in the attractiveness of the bus might have an impact on the observed long-term path of increasing car ownership and use described in section 3. Since we do not sample potential car users, we are unable to measure this impact. Still, in the process of designing the sample we believed that the short-run situation of transportation in Costa Rica was already not viable with the current size of the private car fleet, and we decided therefore to focus on current car owners. Hence, if the policies studied under the current design manage to reduce private commuting by increasing the relative attractiveness of commuting by bus, we can only hypothesize that this might reduce the incentives to buy a car to commute to work.

We also focus on work trips since they, in particular, contribute to congestion problems at peak hours. In our experiment the length and destination of work trips is assumed to be fixed. Additionally, the timing of the trips is assumed to be exogenous, i.e. we do not allow the commuters to adjust their schedule in response to the hypothetical situation presented in the experiment. This assumption is required in order to maintain a higher control of the experiments. If not, we would have to make the mode choice decision a function of the chosen departure time. Since in Costa Rica there is very little work-hour flexibility, both in the public and private sector, and given that peak hours start early in the morning, we do not believe this addition will greatly improve our results. In addition, we note that the hypothetical mode choice situations do not depend on congestion levels[6]. Finally, we restrict the sample to work trips with an origin and destination within the metropolitan area, in order to restrict the analysis to the urban bus system. One limitation of our analysis is that it is partial in nature. For example, our estimates of cost elasticities cannot be regarded as the overall price

[6] For discussions on commuter's scheduling adjustments to changes in congestion, see Small (1982), Arnott et.al. (1993).

elasticity of transport demand for metropolitan San José. Even in the short run, individuals can adjust to higher costs per car trip by carpooling and reductions in non-work trips, among other possibilities. It is actually reasonable to expect that people will reduce the number of trips starting with the less important ones (mainly non-work trips). All of these elements hint to larger effects of the policies that we analyze.

A survey company with more than 10 years of experience conducted the survey between September 1st and October 30th, 2000. All interviews were personal interviews where the enumerator read the questions aloud. The respondents were visited at their homes after office hours. The National Institute of Statistics has divided the metropolitan area into 5700 segments, of which 43 were randomly selected for this survey. The questionnaires were then randomly assigned to each of these segments. The field supervisor would then decide which houses to survey in each locality. If the person at the door did not meet our criteria, the house next door was surveyed. A total of 602 questionnaires were completed, and the statistical error was 4%.

Table 1 presents some descriptive statistics. The average income in Costa Rica was 160,000 colones[7] per month in 1999, whereas the average income of the 10th decile was around 600,000. Our sampled income distribution seems to fit this description, given the exclusion of low-income and rural families from our sample. The gender composition of our sample is simply a reflection of our sampling strategy, which required the respondent to be currently employed. Of particular interest is the extremely high number of respondents who usually use the car to go to work. A large share of the respondents also stated that they sometimes need the car in their line of work.

[7] 1USD corresponds to approximately 300 colones.

Table 1: Sample statistics

Variable		Frequency
Income in colones	0-200,000	31.6%
	200,001-400,000	40.3%
	400,001-600,000	18.3%
	600,001-800,000	5.7%
	800,001+	3.5%
Gender	Male	77.2%
	Female	22.8%
Usual travel mode	Car	90.7%
	Bus	9.3%
Car needed at work	Yes	39.9%
	No	60.1%

The questionnaire consisted of three parts. The first part contained a number of questions regarding their present work trips. The second part of the survey contained the choice experiment. Before the actual experiment was conducted, the enumerator carefully explained each of the attributes in the choice experiment. The respondent also received a written summary of the attributes. The last part of the survey contained questions regarding the respondent's socio-economic status and debriefing questions regarding the choice experiment.

Discussions with experts, several focus groups, informal tests of the questionnaire and a formal pilot study conducted by the survey company preceded the final design of the experiment and the survey. These intended to get a broad picture of the real problem of transportation in metropolitan Costa Rica. They also aimed at identifying the relevant alternatives, their attributes and realistic attribute levels required to create the profiles. Two important conclusions of these discussions were the need to customize some of the attributes, particularly travel time and cost per trip, and the possibility of restricting the number of alternatives to only two: car and bus. All of the attributes selected are factors that a policy maker can directly affect through, for example, a gasoline tax or, indirectly, through contracts with the bus companies, for example. In addition, all these attributes were regarded as relevant based on the information from focus groups. Local newspapers have been discussing the need for a solution to the problem of public and private transportation. Different alternatives have been discussed and described, in particular the program for increasing the quality of bus services. This gave extra realism to our survey.

The selected attributes for the car alternative are: (1) operating costs, (2) travel time per trip and (3) parking cost. The selected attributes for the bus alternative are: (1) travel time, (2) bus fare per trip, (3) punctuality, (4) distance to bus stop, (5) frequency of departures and (6) comfort and security. These attributes and their levels are presented in Table 2.

Table 2: Attributes and attribute levels

Attribute		Levels
Travel cost car, per trip		i. Same as today, ii. 25% increase, iii. 40% increase
Bus fare, per trip		i. 50 colones, ii. 100 colones
Parking cost		i. Free parking, ii. 400 colones per day
Travel time	**Car**	i. Same as today
	Bus	i. Same time as car, ii. 20 minutes longer than car, iii. 30 minutes longer than car, iv. 40 minutes longer than car
Punctuality		i. The bus is always on time, ii. The bus sometimes is more than 15 minutes late
Distance to bus stop		i. 10 minutes, ii. 15 minutes, iii. 20 minutes
No. of departures		Every i. 5 minutes, ii. 10 minutes, iii.15 minutes
Comfort and security		i. Same as today, ii. The Program for Quality Improvement is implemented

The cost per trip and travel time for the car alternative, and travel time by bus, were customized to the current situation and the actual levels are thus not presented in Table 2.[8] Questions regarding the distance to work, type of fuel used and the travel time by car were asked at the beginning of the survey. The enumerator then filled in the relevant information in the choice experiment. The cost per car trip in colones was calculated based on conversion tables, which included the percentage increase in gasoline price. The punctuality attribute was defined and explained such that it was not related to frequency of departures. For a high frequency service, one could argue that punctuality is not important since travelers prefer to go to the bus stop spontaneously. Nevertheless, it was explained that delays in the bus imply that for at least 15 minutes there was no

[8] In principle, the best way to customize travel time for the bus option would have been to define percentage increases relative to the car alternative. Nevertheless, the results from focus groups and the pilot study made us fear that this would be too demanding for the enumerator and possibly for the respondent as well. The levels chosen, in terms of absolute differences, where then the most reasonable ones, based on the results of focus groups. The average travel time for trips to work by bus and car are 60 and 25 minutes, respectively, for the individuals in our sample.

bus stopping, irrespective of the frequency of departures. This is more in line with the actual situation. The Program for Quality Improvement, currently under study by the government, was carefully described. If this program were implemented, it would bring an increase in the quality of the service, including more comfortable buses, and higher security both onboard and at improved bus stops.

The choice situations were constructed with a linear D-optimal design using SAS (see e.g. Kuhfeld, 2001). The eight attributes were combined into 24 choice sets. These where then divided into three groups of eight choice situations, again using a D-optimal criterion. Consequently, each respondent answered eight choice sets. In each choice set they were to choose between going to work by car or by bus, given the combination of attributes.

Since task complexity, learning and fatigue effects are mentioned as criticisms of this type of experiments, we conduct a test of the stability of preferences. Following Carlsson and Martinsson (2001), a simple test was incorporated into the design. This test consisted of surveying half of the respondents with the choice sets in the order {A,B} and the other half in the order {B,A}.[9] A test for stability was performed by comparing the preferences estimated for the choices in subset A, when it was given in the sequence {A,B}, with the preferences obtained when the choices in subset A were given in the sequence {B,A}. This can be formally tested in a likelihood ratio test between the pooled model of the choices in subset A and the separate groups. A similar test can be performed for subset B. If the pooled model cannot be rejected, we can reject the hypothesized presence of the effects mentioned above. Based on the standard MNL, the hypothesis of stable preferences cannot be rejected.[10]

The survey also included debriefing questions, mainly intended to identify respondents who did not like (protestors) or did not understand the experiment. Of all 602 interviewed individuals, 3.8% (23 individuals) expressed a negative perception or understanding of the experiment. These questionnaires are excluded from the estimations presented in the next section.

[9] "A" refers to the original first 4 choice sets, and "B" to the original last 4 choice sets in each experiment.

6. Results

We estimate two Random Parameter Logit models (RPL), in addition to the multinomial logit model. The two RPL models are: one with the attributes independently normally distributed and one with the attributes independently lognormally distributed. In both RPL models the cost attributes are kept fixed for several reasons: (i) we wish to restrict the cost variables to be non-positive for all respondents, hence a normal distribution is not recommended, (ii) a lognormal distribution, which would restrict the sign of the variable, can result in extremely high values-of-time estimates, since values of the cost attribute close to zero are possible (Revelt and Train, 1998), and (iii) the distribution of the marginal value of time is simply the distribution of the time attribute.

The attribute "distance to the bus stop" was consistently insignificant and caused problems with convergence for some models, and was therefore dropped from the analysis.[11] Both specifications include a mode specific intercept for car; in the normal specification the coefficient is fixed, while normally distributed in the model with lognormally distributed attributes.[12] Further, a variable capturing state dependence or inertia is included. This variable is specified as a dummy variable equal to one if the respondent usually uses the car when traveling to work. In both random parameter models, this fixed state dependence variable could not be rejected in a likelihood ratio test. Furthermore, a number of individual characteristics are included as fixed coefficients. These are (i) *need car*, indicating whether the respondent sometimes needs his/her own car in the line of work, (ii) the respondent's *income* and (iii) the respondent's *age*.

---

[10] The statistics are equal to 12.65 and 7.77 respectively, the statistic is χ^2-distributed with 13 degrees of freedom.

[11] This variable was not significant in the pilot study either, although it was mentioned several times during the focus groups. We tried to change the explanation of this attribute in order to have individuals make a trade-off with this characteristic. There are several explanations why this attribute turns out to be insignificant. First, the respondents may actually not see this attribute as important, or the choice experiment may have been too difficult and therefore many respondents chose not to focus on all attributes, including the distance to the nearest bus stop. Second, it may not have been thought of as credible that the government could actually change the distance to the bus stop by changing the routes.

[12] The model with normally distributed attributes did not converge when we included a normally distributed intercept.

In Table 3 we present the results for the three estimated models. The models were estimated with simulated maximum likelihood (see Revelt and Train 1998, Train 1999), based on Halton draws and 500 replications, using Limdep 7.0.2.

Table 3: Econometric results[13]

	MNLogit	RPL (normal)		RPL (lognormal)			
	Coeff	Coeff.	Coeff. Std.	Coeff	Coeff Std.	Median	Mean
Constant: Car	0.098 (0.235)	-1.427 (0.740)[c]	Fixed	-0.488 (0.814)	2.594 (0.266)[a]	-0.488 (0.814)	
Time (minutes): Car	-0.023 (0.004)[a]	-0.032 (0.019)[b]	0.066 (0.011)[a]	-3.221 (0.496)	0.430 (0.467)	-0.040 (0.020)[b]	-0.044 (0.016)[a]
Time (minutes): Bus	-0.017 (0.003)[a]	-0.053 (0.006)[a]	0.043 (0.006)[a]	-3.187 (0.151)	0.988 (0.208)[a]	-0.042 (0.006)[a]	-0.067 (0.018)[a]
Cost (colones): Car	-0.001 (0.0003)[a]	-0.003 (0.090)[a]	Fixed	-0.003 (0.0008)[a]	Fixed		
Cost (colones): Bus	-0.001 (0.002)	-0.002 (0.249)	Fixed	-0.002 (0.002)	Fixed		
Parking (dummy)	-0.190 (0.075)[b]	-0.214 (0.151)	1.446 (0.204)[a]	-1.935 (0.607)	1.522 (0.313)[a]	-0.144 (0.088)[c]	-0.460 (0.490)
Punctuality (dummy)	0.639 (0.076)[a]	1.111 (0.181)[a]	1.864 (0.228)[a]	-0.495 (0.253)	1.325 (0.181)[a]	0.610 (0.154)[a]	1.468 (0.701)[b]
No. of departures (units)	0.031 (0.011)[a]	0.006 (0.030)	0.240 (0.030)[a]	-3.225 (0.646)	1.053 (0.456)[b]	0.040 (0.026)	0.069 (0.019)[a]
Quality program (dummy)	0.054 (0.075)	0.171 (0.132)	0.357 (0.312)	-2.928 (1.458)	1.386 (0.447)[a]	0.054 (0.078)	0.140 (0.289)
Non-random socio-demographic characteristics							
State dependence	1.755 (0.111)[a]		4.300 (0.518)[a]		4.266 (0.555)[a]		
Need car at work	0.553 (0.081)[a]		0.791 (0.307)[a]		1.043 (0.348)[a]		
Income	0.041 (0.011)[a]		0.067 (0.041)[c]		0.050 (0.046)		
Age	-0.081 (0.035)[b]		-0.094 (0.139)		-0.146 (0.154)		
Log-likelihood	2226		1695		1662		
Pseudo R2	0.31		0.47		0.48		

Columns 3 and 4 give the estimated coefficients for the model with the attributes normally distributed and columns 5 and 6 for the model with the attributes lognormally distributed. The last two columns in the table present the estimated median and mean of β_k for the lognormal distribution. Note that the attributes are alternative specific; this

13 Standard errors in parentheses, a-significant at 1% level, b-significant at 5% level, c-significant at 10% level.

implies that for the MNL model and the normal RPL model, the coefficients correspond to the parameters of the utility functions for *each* of the two alternatives. For the lognormal RPL, the mean and median of the corresponding parameters of the utility functions are given in columns 7 and 8. For example, the negative parameter for cost car indicates that an increase in this item will reduce the utility of going by car, and hence the probability of choosing the car to commute. The two RPL models have a substantially higher pseudo-R^2 compared to the MNL model. The MNL is a restricted version of the two RPL models, in which all the coefficients are deterministic. In a likelihood ratio test we can reject the restrictions imposed by the MNL model.[14] The two RPL models give similar results in terms of the significance of coefficient and standard deviations estimates. [15] The only differences are that the mean coefficient for *Number of departures* is insignificant in the normal model while significant in the lognormal model, that the estimated standard deviation for *Travel time* by car is only significant in the normal model and finally that the estimated standard deviation for the *Quality program* is only significant in the lognormal model. Most of the standard deviations are significant, reflecting heterogeneity in the underlying preference structure. Since the estimated standard deviations are large relative to the estimated mean coefficients in the RPL model with normal distribution, a relatively large fraction of the respondents actually has the reverse sign of the preference for the different attributes. This is perhaps not a desirable feature of the model and as such, a justification for preferring the lognormal model.

In the lognormal model, all attributes except the *Bus fare*, and the *Quality program*, for the bus, and *Parking*, for the car have significant mean effects.[16] The relative importance of the level-of-service attributes is revealed by the estimates of the mean in columns (3) and (8) for the normal and lognormal models, respectively, for those attributes that share the same unit of measurement. For example, *Punctuality* has

---

[14] For the RPL(normal) and RPL(lognormal) models, the test statistics are 1062 which is χ^2-distributed with 6 degrees of freedom, and 1128 which is χ^2 with 7 degrees of freedom, respectively.

[15] The significance of the constant term (b_k) in the log-normally distributed coefficients is not reported since the only reasonable test of significance for this parameter is a test of negative infinity. In any case, the standard errors of these estimated coefficients are small compared to the parameters.

relatively a much higher mean effect on the utility derived from a bus trip than the *Quality program*.

For the car mode, *Travel time* and *Cost per trip* are the significant determinants of the mode choice. It is surprising that *Parking cost* is not significant, given our discussion in section 2. A possible explanation for this result is that the levels presented in the experiment might have been regarded as too low. The definition of the levels for the parking attribute proved to be an evasive task throughout the construction of the experiment. Although parking in the city can be regarded as free, there is an informal system of car-watchers who charge a fee for their service. For the bus mode all attributes except for the *Cost per trip* and the *Quality program* are significant determinants of the mode choice, although *Number of departures* is not significant in the normal model. This indicates that the important characteristic of the bus mode is the overall travel time.

An important variable in our model is the state dependence variable, which is highly significant. This variable captures a strong inertia in the use of the car for those individuals already using it, and it has important policy implications. Another way in which we model heterogeneous preferences is by including socio-demographic information about the respondents. As expected, those individuals who sometimes need the car at work are more likely to choose the car mode in the choice experiment. In the MNL model, both the income and the age of the respondent are highly significant. This significance is reduced in the RPL models. The positive coefficient for income indicate that higher income individuals are less likely to take the bus, since they experience a higher cost of traveling time.

7. Analysis of Results

In this section we explore the responsiveness of modal choice to changes in the attributes. We calculate the aggregate elasticities and marginal effects for each attribute and estimate the value-of-time for each mode. The estimates of the value-of-time

[16] For large samples, and under quite general conditions, the sample mean of a sequence of random variables converges to a normal distribution even though the parent distribution is not normal. A one-tail t-student test can therefore be applied on the means to test the hypothesis $\beta^{mean} = 0$, versus $\beta^{mean} > 0$.

provide information about the potential gains for the travelers from the policies. These benefits may actually be as important as the benefits from reduced emissions.

Since the two different RPL models give similar elasticities and marginal effects, we only report the results for the model with the attributes having a lognormal distribution. The aggregate elasticities are computed as a weighted average of the individual elasticities using the choice probabilities as weights and the marginal effects are also weighted by the choice probabilities[17] (see Ben-Akiva and Lerman, 1985; Bath, 1998) Table 4 reports the estimated elasticities and marginal effects. The elasticities are the percentage change in choice probability for a percentage change in the corresponding attribute, and the marginal effects are the marginal change in choice probability for a marginal change in the corresponding attribute.[18]

Table 4: Average elasticities and marginal effects (times 100) for the car and bus modes. RPL model with Lognormal distribution

	Average Elasticities		**Average marginal effects**	
	Direct (Car)	Cross (Bus)	Direct (Car)	Cross (Bus)
Increase in travel time for the car mode	-0.093	0.319	-6.854	8.282
Increase in cost per car trip	-0.064	0.219	-4.613	6.027
Increase in parking costs for car	-0.020	0.069	-1.516	1.673
	Cross (Car)	Direct (Bus)	Cross (Car)	Direct (Bus)
Decrease in travel time for the bus mode	-0.136	0.468	-9.855	12.813
Decrease in cost per bus trip	-0.014	0.049	-1.055	1.239
Increase in punctuality	-0.059	0.202	-4.280	5.461
Increase in frequency	-0.047	0.161	-3.524	3.964
Increase in quality	-0.006	0.022	-0.467	0.551

The elasticities and marginal effects are generally low, mainly due to the effect of state dependence and the limited adjustment to the policies permitted in our study, due to the assumptions of given number of trips, and fixed origin and destination. This is expected, and is in line with other similar studies (see for example Bhat, 2000, 1998; Swait and Eskeland, 1995). Travel time for bus and car have the highest elasticities and marginal effects. A hypothetical 10% decrease in average travel time by bus

[17] Furthermore, since the marginal effects are probability weighted they will not in general sum to one.
[18] Parking, punctuality and quality are dummy variables, so the reported elasticities and marginal effects are only approximations.

(corresponding to an average reduction of 6 minutes) would reduce the probability of car use by 1.36%.

The elasticity and marginal effect of the cost per car trip is much higher than the elasticity and marginal effect of the cost per bus trip, and parking cost is less important than both cost per car trip and per bus trip. Punctuality and frequency of the bus service both have a larger impact on the choice of mode than the Quality Improvement Program. This is also an indication of the importance of travel time, since frequency and punctuality have and impact on overall travel time.

Our analysis allows us to reach several conclusions from a policymaking perspective. First, the model shows an important inertia in travel behavior. The car mode has several advantages in terms of comfort, travel route timing flexibility and safety, among others, and our results confirm that breaking the travel pattern of a car user is difficult. A reduction in travel time for the bus mode emerges as the most important element in a program aimed at attracting commuters towards public transport and away from the private mode. Consequently, measures such as exclusive bus lanes, faster and more accurate connections and traffic light priority, can have an impact on the use of private transport. On the other hand, subsidized bus fares seem to have very low effectiveness according to our analysis. Subsidies should instead be linked to better service, particularly regarding punctuality and frequency of departures, which further reduce overall travel time. Alternatively, the low direct elasticity of bus fares indicates the feasibility of creating a bus system that is more expensive, but faster and with better service, which would target the middle class travelers sampled in our study. Of the monetary incentives that could be used to discourage private transportation, increases in parking costs at work do not seem to be as effective as expected. Its mean effect is not significantly different than zero, and the elasticity of car use with respect to parking costs is correspondingly very low. Due to the reasons mentioned in the previous section, we prefer not to draw a strong conclusion from this result. On the other hand, contrary to public perception, increases in cost per car trip do have an effect on modal substitution, although this effect is rather small.

Finally, the (average) marginal value of time is calculated as the ratio of the coefficient for travel and the coefficient of travel cost by car. The value of time is related to the value of working time for individuals traveling to work. In our sample, the average wage per hour is around 2,000 colones. Table 5 reports the estimated average marginal values of time for all three models. The confidence intervals are based on 9,000 replications of the Krinsky-Robb (1986) method.

Table 5: Average Value of Time, 90% confidence interval

	MNL	RPL (normal)	RPL (lognormal)	
			Mean	Median
Value of time in bus (col/hour)	949 (602 - 4640)	1005 (666 - 1820)	1394 (795 - 3109)	856 (536 - 1715)
Value of time in car (col/hour)	1291 (534 - 2806)	614 (27 - 1878)	908 (441 - 3171)	827 (291 - 2952)

The differences between the two Random Parameter Models become clearer here. Due to the assumption of a log-normal distribution, the estimated mean value-of-time is higher in the lognormal model compared to the normal model, in particular for the value of time when traveling by bus. Furthermore, the confidence intervals are very large, in particular for the lognormal model. The willingness to pay for reduced travel time in both modes is high, around 40 to 50% of the average hourly wage in our sample. This is in line with many previous studies (Small, 1992; Bhat 1998), although some studies have found rather low values of time using choice experiments (see e.g. Calfee and Winston, 1998). In the Random Parameter Models, the willingness to pay for reduced travel time is higher for the bus mode than for the car mode. This result is expected since the disutility of the time spent traveling by bus is likely to be higher. These results show that reductions in travel time due to reduced congestion can have substantial benefits for our sample population. However, we also find that the estimated value of time is sensitive to the econometric specification.[19]

8. Conclusion

In general, the results indicate that mode substitution is sensitive to the characteristics and performance of each mode. In particular, travel time for both modes and travel cost for car are the most important determinants of mode choice. Our estimates of elasticities and marginal effects are rather small. This is in line with other studies, and is partly the result of their short run perspective. Since the aim is to determine which characteristics are more relevant to achieving a switch from private to public modes of transportation, we rather concentrate on the relative importance of each attribute. We therefore conclude that a program aimed at reducing congestion and pollution during peak hours should focus on increasing the cost of private transport and providing faster and more reliable public transport.[20] The possibility of separating public transport by creating a parallel service that provides a more expensive but faster service, is one potential alternative to detract customers from private transportation. As mentioned in section 3, the Costa Rican Ministry of Transport is currently redesigning the system of public transportation. Our study sheds light on the most important features required by that system if it is to attract travelers from private modes. Specifically, emphasis has to be put on designing routes and exclusive bus lanes, and providing traffic priority for buses, faster connections, and more frequent and reliable departures, among other measures intended to reduce travel time. Contrary to popular beliefs, comfort on board seems to be a much lesser relevant characteristic for commuters to work.

Finally, our estimates of the willingness to pay for reduced travel time in both modes show that potentially large benefits can be obtained from a program aimed at reducing congestion. A study of the net social value of such a program is required.

Acknowledgements

We have received valuable comments from two anonymous referees, Olof Johansson-Stenman, Peter Martinsson, Thomas Sterner, Sharon Kuhlman, Victor Castillo and seminar participants at Gothenburg University. Financial support from the Swedish International Development Agency and the Bank of Sweden Tercentenary Foundation is gratefully acknowledged

---

[19] Bhat (2000) also finds that the estimates of value of time is sensitive to the econometric specification, while Bhat (1998) and Calfee et al. (2000) finds little evidence of sensitivity of the estimates of value of time with respect to econometric specification.

[20] This conclusion is contingent on the excess capacity of the current fleet and the environmental performance of new additions to this fleet. Given the discussion in section 2, we believe a mode switch from private to public commuting will actually reduce emissions.

References

Alfaro, M. (1999), "Calidad del Aire en San José y otros Puntos del Gran Área Metropolitana: Informe 1999', Programa de Estudios de Calidad del Aire, Universidad Nacional, Costa Rica.

Arnott; R., A. De Palma and R. Lindsey (1993), "A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand", *American Economic Review*, **83:** 161-179.

Ben-Akiva, M. and S. Lerman (1985), *Discrete Choice Analysis*, Cambridge Mass.: MIT Press.

Bhat, C. (1998), "Accommodating Variations in Responsiveness to Level-of-Service Measures in Travel Mode Choice Modeling", *Transport Research A*. **32**: 495-507.

Bhat, C. (2000), "Incorporating Observed and Unobserved Heterogeneity in Urban Work Travel Mode Choice Modeling´, *Transportation Science* **34**: 228-238.

Carlsson, F. and P. Martinsson (2001), "Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments? – Application to the Valuation of the Environment", *Journal of Environmental Economics and Management* **41**: 179-192.

Calfee, J. and C. Winston (1998), "The Value of Automobile Travel Time: Implications for Congestion Polic", *Journal of Public Economics* **69**: 83-102.

Calfee, J., C. Winston and R. Stempski (2001), "Econometric Issues in Estimating Consumer Preferences from Stated Preference Data: A Case Study of the Value of Automobile Travel Time", *Review of Economics and Statistics* **83**: 699-707.

De Borger, B., I. Mayeres, S. Proost, and S. Wouters (1996), "Optimal Pricing of Urban Passenger Transport: A Simulation Exercise for Belgium", *Journal of Transport Economics and Policy* **30**: 31-54.

Echeverría, J. and G. Solórzano (2000), "Impacto Ambiental de la Importación de Vehículos en Costa Rica: Un Análisis Preliminar" Centro Científico Tropical, Costa Rica.

Eskeland, G. and S. Devarajan (1996), "Taxing Bads by Taxing Goods: Pollution Control with Presumptive Charges", Directions in Development, The World Bank, Washington D.C.

Eskeland, G. and E. Jimenez (1992), "Policy Instruments for Pollution Control in Developing Countries´, *The World Bank Research Observer* **7**: 145-170.

Fullerton, D. and S. West (2001), "Can Taxes on Cars and on Gasoline Mimic an Unavailable Tax on Emissions?", *Journal of Environmental Economics and Management* , forthcoming.

Harrington, W., Krupnick, A. and A. Alberini (1998), "Overcoming Public Aversion to Congestion Pricing", Discussion Paper 98-27, Resources for the Future.

Jiménez, G. (1997), "Análisis del Impacto del Ecomarchamo", School of Civil Engineering, University of Costa Rica.

Krinsky, L. and A.L. Robb (1986), "On Approximating the Statistical Properties of Elasticities", *Review of Economics and Statistics* **68**: 715-719.

Kuhfeld, W. (2001), "Multinomial Logit, Discrete Choice Modeling. An Introduction to Designing Choice Experiments, and Collecting, Processing and Analyzing Choice Data with SAS", SAS Institute TS-643.

Ministry of Health (1997), "Resultados de la Encuesta sobre Principales Causas de Consulta Externa en Costa Rica", Ministry of Health, Costa Rica.

Ministry of Public Infrastructure and Transport (1999), "Reorganización del Transporte Público Colectivo en el Área Metropolitana de San José", Ministry of Public Infrastrucutre and Transport, Costa Rica.

O'Connor, D. (1999), "Applying Economic Instruments in Developing Countries: From Theory to Implementation", *Environment and Development Economics* **4**: 91-100.

Proyecto Estado de la Nación (1999). "Estado de la Nación", San José, Costa Rica.

Pujol, R. (1996), "The Spatial Structure of Urban and Metropolitan Regions as a Critical Variable in the Long Run Reduction of Urban Air Pollution", ProDus, Costa Rica

RECOPE (1999), "Actualización de la Información del Parque Automotor Costarricense", Dirección de Estudios de Demanda Energética, Costa Rica.

Revelt, D. and K. Train (1998), "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level", *Review of Economics and Statistics*, **80**: 647-657.

Small, K.A. (1982), "The Scheduling of Consumer Activities: Work Trips", *American Economic Review* **72**: 467-479.

Small, K.A. (1992), *Urban Transportation Economics*, Luxembourg: Harwood Academic Publishers.

Sterner, T. (1994), "The Price of Petroleum Products", in Sterner, T., ed., *Economic Policies for Sustainable Development*, Kluwer Academic Publishers.

Swait, J. and G. Eskeland (1995), `Travel Mode Substitution in Sao Paulo: Estimates and Implications for Pollution Control´, Policy Research Working Paper 1437, The World Bank.

Thomson, J.M. (1998), `Reflections on the Economics of Traffic Congestion, *Journal of Transport Economics and Policy* **32**: 93-109.

Train, K. (1998), "Recreation Demand Models with Taste Differences over People", *Land Economics* **74**: 230-239.

Train, K. (1999), "Halton Sequences for Mixed Logit", Working Paper Department of Economics, University of California, Berkeley.

Willson, R. and D. Shoup (1990), "Parking Subsidies and Commuter Mode Choice: Assessing the Evidence", *Transportation* **16**: 53-69.

How Much do We Care About Absolute Versus Relative Income and Consumption?

Francisco Alpizar
Fredrik Carlsson
Olof Johansson-Stenman

Abstract
We find, using survey-experimental methods, that most individuals are concerned with both relative income and relative consumption of particular goods. The degree of concern varies in the expected direction depending on the properties of the good. However, contrary to what has been suggested in the previous literature, we find that relative consumption is also important for vacation and insurance, which are typically seen as non-positional goods. Further, absolute consumption is also found to be important for cars and housing, which are widely regarded as highly positional. Implications for Pareto-efficient taxation are illustrated using the results from the experiment.

Keywords: status, relative income, optimal taxes, experiments
JEL-classification: C91, D63, H21

1. Introduction

There is considerable evidence that relative income and consumption are important determinants of both individual well-being and behavior, and even Adam Smith in his day noted that women in England required better clothing to appear in public without shame than women in Scotland did. Seminal contributions include Veblen (1899) who used the term conspicuous consumption to refer to expenditure in goods that signal the consumer's position in society, and Duesenberry's (1949) theory of consumption, which emphasized the importance of relative standings in determining consumption and savings patterns over time. Weiss and Fershtman (1998), Holländer (2001) and Brekke and Howarth (2002) provide good overviews of more recent literature on and implications of social status.

Still, the issue of *how* important relative income and consumption are for individual well-being, compared to their absolute counterparts, is much more difficult to answer, and this is the main concern of this paper. On one extreme, standard economic theory typically assumes, based on no empirical evidence, that only absolute income and consumption matter. On the other extreme, Easterlin (1974, 1995), Oswald (1998) and others conclude that only relative income seems to matter. This conclusion is largely based on a large number of survey-based psychological studies, where it is found that happiness increases with income in a given country and in a given year, but also that average happiness in a given country seems to be roughly constant over time, even though average income increases. Furthermore, people also seem to be about equally as happy in different countries with different incomes. These studies have been criticized, however, and Brekke (1997) and Osmani (1993) argue that people may respond to questions about happiness relative to a subjective happiness norm, and that this norm may be income-dependent too. A "very happy" response in a rich society would then reflect a happier person than the same response in a poorer society. Happiness would then still depend, at least partly, on absolute income. This clearly illustrates the need for alternative empirical methods.

An alternative strategy for measuring the importance of relative standing is to ask individuals hypothetical questions regarding their choice among alternatives states or outcomes, where their choice reveals their concern for relative positions (Johansson-Stenman et al., 2002; Solnick and Hemenway, 1998). Johansson-Stenman et al. used an experiment where Swedish students made repeated choices between alternative societies, described by an imaginary grandchild's income and the average income. For example, in one society the grandchild's income was $2500/month which was lower than the average income of $3000/month. In the other society the grandchild's income was $2300 which is higher than the average income of $2000/month. The task was to decide in which society the grandchild would be most content. It was found that most people do care about relative income, but also that absolute income is in most cases important for individual well-being. Here we largely follow the same experimental design when measuring the positionality degree for income, although we apply it to a student sample from Costa-Rica.

However, if income contributes to utility solely instrumentally through consumption, as is typically assumed, then utility must depend on relative consumption. The literature on positional goods (Hirsch 1976, Frank 1985a, b) dealing with this issue, typically assumes that some goods are much more positional than others, so that relative consumption is more important for goods such as jewelry, cars and houses compared to more non-positional goods such as bread, insurance and leisure. If this is correct, important policy implications follow, as extensively discussed by Frank (1985a, b). For example, it can be socially optimal to have higher taxes on goods that are more positional, as demonstrated in Section 5 of this paper. Boskin and Sheshinski (1978) and Ireland (2001) show that the optimal amount of redistribution from income taxes increases when income is partly positional and leisure is purely non-positional, and Persson (1995) attempts to explain the very high income taxes in some countries, such as Sweden, using similar assumptions. Ljungqvist and Uhlig (2000) find arguments for a Keynesian stabilization tax policy, where it is again assumed that leisure is purely non-positional. Ng (1987) and Ng and Wang (1993) show that publicly provided goods, such as environmental quality, should be over-provided relative to the basic cost-benefit rule in the case where relative private consumption or income matters for utility, and where the publicly provided good is non-positional.

The empirical evidence for the hypothesis of varying degrees of positionality, besides introspection, is scarce, however. The only study, to our knowledge, that attempts to empirically investigate the positionality degree for different goods is Solnick and Hemenway (1998). They let American students choose between one “relative alternative” and one “absolute alternative” for each good. They found that the concern for relative standings was strongest for the students’ own attractiveness and supervisor's praise, while weakest for vacation time. They also found that on average 48% of the respondents preferred the society with a higher relative but lower absolute income, which is perfectly consistent with the findings of Johansson-Stenman et al. (2002) if parameterized in the same way. However, since Solnick and Hemenway did not ask repeated questions for each good, they did not produce any parameter estimates of the positionality degree. This is therefore done here for the first time (as far as we know), by extending the experimental design from Johansson-Stenman et al. to different goods.

Section 2 describes the models underlying the experiments on the importance of relative income and consumption. Section 3 presents the set-ups and Section 4 the results of the experiments. Section 5 derives and calculates Pareto-efficient consumption taxes based on the obtained positionality estimates. Finally, Section 6 draws conclusions.

2. Modeling Relative Consumption and Relative Income

There are many ways to incorporate relative standing into the utility function. Most studies have either used some kind of *ratio comparison* utility function, $U = v(x,r) \equiv v(x, x/\bar{x})$, where x is the individual's income (or consumption vector of different goods) and $\bar{x}$ is the average income in society (e.g. Duesenberry, 1949; Boskin and Sheshinski, 1978; Layard, 1980; Persson, 1995; Carroll et al., 1997), or some kind of *additive comparison* utility function, $U = v(x,r) \equiv v(x, x-\bar{x})$ (e.g. Akerlof, 1997; Corneo and Jeanne, 1997; Knell, 1999; Ljungqvist and Uhlig, 2000). Clark and Oswald (1998) show important theoretical differences between these formulations, but the empirical evidence is scarce. The only result we are aware of is Johansson-Stenman et al. (2002) who, in a simple test, compared the ratio comparison utility function,

$$v = x^{1-\gamma}\left(\frac{x}{\bar{x}}\right)^{\gamma}, \tag{1}$$

to the additive comparison utility function,

$$v = x - \delta\bar{x}. \tag{2}$$

Although the ratio-formulation performed better in terms of explaining respondent behavior, they concluded that more research is clearly needed. Lacking clear empirical evidence, we will present the results for both of these functional forms.[1]

It is likely that the positional concern is more complex in reality than in the stylized models most often used. For example, Knell (1999) analyses theoretically "within-class comparison", "upward comparisons" and "society-wide comparisons". The first two cases include people who care for their status as members of a specific group, and those who want to be like the ones with higher status, respectively. Nevertheless, for empirical and experimental simplicity we solely deal with the latter type of comparison, i.e. we assume that people compare themselves to the average in society, which enables us to estimate the positionality degree for each good with only one parameter. It is easy to verify that the parameters γ and δ in (1) and (2) measure the *marginal degree of positionality*,

$$\alpha = \frac{\frac{\partial v}{\partial r}\frac{\partial r}{\partial x}}{\frac{\partial v}{\partial x} + \frac{\partial v}{\partial r}\frac{\partial r}{\partial x}}, \tag{3}$$

i.e., the fraction of the total utility change which comes from increased relative consumption from the last dollar spent. Still, in future research it would be of interest to consider less restrictive formulations.

2.1. Measuring Positionality

Consider a hypothetical choice between two societies, where in society A the individual's income is \$2500 per month and the average income is \$3000, while in society B the individual's income is \$2040 per month and the average income is \$2000. In all other respects, the societies are identical. If an individual is indifferent between these two societies, we have in the case of the ratio comparison utility function that

[1] Since we are only interested in the ordinal properties of the utility function, any monotonic transformation of (1) and (2) are equally valid utility functions.

$$(x_A)^{1-\gamma}\left(\frac{x_A}{\bar{x}_A}\right)^{\gamma} = (x_B)^{1-\gamma}\left(\frac{x_B}{\bar{x}_B}\right)^{\gamma} \Rightarrow \gamma = \ln\left[\frac{x_B}{x_A}\right] \Big/ \ln\left[\frac{\bar{x}_B}{\bar{x}_A}\right] = 0.5, \qquad (4)$$

while for the additive comparison utility function we have that

$$x_A - \delta\bar{x}_A = x_B - \delta\bar{x}_B \Rightarrow \delta = \frac{x_A - x_B}{\bar{x}_A - \bar{x}_B} = 0.46. \qquad (5)$$

Hence, if an individual is indifferent between the two societies we have that $\gamma = 0.5$ (or that $\delta = 0.46$). Consequently, if he/she prefers society A then $\gamma < 0.5$ (or $\delta < 0.46$), and vice versa. By letting individuals make repeated pair-wise choices between hypothetical societies with different implicit marginal degrees of positionality, it is possible to more precisely estimate the degree of positionality for income.

We use similar experiments to measure the degree of positionality for different goods, by letting individuals make tradeoffs between their own consumption and their relative consumption, keeping everything else equal. In Appendix 1 it is shown that the marginal degree of positionality for income is equal to a weighted sum of the marginal positionality for each good, where the weights are the income shares spent on each good. This holds irrespective of the structure of relative consumption, e.g. whether ratio or additive comparisons, or a combination of these, is the correct formulation. This implies that if only relative income matters for happiness, then all market goods must have a value of α equal to one, given that α cannot be larger than one for any single good. Conversely, if α is not equal to one for all goods, it follows that absolute income as well must matter for utility. On the other hand, if α varies among the goods, i.e. is not equal to zero for all goods, it follows that utility must depend on *both* absolute and relative income.

3. Design of the Experiments

A total of 325 students from The University of Costa Rica took part in the experiment. The students were interviewed in the classroom as part of a lecture, and the average group size was 30 students.[2] The survey lasted for approximately 30 minutes, and there was no show-up fee paid. The respondents were given verbal information before each section of the experiment, in addition to printed information.

The experiment consisted of three sections: (i) the relative income experiment, (ii) the relative consumption experiment and (iii) questions regarding the respondent's socio-economic status. Following Johansson-Stenman et al. (2002), the respondents were instructed to consider *the well-being of their imagined grandchild* when making their choices. This was to help the respondents liberate themselves from their current circumstances. Further, it is possible that utility may also depend on income and consumption changes over time (i.e. positional in the time-dimension); see e.g. Frank and Hutchens (1993). They were frequently reminded that they should *not* choose what they considered the overall best society, but the society that would be the best for their grandchild. The students were told that the societies were identical in all respects, except the issue being analyzed. It was also stressed that prices and goods were the same in all societies.

Our main hypothesis is that the students will use their own preferences when responding, since they for obvious reasons have very limited information regarding their hypothetical grandchild's preferences. However, it is possible that some base their answers on their view of what matters for peoples' well being in general, i.e. on some kind of perceived average preferences, rather than on their own preferences.

3.1. The Relative Income Experiment

In the relative income experiment, the respondents make repeated choices between two societies, A and B, described by the average income and the grandchild's income. In all other respects, the societies are identical. Society A is a fixed alternative where the average income is 360,000 Colones/month and the grandchild's income is 300,000 Colones/month.[3] This society is then compared with seven different B societies with varying individual incomes but a given average income (an example is presented in Appendix 2). The grandchild's income in society B was chosen such that it corresponds to a certain degree of positionality if the individual is indifferent between the two

[2] This was a general course offered to students in several different areas including law, economics, engineering, political science, and natural sciences, among others.
[3] 1 US Dollar corresponds to 300 Colones using the exchange rate of November 2001.

societies. This is calculated both for the ratio comparison in (1) and for the additive comparison in (2). The societies are presented in Table 1, which also shows the order in which they were presented in the experiment.

Table 1: Societies in relative-income experiment, amount of Colones per month

Set		Grandchild income	Average income	Marginal degree of positionality if indifference between A and B	
				Ratio comparison, γ	Additive comparison, δ
	Society A	300,000	360,000		
1	**Society B**	300,000	240,000	0	0
2	**Society B**	285,000	240,000	0.125	0.125
3	**Society B**	270,000	240,000	0.25	0.25
4	**Society B**	245,000	240,000	0.5	0.45
5	**Society B**	220,000	240,000	0.75	0.67
6	**Society B**	210,000	240,000	0.9	0.75
7	**Society B**	200,000	240,000	1.0	0.83

If an individual chooses society B in questions 1 and 2, but society A in question 3, then the individual is willing to sacrifice 15,000 Colones/month in order to be above the average, but not as much as 30,000 Colones/month. This corresponds to a marginal degree of positionality between 0.125 and 0.25 for both the additive comparison and the ratio comparison case.

3.2. The Relative Consumption Experiment

In the experiment on relative consumption, the respondents make choices between two societies described by the average consumption and the grandchild's consumption of a particular good. Again, the societies are identical in all other respects. To test the hypothesis that the degree of positionality depends on the type of good that is consumed, we include four different types of consumption in the experiment: cars, housing, insurances and days of vacation.

But how much to spend on each good is of course up to the consumer in a free society. Therefore it is not straightforward to generalize the relative income experiment to measure the importance of relative consumption of different goods. We deal with this problem by letting the company at which the grandchild works provide the goods as a fringe benefit. For example, the grandchild could receive a car at a value of 3 million colones as a benefit. The grandchild is then still free to choose characteristics such as size and brand according to his/her taste, as long as the value corresponds exactly to the one specified for that good. For each good the respondent makes three choices between a fixed alternative A and a varying alternative B (an example is presented in Appendix 2). The construction is, in all other respects, the same as for the relative income experiment, and the choices will reveal the marginal degrees of positionality. The societies are presented in Table 2, along with the implied marginal degree of positionality if the respondent is indifferent between the two societies.

Table 2: Societies in relative-consumption experiment. Amounts in Colones except for vacation, which is in days per year

Set		Grandchild consumption	Average consumption	Marginal degree of positionality if indifference between society *A* and *B*	
				Ratio comparison, γ	Addit. comparison, δ
	Car A	4,500,000	5,400,000		
1	**Car B**	4,065,000	3,600,000	0.25	0.24
2	**Car B**	3,670,000	3,600,000	0.5	0.46
3	**Car B**	3,320,000	3,600,000	0.75	0.66
	Housing A	36,000,000	43,200,000		
1	**Housing B**	32,500,000	28,000,000	0.25	0.24
2	**Housing B**	29,400,000	28,000,000	0.5	0.46
3	**Housing B**	26,500,000	28,000,000	0.75	0.66
	Insurance A	900,000	1,080,000		
1	**Insurance B**	813,000	720,000	0.25	0.24
2	**Insurance B**	735,000	720,000	0.5	0.46
3	**Insurance B**	663,000	720,000	0.75	0.66
	Vacation A	20	25		
1	**Vacation B**	18	16	0.25	0.24
2	**Vacation B**	16	16	0.5	0.46
3	**Vacation B**	14	16	0.75	0.66

4. Results of the Experiment

4.1 Income Experiment

Of the 325 responses, 13.5% were inconsistent in the sense that they switch from choosing society A, to choosing society B in a later choice, which violates the monotonicity assumption of the utility function. Some potential explanations for such behavior are learning and fatigue effects.[4] Irrespective of the cause, we exclude these responses from the analysis. The results of the experiment are presented in Table 3.

[4] Using a binary probit model, we tried to explain the occurrence of inconsistent responses, but none of the parameters were significant.

Table 3: Results relative income experiment

Parameter values		No.	Freq.	Cum. Freq.
Ratio comparison	Additive comparison			
$\gamma < 0$	$\delta < 0$	62	0.22	0.22
$0 < \gamma < 0.125$	$0 < \delta < 0.125$	61	0.22	0.44
$0.125 < \gamma < 0.25$	$0.125 < \delta < 0.25$	12	0.04	0.48
$0.25 < \gamma < 0.5$	$0.25 < \delta < 0.45$	27	0.10	0.58
$0.5 < \gamma < 0.75$	$0.45 < \delta < 0.67$	23	0.08	0.66
$0.75 < \gamma < 0.9$	$0.67 < \delta < 0.75$	3	0.01	0.67
$0.9 < \gamma < 1$	$0.75 < \delta < 0.83$	8	0.02	0.69
$\gamma \geq 1$	$\delta > 0.83$	87	0.31	1.00

The median degree of positionality is between 0.25 and 0.5 (0.25 and 0.45 for the additive comparison case). However, the distribution of responses is almost bipolar; 44% of the respondents have a degree of positionality of less than 0.1, and 31% have a degree of positionality higher than 1.0. The mean degree of positionality is 0.45 and 0.40, respectively.[5]

In Section 2 we defined the marginal degree of positionality as the fraction of the total utility change from the last dollar spent which comes from increased relative income or consumption. Our results show that, on average, 45% of the utility increase from a small income increase arises from enjoying a higher relative income (40% for the additive formulation). This fraction is clearly lower than the hypothesis of 100% corresponding to the hypothesis that only relative income matters, but it is also higher than the 0% hypothesis stating that only absolute income matters.

[5] The individual marginal degree of positionality is set to the median value of the corresponding interval. For the extreme cases $\gamma \leq 0$ and $\gamma \geq 1$ we set the degree to - 0.0625 and 1.05 respectively, and for the extreme cases $\delta \leq 0$ and $\delta > 0.83$ we set the degree to -0.0625 and 0.95 respectively. In our theoretical model, both δ and γ are in the closed interval [0,1], but the construction of the experiment does not allow for indifference, therefore the adjustment described above.

4.2 Consumption Experiments

The results of the consumption experiments are presented in Table 4.[6]

Table 4: Results of consumption experiments

	Car			Housing			Insurance			Vacation		
	No.	**Freq.**	**Cum freq.**	**No.**	**Freq.**	**Cum freq.**	**No.**	**Freq.**	**Cum freq.**	**No.**	**Freq.**	**Cum freq.**
γ **< 0.25** δ **<0.24**	96	0.32	0.32	100	0.33	0.33	158	0.51	0.51	157	0.51	0.51
0.25 < γ **< 0.5** **0.24 <** δ **< 0.46**	40	0.13	0.45	24	0.08	0.41	27	0.09	0.60	26	0.09	0.60
0.5 < γ **< 0.75** **0.46 <** δ **< 0.66**	35	0.12	0.57	37	0.12	0.53	31	0.10	0.70	28	0.09	0.69
γ **> 0.75** δ **>0.66**	130	0.43	1.00	141	0.47	1.00	96	0.30	1.00	96	0.31	1.00
Mean values	$\overline{\gamma} = 0.55,\ \overline{\delta} = 0.50$			$\overline{\gamma} = 0.56,\ \overline{\delta} = 0.51$			$\overline{\gamma} = 0.41,\ \overline{\delta} = 0.37$			$\overline{\gamma} = 0.41,\ \overline{\delta} = 0.37$		

The median degree of positionality is between 0.5 and 0.75 (0.46 and 0.66) for cars and housing, while lower than 0.25 (0.24) for insurance and vacation. The same pattern holds for the mean degree of positionality.[7] Consequently, there are differences in the degree of positionality in the expected direction among different types of goods; the more visible goods, cars and housing, are more positional than less visible ones. At the same time it is surprising that the mean degree of positionality is relatively high even for goods such as vacation and insurance. Solnick and Hemenway (1998) asked American students to choose between 1 week of vacation for themselves given that others had none, and 2 weeks given that others had 4 weeks on average; this implies that $\delta = 0.25$ in the case of indifference. Only 20% went for the positional alternative, compared to about 50% in Table 4.

As expected, the mean degree of positionality for income is in between the corresponding values for the less positional goods, vacation and insurances, and the more positional goods, cars and housing (remember that marginal degree of

[6] The share of inconsistent responses varied between 4% and 7% in the different consumption experiments. Again, we exclude these responses from the analysis.

[7] The individual marginal degree of positionality is set to the median value of the corresponding interval. For the extreme cases $\gamma < 0.25$ and $\gamma > 0.75$ we set the degree to 0.05 and 0.95 respectively, and for the extreme cases $\delta < 0.25$ and $\delta > 0.66$ we set the degree to 0.05 and 0.86 respectively. The estimated mean is not too sensitive to this setting.

positionality for income is a weighted sum of the marginal degrees of positionality for each good, as shown in Appendix 1).

4.3 Hypothetical Bias and Identity

There are several potential explanations for the large share of "extreme" answers, besides the fact that these might reflect true preferences. It appears reasonable that some respondents used a lexicographic strategy as a cognitively cheap attempt at solving the exercise, even though their underlying true preferences may be more complex (see e.g. Payne et al. 1993). Another possibility is related to the desire to obtain a certain self-image. Akerlof and Kranton (2000) argue that individuals' self-image, or identity, is an important factor in explaining many kinds of human behavior. If individuals wish a self-image of either being concerned or not being concerned about status, lexicographic responses seem rational. Similarly, to understand a complex world we clearly need to simplify it, implying that more simplistic explanations are often attractive. Arguably, a world where individual well-being depends solely on absolute consumption, or solely on relative consumption, is easier to comprehend than a world where well-being depends on both absolute and relative consumption. Lexicographic responses are useful in maintaining or amplifying both of these views.

In Table 5, we therefore report the results of the experiments when removing two types of lexicographic responses: (i) respondents who always opted for society A, i.e. they are non-positional for all goods and (ii) respondents who always opted for society B, i.e. they are highly positional for all goods. After removing the lexicographic responses the positionality differences increase among the goods. We only report the result for the ratio-comparison utility function, but the pattern is the same for the difference-comparison case.

Table 5: Impact of removing lexicographic responses on the mean marginal degree of positionality

	Removing always non-positional	Removing always highly positional	Removing both types
Car	0.68	0.47	0.60
Housing	0.70	0.49	0.63
Insurance	0.50	0.31	0.38
Vacation	0.50	0.30	0.38

4.4 Econometric Analysis

We now turn to the question of which individual factors determine the responses in terms of positionality. This is analyzed with an ordered probit for each of the experiments. For presentational ease we group the responses from the income experiment corresponding to the responses of the consumption experiments - the qualitative results are not affected by this grouping. Since the experiment is complex, we feared that there could be order effects and enumerator effects (there were three different enumerators). Therefore, the order of the goods was shifted in the second part of the experiment. However, after testing for order and enumerator effects in the ordered probit models we cannot reject the absence of such effects, except for the housing experiment where we could not reject an interview or an order effect.

Table 6 reports the estimated marginal effects for the income and the consumption experiments.[8] A positive marginal effect for a given interval indicates that an increase in the explanatory variable increases the probability that an individual will have a marginal degree of positionality located in that interval.

[8] The models were estimated using Limdep 7.1. For details on estimation and calculation of marginal effects, see Greene (2000).

Table 6: Ordered probit marginal effects at sample means for the income and consumption experiments[9]

Variable [Mean]	γ < 0.25	0.25 < γ < 0.5	0.5 < γ < 0.75	γ > 0.75
			Income	
Female [0.478]	-0.1636***	0.0017	0.0097	0.1522***
Parents' Income [4.111]	0.0233**	-0.0003	-0.0014	-0.0216**
Left [0.399]	-0.1072***	0.0007	0.0060	0.1005**
Economics [0.097]	-0.1628***	-0.0050***	0.0044	0.1634***
Law [0.201]	-0.1877***	-0.0045***	0.0060	0.1862***
Social Science [0.302]	-0.1115***	-0.0001	0.0058	0.1060**
			Car	
Female [0.479]	-0.0556*	-0.0065**	0.0000	0.0621
Parents' Income [4.020]	0.0052	0.0006	0.000	-0.0058
Left [0.392]	-0.0025	-0.0003	0.000	0.0028
Economics [0.104]	-0.0893***	-0.0140***	-0.0032	0.1065*
Law [0.208]	-0.1105***	-0.0169***	-0.0036	0.1309**
Social Science [0.285]	-0.1057***	-0.0149***	-0.023	0.1230**
			House	
Female [0.493]	-0.1089***	-0.0079**	-0.0037	0.1206**
Parents' Income [4.035]	0.0135	0.0010	0.0005	-0.0149
Left [0.396]	-0.0711*	-0.0055*	-0.0029***	0.0795
Economics [0.101]	00013	0.0001	0.0000	-0.0015
Law [0.215]	-0.0936*	-0.0083*	-0.0057***	0.1076*
Social Science [0.292]	-0.0211*	-0.0016	-0.0008	0.235
			Insurance	
Female [0.492]	-0.0995***	0.0020	0.0088	0.0887**
Parents' Income [4.043]	0.0167	-0.0003	-0.0015	-0.0148*
Left [0.371]	0.0126	-0.0003	-0.0011	-0.0112
Economics [0.097]	-0.1423***	-0.0016	0.0076	0.1363***
Law [0.204]	-0.1322***	-0.0001	0.0086	0.1237***
Social Science [0298]	-0.1571***	0.0005	0.0109	0.1457***
			Vacation	
Female [0.493]	-0.1623***	0.0047	0.0137	0.1439***
Parents' Income [4.075]	0.0198*	-0.0006	-0.0017	-0.0175**
Left [0.378]	-0.0268	0.0008	0.0023	0.0237
Economics [0.099]	-0.1735***	-0.0019	0.0082	0.1673***
Law [0.207]	-0.0795***	0.0013	0.0059	0.0723
Social science [0.299]	-0.1015***	0.0019	0.0076	0.0920**

The results from both the income and consumption experiments suggest that women care relatively more about status than men do. One could speculate that there might be decreasing returns to status, so that status is relatively more important for low-status individuals. In a society traditionally dominated by men, such as Costa Rica, a higher status might therefore be especially beneficial for women. It is also possible, however, that the female responses reflect a more socially oriented world-view, where comparisons to others are more important, rather than differences in preferences. The same arguments can be applied to left-wing voters (defined as those who would vote for

[9] *** significant at a 1% level, ** significant at a 5% level, * significant at a 10% level

either of the two parties PLN and PFD), who appear to care more about status in income and housing than others do. Indeed, there is a tradition in socialist writings to emphasize social comparisons, expressed here by Karl Marx (1891, Ch. 6):

> A house may be large or small; as long as the neighbouring houses are likewise small, it satisfies all social requirements for a residence. But let there arise next to the little house a palace, and the little house shrinks to a hut. The little house now makes it clear that its inmate has no social position at all to maintain, or but a very insignificant one; and however high it may shoot up in the course of civilization, if the neighbouring palace rises in equal or even in greater measure, the occupant of the relatively little house will always find himself more uncomfortable, more dissatisfied, more cramped within his four walls.

We also find that the concern for status is decreasing in the parents' income (since most students do not have an income of their own, we asked for an estimate of their parents' income). For the income experiment, we obtain a positive and significant marginal effect for the interval $\gamma < 0.25$, and a negative and significant one for the interval $\gamma > 0.75$. A similar pattern holds for the consumption experiments, although most marginal effects are insignificant. There are again many possible explanations. One could argue that the parents' income, in part, is endogenous with respect to the preferences, and that the likelihood of obtaining a high income decreases with the resources spent on status-seeking behavior, and that these preferences are carried over from parents to children. However, this argument could be reversed, since high status can also be a strong motivation for becoming rich. Indeed, Mandeville (1723) in *The Fable of the Bees* has already claimed that conspicuous consumption encouraged people to work harder. Another possibility is that high-income individuals care less about comparisons to the mean income in society; instead, they might be concerned with status in a richer subgroup.

It also turns out that the students' majors reveal large differences in concerns regarding status. Students majoring in economics, law and social sciences are more concerned about status than students majoring in technology, natural sciences and other subjects. Since all of these students are in the beginning of their programs, these differences probably reflect sample-selection effects, rather than influences from the education.

5. Implications for optimal taxation

Both commodity and income taxation are typically seen as distortionary (e.g. Atkinson and Stiglitz, 1980). However, as argued by Frank (1985a), in a situation where people care about relative consumption, implying positional externalities, there are pure efficiency arguments in favor of differential commodity taxation, rather than lump-sum taxes. In this section we derive optimal tax expressions as functions of the marginal degree of positionality for different goods. We then use our empirical results to calculate Pareto efficient tax levels for the goods considered.

Assume that there are n individuals consuming g goods. The utility function for an arbitrary individual K with positional concerns can be written as:

$$U^K = v^K(x^K, r^K) = u^K(x^K, \bar{x}), \tag{6}$$

where $x^K = (x_1^K, x_2^K, ..., x_g^K)$, $r^K = (r_1^K, r_2^K, ..., r_g^K)$ and $\bar{x} = (\bar{x}_1, \bar{x}_2, ...\bar{x}_g)$ are vectors of absolute consumption, relative consumption and mean (absolute) consumption for individual K. We assume that v is quasi-concave in all elements of x and r. Under an additive formulation we have that $r^K \equiv x^K - \bar{x}$ so that $U^K = v^K(x^K, x^K - \bar{x})$, and under a ratio formulation we have that $r^K = x^K / \bar{x}$, implying that $U^K = v^K(x^K, x^K / \bar{x})$.

A Pareto efficient allocation can be found by maximizing utility for individual K while holding utility constant for everybody else, and given the budget constraint with prices normalized to one. This implies the following Lagrangean:

$$u^K(x^K, \bar{x}) - \sum_{i_{-K}} \lambda_i \left(U_0^i - u^i(x^i, \bar{x}) \right) + \mu(Y - \sum_i x^i). \tag{7}$$

The first order conditions for an arbitrary individual L and goods x_j and x_0 can be combined to:

$$\frac{\partial u^L}{\partial x_j^L} \Bigg/ \frac{\partial u^L}{\partial x_0^L} = \frac{1 + \frac{1}{n} \sum_i \frac{\partial u^i / \partial \bar{x}_0}{\partial u^i / \partial x_0^i}}{1 + \frac{1}{n} \sum_i \frac{\partial u^i / \partial \bar{x}_j}{\partial u^i / \partial x_j^i}} . \tag{8}$$

The marginal rate of substitution (MRS) between the goods should hence not equal the marginal rate of transformation (MRT), but rather an expression that is corrected for the negative externalities arising from the importance of relative consumption. Given the standard competitiveness assumption that an individual takes average consumption and prices as given, individual utility maximization implies that:

$$\frac{\partial u^L / \partial x_J^L}{\partial u^L / \partial x_0^L} = 1 + t_j . \tag{9}$$

Optimality can be achieved with a tax on good *j* taking good 0 to be an untaxed numeraire, which is natural to interpret as leisure. Combining (8) and (9) gives

$$t_j = \frac{1}{n} \sum_i \frac{\partial u^i / \partial \bar{x}_0}{\partial u^i / \partial x_0^i} - \frac{1}{n} \sum_i \frac{\partial u^i / \partial \bar{x}_j}{\partial u^i / \partial x_0^i} \equiv MEC_j - MEC_0 . \tag{10}$$

The optimal tax is simply equal to the difference in marginal external cost, MEC, between good *j* and leisure (the untaxed numeraire good) in terms of summed marginal willingness to pay. In the case where leisure is purely non-positional, (10) reduces to a standard Pigouvian externality-correcting tax.

Using (6) we can explicitly use the marginal degree of positionality to get some more insight on the optimal tax. For the additive-comparison case we have that

$$\frac{\partial u^i / \partial \bar{x}_j}{\partial u^i / \partial x_j^i} = -\frac{\partial v^i / \partial r_j^i}{\partial v^i / \partial x_j^i + \partial v^i / \partial r_j^i} = -\alpha_j^i , \tag{11}$$

where α_j^i is the marginal degree of positionality as defined in equation (3), implying that (10) can be re-written as

$$t_j = \frac{\overline{\alpha}_j - \overline{\alpha}_0}{1 - \overline{\alpha}_j}, \tag{12}$$

which is increasing in the positionality for good *j*, as expected. Hence, in the additive comparison case the optimal tax can be expressed as a simple expression of the mean degrees of positionality.

In the ratio-comparison case we have that:

$$\frac{\partial u^i / \partial \overline{x}_j}{\partial u^i / \partial x^i_j} = -\frac{x^i_j}{\overline{x}_j} \frac{\partial v^i / \partial r^i_j \, 1/\overline{x}_j}{\partial v^i / \partial x^i_j + \partial v^i / \partial r^i_j \, 1/\overline{x}_j} = -\frac{x^i_j}{\overline{x}_j} \alpha^i_j \,. \tag{13}$$

In this case, (10) can thus be re-written as

$$t_j = \frac{\frac{1}{n}\sum_i \frac{x^i_j}{\overline{x}_j}\alpha^i_j - \frac{1}{n}\sum_i \frac{x^i_0}{\overline{x}_0}\alpha^i_0}{1 - \frac{1}{n}\sum_i \frac{x^i_j}{\overline{x}_j}\alpha^i_j} = \frac{\overline{\alpha}_j\left[1 + \mathrm{cov}(\frac{\alpha_j}{\overline{\alpha}_j}, \frac{x_j}{\overline{x}_j})\right] - \overline{\alpha}_0\left[1 + \mathrm{cov}(\frac{\alpha_0}{\overline{\alpha}_0}, \frac{x_0}{\overline{x}_0})\right]}{1 - \overline{\alpha}_j\left[1 + \mathrm{cov}(\frac{\alpha_j}{\overline{\alpha}_j}, \frac{x_j}{\overline{x}_j})\right]}. \tag{14}$$

Hence, in the ratio-comparison case the optimal tax does not only depend on the mean degrees of positionality, but also on the normalized covariances between positionality and consumption of the good. The intuitive reason is that individuals' marginal willingness to pay for a reduction in average consumption of the good increases with their own consumption of the good; contrary to the additive comparison case, where the marginal willingness to pay is independent of their own consumption. The tax expression reduces to the same as that of the additive comparison case if the covariance terms are zero. To provide some insight into the magnitude of these terms, consider the following two good Cobb-Douglas utility functions:

$$u = xr^{\gamma/(1-\gamma)}y = \quad \Rightarrow \quad x^* = \frac{M}{2-\gamma}, \tag{15a}$$

$$u = x^{1-\gamma}r^{\gamma}y \quad \Rightarrow \quad x^* = \frac{M}{2}, \tag{15b}$$

$$u = x^{(1-\gamma)/\gamma}ry \quad \Rightarrow \quad x^* = \frac{M}{1+\gamma}, \tag{15c}$$

where x is consumption of a good with a marginal degree of positionality equal to γ, $r \equiv x/\bar{x}$ is relative consumption of this good, y is consumption of non-positional goods, M is individual income and x^* is the utility-maximizing demand of good x in each case when prices are normalized to unity. Assume first that all individuals have preferences according to (15a), but that γ varies in the population.[10] Disregarding income differences, this means that those who are more positional with respect to good x will also consume more of it, since $x^* = M/(2-\gamma)$ is increasing in γ. The *direct* marginal utility of consuming one more unit of x would then be the same for all with the same consumption irrespective of γ, since $MRS_{xy|r} = y/x$.The marginal *status* value of x is larger for those with a high γ, since $MRS_{ry|x} = \frac{\gamma}{1-\gamma}\frac{y}{r}$ (which increases in γ for a given consumption bundle). This case implies a positive covariance expression. One can for example consider fashion clothes, where the direct utility may be about the same for most people.

Now consider (15c) instead. Here we have that those who are more positional with respect to good x will consume less of it, since $x^* = M/(1+\gamma)$ is decreasing in γ. In this case, those with a high γ would have a lower *direct* marginal utility of consuming x compared to other goods, since $MRS_{xy|r} = \frac{1-\gamma}{\gamma}\frac{y}{x}$. The marginal *status* value of x would be the same irrespective of γ, since $MRS_{ry|x} = \frac{y}{r}$. This case implies a negative

[10] We have suppressed individual subscripts. Note that we are only interested in the ordinal properties of the utility function. Hence, every monotonic transformation of each separate utility function is equally

covariance expression. Possible examples include heavy trucks, where those who buy them on average probably do so due to high "use-values" and not because they derive more status than others from buying them. This case illustrates that a high positionality factor for a certain good may not be due to high status values, but rather to low direct use values. Equation (15b) presents an intermediate case, where consumption of x is independent of γ. In this case, a high γ reflects a combination of low use values and high status values, and the covariance expression is zero here. On average, it appears reasonable that private cars could belong to this group. Some may buy new cars due to high use values, implying a low γ (for an average status concern), while others buy them largely for the status effect.

Let us now use the results from the experiments for calculating the optimal taxes. In all cases we assume that vacation is untaxed (and it is often assumed to be untaxable too). Table 7 reports the optimal taxes for both utility functions. For the ratio-comparison case we assume, in the absence of any other information, that the normalized covariances are all zero. In this case the optimal tax *expressions* are the same for the additive and the ratio comparison cases, but the estimated mean positionality *degrees* are different.

Table 7: Optimal taxes correcting for the positional externality with and without lexicographic preferences

	All consistent responses included		**Lexicographic responses excluded**	
	Ratio comparison	**Additive comparison**	**Ratio comparison**	**Additive comparison**
Car	0.31	0.28	0.55	0.41
Housing	0.34	0.31	0.67	0.51
Insurance	0.00	0.02	0.00	0.00

The results show that relatively positional goods such as cars and housing should, based on pure efficiency considerations, have a corrective tax equal to approximately 30% of the production price, or around 50% if we exclude the lexicographic responses, while there is no need to tax relatively non-positional goods such as insurance. The tax levels obtained for housing and cars are non-negligible, without being extreme. Introducing a set of differential consumption taxes would thus imply a Pareto

valid. We can then not say anything from equations (15a-c) about cardinal properties such as the marginal utility of income, or about interpersonal utility comparisons.

improvement compared to relying solely on individual-specific lump-sum taxation. This result is based on the standard externality argument, i.e. increasing consumption for one individual implies decreased relative consumption for others.

The result would of course be more complicated in a more realistic second-best world with no possibilities of using differential lump-sum taxation. Still, the first-best optimal externality-correcting taxation has been shown to often provide a good approximation to the second-best taxation (cnf. Sandmo, 1975; Bovenberg and de Mooij, 1994; Kaplow, 1996). In addition to correcting for positional externalities among private goods, there are also arguments raised for increased redistribution due to relative-income effects. Boskin and Sheshinski (1978) and Ireland (2001) find for optimal linear and non-linear income taxes, respectively, that the marginal tax rates are typically higher in the presence of relative income effects. The basic intuition is straightforward: When utility depends on relative income, but not on relative leisure, the negative efficiency effect of an increased income tax is partly off-set by a positive externality effect. Since the optimal tax in itself is a tradeoff between efficiency and equity effects, this tradeoff will now imply more equity, which is obtained by higher taxes. In our case, we have found that leisure is also positional, but to a lesser degree compared to income. Consequently, the distributional argument by Boskin and Sheshinski and Ireland is still valid, although to a smaller extent due to the positionality in leisure. Nevertheless, more research is clearly needed both in relation to optimal taxation in a second-best world, and on measuring positionality per se.

6. Conclusions

We have found that, on average, both relative and absolute income, as well as consumption, matter for individual utility, or well-being. The differences are in the expected direction, so that goods widely considered to be positional, like houses and car ownership, are also found to be more positional than goods typically seen as non-positional, such as vacation and insurance. Income is in between, which is also expected. In addition, however, we found the mean estimates of the marginal degree of positionality for houses and car ownership to be much smaller than unity, implying that absolute consumption is still important. At the same time, the mean degree of positionality is considerable for vacation and insurance, implying that relative consumption is important here too, contrary to what is typically assumed in the literature. We do not consider this finding *per se* unintuitive. Indeed, would you be equally content with a three-week vacation if others have five weeks, as if they were to have only one week? And wouldn't you feel less secure for the same insurance level if you knew that most others are better insured than you?

However, the results from experimental and survey-based studies, including this one, should always be treated with care. One reason that goods such as cars are expected to be more positional than insurance is that cars are more visible. In the experimental context, all goods are in a sense equally visible for the respondents when they make their choices. This may imply that vacation and insurance are less positional in reality than what the results here indicate. Still, as far as we know this is the first attempt to explicitly estimate the positionality degree for different goods, and we therefore strongly welcome other empirical strategies, experimental set-ups, and samples to increase the knowledge in this important area, and also to see how robust the findings in this paper are.

Based on our empirical results, we derived and computed Pareto efficient tax levels in a first-best economy. The more positional goods, houses and cars, should then have a corrective tax of 28-67% based on the production price, depending on assumptions. Needless to say, there are many practical problems associated with detailed differential commodity taxation, as discussed e.g. by Frank (1999) and Ireland (2001). Even so, it is important to know where to begin simplifying when designing tax systems or economic policy in general.

Appendix 1: Proof of the relationship between relative income and relative consumption

Let us consider the consumer choice with *n* different goods, which are positional to a varying degree. It is convenient to write the utility functions as follows:

$$U = v(x_1,...,x_n, r_1,...,r_n) = u(x_1,...,x_n \bar{x}_1,...,\bar{x}_n), \tag{A1}$$

where v is quasi-concave in its arguments. Since we are not interested in price changes, we can normalize the consumer price for all goods to unity. It is also useful to express some (quasi-) indirect utility functions as follows:

$$U = \omega(m, r_m, a_1,...,a_n) \equiv \omega(m, r_m, \frac{\bar{x}_1}{m},...,\frac{\bar{x}_n}{m}) = V(m, r_1,...,r_n), \tag{A2}$$

where *m* is the income needed to obtain utility *U* at the relative consumption levels $r_1,...,r_n$, or at fixed relative income r_m and market shares for different goods $a_1,...,a_n$. Differentiating (A2) with respect to *m*, and using (3) for each good *j*, we get:

$$\frac{\partial \omega}{\partial m} + \frac{\partial \omega}{\partial r_m}\frac{\partial r_m}{\partial m} = \frac{\partial V}{\partial m} + \sum_j \frac{\partial V}{\partial r_j}\frac{\partial r_j}{\partial x_j}\frac{\partial x_j}{\partial m} = \frac{\partial V}{\partial m} + \sum_j \alpha_j \frac{\partial u}{\partial x_j}\frac{\partial x_j}{\partial m} = \frac{\partial V}{\partial m} + \rho \sum_j \alpha_j \frac{\partial x_j}{\partial m}. \tag{A3}$$

We have directly that $\frac{\partial \omega}{\partial m} = \frac{\partial V}{\partial m}$. From individual utility maximization, taking consumption of others as given and price normalized to 1, we have that $\rho = \frac{\partial u}{\partial x_j} = \frac{\partial v}{\partial x_j} + \frac{\partial v}{\partial r_j}\frac{\partial r_j}{\partial x_j} = \frac{\partial \omega}{\partial m} + \frac{\partial \omega}{\partial r_m}\frac{\partial r_m}{\partial m}$ at the optimum. Hence, we can re-write (A3) as:

$$\frac{\frac{\partial \omega}{\partial r_m}\frac{\partial r_m}{\partial m}}{\frac{\partial \omega}{\partial m} + \frac{\partial \omega}{\partial r_m}\frac{\partial r_m}{\partial m}} = \sum_j \alpha_j \frac{\partial x_j}{\partial m}. \tag{A4}$$

The left-hand side of (A4) is a measure of the marginal degree of positionality for income, given that the overall market shares of different goods are held constant. Thus we have:

$$\alpha_m = \sum_j \alpha_j \frac{\partial x_j}{\partial m} = \sum_j \sigma_j \gamma_j \alpha_j \ , \qquad \text{(A5)}$$

where $\sigma_j \equiv \frac{\partial x_j}{\partial m} \frac{m}{x_j}$ is the income elasticity for good *j*, and $\gamma_j \equiv \frac{x_j}{m}$ is the individual expenditure share for good *j*.

Appendix 2: Sample questions

Income experiment

Question 2. Choose between society A and B for your grandchild.

Society A: Your grandchild's income is 300.000 Colones/month
The average income in society A is 360.000 Colones/month

Society B: Your grandchild's income is 288.000 Colones/month
The average income in society B is 240.000 Colones/month

Given the prerequisites described in the introduction, choose the society you consider to be in the best interest of your grandchild, i.e. the society in which your grandchild will be most content.

❒ Society A

❒ Society B

(It is important that you focus only on the best interest of your grandchild, not on the society that is best for others or on the better society overall. Also, remember that prices are the same in both societies.)

Consumption experiment

The societies A and B are the same except for the information given below. Hence, even though other people spend more (on average) on cars in society A, their consumption of other goods is the same in society A and B.

Question 2. Choose between society A and B for your grandchild.

Society A
- The company provides a car with a market value of 4,500,000 Colones for your grandchild.
- In the society the average market value of cars is 5,400,000 Colones.

Society B
- The company provides a car with a market value of 3,675,000 Colones for your grandchild.
- In the society the average market value of cars is 3,600,000 Colones.

Choose the society you consider to be in the best interest of your grandchild, i.e. the society in which your grandchild will be most content.

❐ Society A

❐ Society B

Acknowledgements

We are grateful for valuable comments from seminar participants at Göteborg University and Resources for the Future (RFF). Financial support from the Swedish Transport and Communications Research Board (KFB), the Swedish International Development Agency (SIDA) and the Bank of Sweden Tercentenary Foundation is gratefully acknowledged.

References

Akerlof, G.A. (1997). "Social Distance and Social Decisions". *Econometrica* 65, 1005-1027.

Akerlof, G.A. and R.E. Kranton (2000). "Economics and Identity". *Quarterly Journal of Economics* 115: 715-753.

Atkinsson, A.B. and J.E. Stiglitz (1980). Lecture in Public Economics. Singapore: McGraw Hill.

Boskin, M.J. and E. Sheshinski (1978). "Individual Welfare Depends upon Relative Income". *Quarterly Journal of Economics*. 92, 589-601.

Bovenberg, A. L. and R. A. de Mooij (1994). "Environmental Levies and Distortionary Taxation." *American Economic Review* 94, 1085-89.

Brekke, K.-A. (1997). *Economic Growth and the Environment*. Cheltenham: Edward Elgar.

Brekke, K.-A. and R.B. Howarth (2001). *Affluence, Well-Being and Environmental Quality*. Cheltenham, Edward Elgar.

Carroll, C. J. Overland and D. Weil (1997). "Comparison Utility in a Growth Model". J*ournal of Economic Growth* 2, 339-367.

Clark, A.E. and A.J. Oswald (1998). "Comparison-Concave Utility and Following Behaviour in Social and Economic Settings". *Journal of Public Economics* 70, 133-155.

Corneo, G. and O. Jeanne (1997). "Conspicuous Consumption, Snobbism and Conformism". *Journal of Public Economics* 66, 55-71.

Duesenberry, J.S. (1949). *Income, Savings, and the Theory of Consumer Behaviour*, Cambridge: Harvard University Press.

Easterlin, R.A. (1974). "Does Economic Growth Enhance the Human Lot? Some Empirical Evidence", in P.A. David and M. Reder (eds.) *Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz*. Palo Alto: Stanford University Press: 89-125.

Easterlin, R.A. (1995). "Will Raising the Incomes of All Increase the Happiness of All?". *Journal of Economic Behavior and Organization* 27, 35-47.

Frank, R.H (1985a). *Choosing the Right Pond: Human Behavior and the Quest for Status*. New York: Oxford University Press.

Frank, R.H. (1985b). "The Demand for Unobservable and Other Nonpositional Goods". *American Economic Review* 75, 101-16.

Frank, R. H. (1999). *Luxury Fever*. New York: Free Press.

Frank, R.H. and R.M. Hutchens (1993) "Wages, seniority, and the Demand for Rising Consumption Profiles". *Journal of Economic Behavior and Organization* 21, 251-76.

Greene, W. (2000). *Econometric Analysis*. New Jersey: Prentice-Hall.

Hirsch, F. (1976). *Social Limits to Growth*. Cambridge: Harvard University Press.

Holländer, H. (2001). "On the Validity of Utility Statements: Standard Theory Versus Duesenberry's". *Journal of Economic Behavior and Organization* 45, 227-249.

Ireland, N.J. (2001). "Optimal Income Tax in the Presence of Status Effects". *Journal of Public Economics* 81, 193-212.

Johansson-Stenman, O., F. Carlsson and D. Daruvala (2002). "Measuring Hypothetical Grandparents Preferences for Equality and Relative Standings". *Economic Journal*, forthcoming.

Kaplow, L. (1996). "The Optimal Supply of Public Goods and the Distortionary Cost of Taxation". *National Tax Journal* 49, 513-33.

Knell, M. (1999). "Social Comparisons, Inequality, and Growth". *Journal of Institutional and Theoretical Economics* 155, 664-695.

Layard, R. (1980). "Human Satisfaction and Public Policy". *Economic Journal*, vol. 90, 737-750.

Lungqvist and Uhlig (2000). "Tax Policy and Aggregate Demand Management Under Catching Up with the Joneses". *American Economic Review* 90, 356-366.

Mandeville, B. (1723). *The Fable of the Bees; or, Private Vices, Public Benefits*. London: J Tonson.

Marx, K. (1891). *Wage Labour and Capital,* the first English translation of the pamphlet, first published (in German) in *Neue Rheinische Zeitung*, Nos. 264-267 and 269, April 5-8 and 11, 1849.

Ng, Y.-K. (1987). "Relative Income Effects and the Appropriate Level of Public Expenditure", *Oxford Economic Papers* 39, 293-300.

Ng, Y.-K. and J. Wang (1993). "Relative Income, Aspiration, Environmental Quality, Individual and Political Myopia - Why May the Rat-Race for Meeting Growth be Welfare-Reducing?". *Mathematical Social Sciences* 26, 3-23.

Osmani, A.J. (1998). "Happiness and Economic Performance". In (M. Nussbaum and A. Sen, eds) *The Quality of Life*. Oxford: Oxford University Press.

Oswald, A.J. (1998). "Happiness and Economic Performance". *Economic Journal* 107, 1815-1831.

Payne, J., J. Bettman and E. Johnson (1993). The Adaptive Decisionmaker. Cambridge: Cambridge University Press.

Persson, M. (1995), "Why are Taxes so High in Egalitarian Societies?" *Scandinavian Journal of Economics* 97, 469-76.

Sandmo, A. (1975). "Optimal Taxation in the Presence of Externalities." *Swedish Journal of Economics*. 77, 86-98.

Solnick, S. and D. Hemenway (1998). "Is More Always Better?: A Survey on Positional Concerns". *Journal of Economic Behavior & Organization* 37, 373-83.

Veblen, T. (1899). *The Theory of the Leisure Class*, New York: MacMillan.

Weiss, Y. and C. Fershtman (1998). "Social Status and Economic Performance: A Survey". *European Economic Review* 42, 801-820.

Collective versus Random Fining: An Experimental Study on Controlling Non-Point Pollution

Francisco Alpízar

Till Requate[a]

Abstract

This paper presents an experimental study of two different pollution compliance games: collective *vis-à-vis* random fining as a means to regulate non-point pollution. Using samples from both Costa-Rican coffee mill managers and Costa-Rican students, we confirm the hypothesis that the two games are equivalent and, in the majority of cases, lead to efficient outcomes through Nash play. However, we reject the hypothesis that managers and students behave equally. Off the equilibrium, managers tend to over-abate, whereas students tend to under-abate. This result suggests the importance of considering subject pool differences in the evaluation of environmental policies by means of experiments, particularly if those policies involve certain forms of management decisions.

Keywords: Non-point pollution, environmental regulation, experimental economics, subject pool.

JEL: B4, C9, Q28, H2,

1. Introduction

Economic theory provides a wide array of policy instruments such as emission standards, effluent charges and tradable permits to control pollutants from point sources. The problem of efficient pollution abatement turns out to be more difficult under circumstances of non-point pollution. In that case the general problem is that only ambient pollution levels, albeit not the discharges of any individual polluter, can be observed. A common suggestion to deal with this problem is the use of target-based mechanisms like forcing contracts. For example, all potential polluters can be collectively punished if ambient pollution exceeds a certain threshold, exogenously

[a] Interdisciplinary Institute of Environmental Economics, Heidelberg University; email: requate@mail.eco.uni-heidelberg.de

defined by the regulatory authority. Such a mechanism has been suggested by Segersen (1988) and independently by Meran and Schwalbe (1987). Alternatively, it would also be possible to subsidize the firms according to the marginal benefit of abatement. A special characteristic of mechanisms like these is that they are not budget balancing in the sense that the sum of charges exceeds the social damage, or the sum of subsidies paid to the firms exceeds the social benefit.

By contrast, Xepapadeas (1991) suggests a mechanism with subsidies per unit of abatement and random fines. If ambient pollution exceeds the socially optimal level or some other fixed threshold, one of the suspected polluters is picked out randomly and charged a fine, irrespective of being one of the true polluters. Thereafter, the fine is redistributed to the other suspects. Xepapadeas calls this mechanism budget balancing, since the total payments to/from dischargers in case of deviations between desired and observed ambient concentrations, equal the corresponding society's valuation of reduced abatement. In addition to the random fining scheme, Xepapadeas also discusses a non-budget-balancing mechanism based on subsidies and collective penalties, similar to Meran and Schwalbe and Segersen.

From a more general point of view the underlying problem is a typical problem of moral hazard in teams, which was first treated by Holmström (1982) in a general principal-agent environment. Holmström demonstrates that under fairly general conditions, particularly for risk neutral agents, no budget-balancing mechanism exists which implements the efficient outcome in Nash equilibrium. Kritikos (1993) and Herriges et al. (1994) point out that Xepapadeas' results regarding the efficiency of the random fining mechanism are in contradiction with Holmström's findings, since risk neutrality is an assumption in his paper. Kritikos shows, however, that Xepapadeas' idea of random fining can be restored if the revenues from fining are not redistributed to the other suspects, i.e. by giving up budget balancing. In contrast, referring to Rasmussen's (1988) work on risk averse agents, Herriges et al. (1994) argue that a budget balancing scheme of random fining can be maintained if polluters are risk averse.

In this paper we present an experimental study of both mechanisms proposed by Xepapadeas (1991) for cases where only ambient pollution levels can be observed. The main issue is the analysis of the players' behavior under collective *vis-à-vis* random fining, where the expected pay-off in both mechanisms is identical. We designed two non-cooperative games that closely reflect the proposed mechanisms. One of the games reflects the non-budget-balancing mechanism, where firms face collective penalties if ambient pollution exceeds the desired target. The second game mimics the mechanism, where one firm is randomly chosen to bear the fine whenever measured ambient pollution is not optimal. Following Kritikos, we did not incorporate budget balancing, i.e. we did not refund the fines to other players, mainly because we are interested in the response of agents to the scheme, and because risk neutrality cannot be excluded for small stakes.

We run the experiments with two different pools of subjects. Our first sample consists of managers of coffee mills in Costa Rica. Coffee milling is associated with water pollution because water is used to peel and wash the fruit. Although the pollution of rivers due to coffee milling is not a typical non-point pollution problem, the fact that several mills generally share the same river basin, and the agency's lack of funds to monitor each mill individually, added realism to the experiments. In any case, our experiment was not constructed for any particular industry, and we view our sample of managers as a representative sample of this type of individual in any given industry. Our second sample consists of Costa Rican students,[1] where the participants were subjected to the very same treatment that was applied to the sample of managers.

Due to sample size constraints in the sample of managers, we were only able to form "teams" of two, i.e. there were only two potential polluters. Although in practice two polluters obviously do not constitute a non-point pollution problem, the fact that individual pollution is not observable is a central element of the design of this controlled experiment. In addition, even if we form larger groups each player will obtain information about the behavior of the other players, based on his/her own payoffs. Still we acknowledge that larger groups provide less detailed information about the behavior of the "rest of the team" and therefore might better capture a non-point pollution situation.

According to Smith (1982), the general view is that the evaluations of the performance of institutions, obtained from laboratory tests, should also apply to non-laboratory settings where similar conditions hold. Still, some authors have argued that different subject pools might introduce different conditions into the experiments, some of which are not possible to control for in the design (Coller and Harrison, 1996). For example, students, although readily available and cheap to work with, might have a particular sociological behavior (Cunningham, 1974), might use different heuristics to solve the exercise (Friedman and Sunder, 1994), and exhibit different degrees of risk aversion than, for example, firm managers. This calls for special caution when generalizing the results from these experiments. We believe that our samples provide a unique opportunity to explore differences in the behavior of convenience samples and industry people, who are involved in decisions about discharging pollutants.

The two samples of managers and students played several rounds of the collective fine game first, and several rounds of the game with random fining second. Since both games can be designed such that their Nash equilibria induce the socially optimal outcomes, our first conjecture is that the two mechanisms lead to Nash play. Due to the small sample size and the categorical nature of the data, a formal statistical analysis within game play was precluded. Hence, we looked at the frequency of Nash play compared to that of other strategies. We find that the frequency of Nash outcomes is higher than the frequency of any other pair of strategies. In other words, the games lead to Nash play if we employ the criterion of simple majority. In the games played second, we even observe Nash play in the absolute majority of cases. Still, we acknowledge that criteria such as simple or absolute majority could be regarded by some as rather weak.[2]

Since the two games are theoretically equivalent if the players are risk neutral, our next, and stronger, hypothesis is that the behaviors of the players are equal in both types of game. This hypothesis seems to be rejected at first glance. Since it cannot be excluded, however, that the differences in play are due to order effects, we tested for them by convoking a second sample of students who played the two series of games in the reversed order; that is, they played the game with random fining first, and the collective fine game second. Unfortunately, it was not possible to convoke managers for a similar test. We employ the two student samples to compare the games after

[1] We conducted the experiments with students from Costa Rica instead of recruiting students from Göteborg or Heidelberg in order to exclude cultural differences.

[2] If all players predict that a particular Nash equilibrium will occur, then no player has an incentive to play differently, i.e. in theory one would expect a frequency of 100% for the Nash strategy. Accepting that some players will not play their Nash strategy implicitly also accepts that some players will make a mistake at some stage of their optimizing process. Such mistakes might be likely in some situations, since the outcome of the game depends on more information than is provided in the strategic form, e.g. experience, culture and expectations about the game (Fudenberg and Tirole, 1998). Hence, we prefer to refer to conjectures rather than hypothesis testing.

accounting for order effects. By using a Chi-square test, we cannot reject the hypothesis that the two games of collective and random fining are equivalent. Applying, once again, a Chi-square test to the two student samples, we reject the hypothesis that the order in which the games are played is irrelevant.

Finally we tested for differences in the subject pool under the hypothesis that the behavior of managers is equivalent to the behavior of students. Employing another Chi-square test, we reject this hypothesis. Looking at the outcome tables in more detail, we even find that off the equilibrium, managers tend to over-abate, whereas students tend to under-abate. Furthermore, we find that managers and students react differently to the random fine.

The paper is organized as follows: in Section 2 we briefly set up the theoretical background of the games of collective and random fining. In Section 3 we describe the design of the experiment and state our hypotheses. Section 4 presents the results of our analyses. In Section 5 we draw our conclusions.

2. Theoretical Background

Following Xepapadeas' (1991, Section 3) simple framework, we assume a finite number n of identical polluters, where A_i denotes the abatement level of polluter n and $C(A_i)$ his or her abatement cost function, which is assumed to be increasing and convex. If λ denotes the (constant) marginal damage of pollution, the social benefit of abatement is determined by

$$SB = \lambda \sum_i A_i \ . \tag{1}$$

The regulator only observes ambient pollution, given by

$$W = \overline{W} - \sum_i A_i \ , \tag{2}$$

where $\overline{W}$ is the unregulated ambient pollution level. A first best allocation requires $C'(A_i)=\lambda$, i.e. marginal abatement cost equals marginal social benefit. Let A^* denote the socially optimal abatement level of each firm, which by symmetry is unique and identical for each firm. Segersen suggests a tax/subsidy scheme of the following kind:

$$\Pi_i(A_1,...,A_n) = \Pi_i^o - C(A_i) + \sigma\left(\sum_{i=1}^{n} A_i - nA^*\right), \tag{3}$$

where σ is the rate of subsidy paid to each firm in case there is collective over-abatement. In case of collective under-abatement, firms are taxed by the rate σ. It is easy to see that the strategy profile where each firm abates the optimal level A^* is a Nash equilibrium, where the first order condition of each firm in equilibrium satisfies $C'(A_i) = \sigma$. Clearly if σ is substituted by σ/n, i.e. each firm does not bear the full marginal cost of pollution, but a corresponding share of it, then compliance with A^* does not form an equilibrium.

Xepapadeas explores two mechanisms in which the agency subsidizes abatement in case of collective compliance, but fines the firms in case of collective under-abatement, distinguishing between collective and random fines.

2.1 Game with collective penalties

Under collective penalties, firms are punished collectively if the regulator observes aggregate under-abatement. Each player receives a uniform share θ of the total subsidy minus the fine F^c. If there is optimal or over-abatement, the firms receive the corresponding share of the subsidy. Thus the profit function in this game with collective penalties is given by the following scheme:

In case of under-abatement, i.e. if

$$\sum_{i=1}^{n} A_i < nA^*, \tag{4}$$

then we get

$$\Pi(A_i) = \Pi^0 - C(A_i) + \theta \cdot \{\sigma \cdot \sum_i A_i - F^C\}. \tag{5}$$

In case of aggregate compliance or over-abatement, i.e. if

$$\sum_{i=1}^{n} A_i \geq nA^*, \tag{6}$$

then we get

$$\Pi(A_i) = \Pi^0 - C(A_i) + \theta \cdot \sigma \sum_i A_i . \tag{7}$$

Compliance is an equilibrium if and only if

$$\Pi^0 - C(A^*) + \theta\sigma nA^* > \Pi^0 - C(\widetilde{A}) + \theta\left[\sigma\left[\widetilde{A} + (n-1)A^*\right] - F^c\right] \tag{8}$$

for any $\widetilde{A} \neq A^*$. The incentive compatibility condition in Equation (8) is equivalent to

$$F^c > \frac{1}{\theta}\left[C(A^*) - C(\widetilde{A})\right] + \sigma\left[\widetilde{A} - A^*\right] \text{ for each } \widetilde{A}. \tag{9}$$

Obviously, F^C can be chosen sufficiently large to induce compliance with the optimal abatement level by each firm as the Nash Equilibrium. It is important to note that compliance to the socially optimal outcome (A*,...,A*) is not the only equilibrium. Any other strategy profile (A_1,...,A_n) with $\sum_{i=1}^{n} A_i = nA^*$ is also an equilibrium as long as inequality (9) holds for the firm with the highest A_i ,i.e. with the highest abatement cost. However, if firms are symmetric, the socially optimal outcome seems to be a focal point.

2.2 Game with random fines

In this game, if firms do not abate optimally, one of the players is picked randomly with probability ξ and has to pay the exogenously determined fine F^R. The rest of the players receive the subsidy corresponding to the observed total abatement. The profit function in this game is given by the following conditions.

In case of under-abatement, i.e. if (4) holds, we have

$$\Pi = \Pi^0 - C(A) - F^R \text{ with Prob } \xi,$$

and (10)

$$\Pi = \Pi^0 - C(A) + \theta_{n-1}\sigma^* \sum_i A_i \text{ with Prob } 1-\xi,$$

where θ_{n-1} is the share of social benefits distributed among the other firms.

In case of aggregate compliance or over-abatement, i.e. if (6) holds, we have

$$\Pi(A_i) = \Pi^0 - C(A_i) + \theta\sigma \sum_i A_i$$

Note that in the case of only two players, the firm that is not fined receives the full benefit from the subsidy. The only difference with the payoff functions in Xepapadeas (1991) is that the fine F^R is not returned to the lucky firm.

In case of aggregate compliance or over-abatement, the expected profit of firm *i* is given by

$$E\Pi = \Pi^0 - C(A_i) + \theta\sigma \sum_{i=1}^n A_i ,$$

and for the case of non-compliance by

$$E\Pi = \Pi^0 - C(A_i) + \xi\left[-F^R\right] + (1-\xi)\left[\theta_{n-1}\sigma \sum_{i=1}^n A_i\right]. \tag{11}$$

For the risk neutral firm, compliance to the socially optimal action is an equilibrium if and only if

$$-C(A^*) + \theta\sigma n A^* > -C(\widetilde{A}) - \xi F^R + (1-\xi)\theta_{n-1}\sigma\left[\widetilde{A} + (n-1)A^*\right] \quad \text{for any } \widetilde{A}. \tag{12}$$

It is obvious again that F^R can be chosen sufficiently large to guarantee that the strategy profile where each firm complies with the socially optimal outcome is an equilibrium. Again, note that this equilibrium is not unique.

3. Design of the experiment, hypotheses and application

3.1 Design of the experiment

The games were constructed to be played repeatedly in a non-cooperative setting with two players (*A* and *B*). In principle, both games could be designed to be played by more than two individuals, given some minor adjustments. We decided to have two players, mainly due to limitations in sample size, particularly for the managers. We acknowledge that larger groups might introduce new interesting features, particularly with regards to the information available to the participants.

Throughout the experiment we set the default profit to $\Pi_i^o = 34$. Table 1 contains our assumed abatement cost schedule, which is the same for both players in both games.

Table 1: Abatement Cost Schedule

Abated Emissions	Marginal Abatement Costs (MA)	Total Abatement Costs (AC)
0	0	0
1	20	20
2	40	60
3	60	120

The social benefit per unit of abatement is assumed to be constant and equal to 50. Then, in the social planner's optimal solution, each polluter should abate 2 units in each period (round of the game). Next we set the subsidy σ =50 per unit of reduction in ambient pollution,[3] in combination with a fine $F^C=F^R=34$ if ambient pollution is not reduced by a total of 4. In this setting, each firm has full information about its pay-off function and that of the other firms in the industry (firms are homogeneous). In addition, the observed ambient concentrations provide information about the behavior of the "other" firms in each monitoring round.

[3] We assume a deterministic one to one relation between firms' abatements and reductions in ambient pollution.

3.1.1 Game with collective penalties

In each round of this game, the two players (firms) are punished collectively by F^C=*34* if there is under-abatement. Each player receives the corresponding uniform share θ=*1/2* of the total subsidy minus the fine. Table 2 provides the payoff matrix resulting from this profit function.

Table 2: Payoff matrix for game with collective penalty

Abated Emissions		**Player B**			
		0	1	2	3
Player A	0	**17** **17**	**22** **42**	**7** **67**	**-28** **92**
	1	**42** **22**	**47** **47**	**32** **72**	**14** **114**
	2	**67** **7**	**72** **32**	**74** **74**	**39** **99**
	3	**92** **-28**	**114** **14**	**99** **39**	**64** **64**

3.1.2 Game with random fines

In this game, if firms do not abate optimally, one of the players is picked randomly with probability $\xi = 0.5$ and has to pay the fine, F^R=34. Moreover, we choose $\theta_{n-1} = 1$. The payoffs are presented in Table 3. Each cell of the matrix contains three possible payoffs depending on total observable abatement. The smallest (top) value is the result of being randomly selected to pay the fine. The highest (bottom) value is the result of not being fined and receiving the full subsidy. If $\sum_i A_i > 4$, both values are the same since no fine is imposed. The middle value is the expected payoff for choosing a certain amount of emission reduction.

Table 3: Payoff matrix for game with random fines

Abated Emissions		**Player B**			
		0	1	2	3
		0 **17** 34 0 **17** 34	-20 **22** 64 0 **42** 84	-60 7 74 0 67 134	-120 -28 64 0 92 184
	1	0 **42** 84 -20 **22** 64	-20 **47** 114 -20 **47** 114	-60 **32** 124 -20 **72** 164	14 14 14 114 114 114
Player A	2	0 **67** 134 -60 **7** 74	-20 **72** 164 -60 **32** 124	74 **74** 74 74 **74** 74	39 39 39 99 99 99
	3	0 92 184 -120 -28 64	114 114 114 14 14 14	99 99 99 39 39 39	64 64 64 64 64 64

3.2 Hypothesis

Both games described above capture efficient mechanisms between the regulator and a polluting industry for the case where individual emissions are unknown to the agency. An efficient mechanism "…induces the dischargers to adopt optimal abatement policy in the absence of effective individual monitoring by the agency. The agency can monitor only whether the dischargers as a group follow the optimal policy, by monitoring the deviations between desired and observed ambient concentrations" (Xepapadeas, 1991, p.120). Since both games are theoretically equivalent for risk neutral players,[4] our first hypothesis is that both games successfully achieve the optimal pollution reduction ($A_i=2,\ i=1,2$) and that there are no differences between the games.

Our second hypothesis concerns the behavior of the student sample versus the sample of coffee mill managers, i.e. a more realistic sample of potential affected parties from this type of regulation. Most experimental studies use students as a representative

[4] If the agents were risk averse, the fine chosen for the random game would provide a stronger incentive to comply with the regulation.

sample, and use the conclusions from such studies to draw conclusions about the behavior of firms (Smith, 1982; Plot,[5] 1982). Hence, our hypothesis is that both of our samples should render similar results for both games in terms of pollution reduction and Nash behavior.

3.3 Application of the games

The games were applied to three different samples. One sample consisted of 16 pairs of coffee mill managers in Costa Rica (i.e. a total of 32 individuals), who participated in a one day seminar organized by ICAFE (Coffee Institute,[6] Costa Rica). The participants were told in advance that part of the seminar, approximately 2 hours, was going to be dedicated to experimental sessions that involved non-negative monetary payments. The second sample consisted of 21 pairs of students from the University of Costa Rica who responded to the advertising of the games, which was posted all around the university campus. The students were convoked for a two hour session of games involving non-negative payments. In both of these cases, the two games discussed here were part of a larger set of exercises, which included trust experiments. Approximately 30 minutes of the two hours were used for the two games described in this paper. Finally, the third sample consisted of 16 pairs of students from the same university, which were convoked in the same way as above. This session lasted for approximately one hour, and was intended to allow for testing of order or learning effects.

All sessions were conducted in large rooms that allowed us to physically separate the group of players A from players B.[7] The participants were arbitrarily assigned to each group. In the session with students, the participants who came accompanied were assigned to the same group. Additionally, we repeatedly requested that there should be no communication during the experiments.[8]

---

[5] "General theories must apply to special cases. […] Theories which do not apply to the special cases are not general theories and cannot be advocated as such" (Plot, 1982, p.1522).

[6] This institute is the organizing body in charge of coffee production and processing, as well as regulation of the industry.

[7] A pair of players always consisted of a member of group A and a member of group B.

[8] This request was particularly difficult for the sample of coffee mill managers —mostly senior businessmen not used to being silenced. We downgraded our request to not discussing the games or the strategies under the promise of a later explanation. We believe that this strategy was more successful, and we never heard discussions about the optimal strategies. Nevertheless, this possibility cannot be rejected. Note that these were within player-type conversations. The groups of players were physically separated.

Each game was carefully explained in Spanish and the experimenters provided several examples, putting care into not biasing the results by choosing specific numbers. Special emphasis was added to the fact that the regulator was unable to monitor individual emissions and was therefore limited to ambient emissions monitoring. The actual game did not start until all questions were answered and the researchers were fully satisfied with the level of understanding. The same individuals conducted the experiments for all samples, following the same script.

Another important element of the games was the monetary compensation. We explained that there would be a monetary payment according to the results of the games. We also explained that one of the rounds was going to be randomly selected for payment, and that each person would be paid individually and privately. Beattie and Loomes (1997) called this practice a *random problem selection procedure,*[9] which is intended to encourage the respondents to treat each problem/round in isolation. In this way, the researcher is able to control for income effects, and, most importantly, each response can be treated as independent of the others, hence allowing for "direct within-subjects tests" (p.156). Finally, given the differences in income level between our samples, we decided to use a different "exchange rate" for each sample. For the case of coffee mill managers, we decided to have an exchange rate of 12 units per US$1, and for the students this rate was 120 units per US$1. We strongly believe that the amount of money at stake was non-negligible both for the managers and the students, such that dominance[10] was achieved.

We also described the dynamics of the games. Each of the two games would be played for 5 initial rounds. At the end of the fifth round a flip of a coin would decide if another round would take place.[11] After each round the assistants would collect the decision sheets and the pay-offs for that round would be calculated.[12] The sheets would

[9] Starmer and Sugden (1991) call this practice *random lottery incentive system*. They find evidence that subject responses are not significantly different between designs based on this system and single choice designs.
[10] Smith (1982) states that dominance is achieved when: "The reward structure dominates any subjective costs (or values) associated with the participation in the activities of the experiment" (p.934).
[11] This strategy was used to reduce the number of deviations from the Nash strategy in the last round of each game. Since the two games described in this paper were the last to be played during the session, we took the liberty of stopping the second of our games after the 5th round, without flipping a coin. The rest of the games ended according to the flip of the coin. Merely by chance, all the games played first also ended after round 5.
[12] The use of computers would certainly simplify this. Computers were not available.

then be returned to the players for the next round. The players would keep their unidentified randomly selected partner through all rounds of a game.

Finally, for the sample of managers and the sample of 21 pairs of students (henceforth students-1), the games were played in the following order: first, the game with collective penalties and second, the game with random fines. This order was initially selected because we believed this was the more natural way for policy to evolve over time. Still, we were interested in testing the impact of playing the games in a different order. Therefore, in the sample of 16 pairs of students (henceforth students-2), we played the games in the inverse order, i.e., first, the game with random fines and second, the game with collective fines.

4. Analysis of the results

4.1 "Within sample-within game" analysis

We start our analysis by exploring the different ways in which both the managers and the students played both games. The primary objective is to determine whether the games achieve the optimal Nash equilibrium. We should recall from the previous section that optimal abatement occurs when both firms reduce pollution by two units. Furthermore, the (2,2) outcome is the optimal Nash equilibrium in both games. As mentioned in Section 3.2, our conjecture is that both games are efficient in achieving optimal abatement. The categorical nature of the data and the small sample size preclude a formal statistical analysis of this conjecture.[13] Hence, our analysis will be based on the following criteria: (i) simple majority, i.e. if the Nash equilibrium is the most frequent outcome of all possible outcomes, then we do not reject the conjecture; (ii) absolute majority, i.e. if the frequency of pairs playing Nash is higher than the sum of the frequencies of all other possible outcomes, then we do not reject the conjecture. Tables 4, 5 and 6 contain the percentage of pairs that played each possible outcome in each round (1,2,3,4,5) of both games (c-collective and r-random), for the sample of managers, students-1 and students-2, respectively.[14]

[13] Table A1 in Annex 1 contains the net social benefits associated with all possible outcomes and categories of the games.

[14] Empty cells in the following tables stand for 0%.

Table 4: Structure of responses in games with collective penalties (C1-C5) and random fines (R1-R5): Sample of Managers.

Strategy	C1	C2	C3	C4	C5	R1	R2	R3	R4	R5
% in 3-3	12.5	6.25	18.75	12.5						
% in 3-2	25	31.25	31.25	12.5	25	25	25	6.25	12.5	12.5
% in 3-1	6.25		12.5	12.5						6.25
% in 2-2	**31.25**	**18.75**	**12.5**	**25**	**37.5**	**50**	**56.25**	**68.75**	**68.75**	**56.25**
% in 3-0		6.25			6.25				6.25	
% in 2-1	25	31.25	18.75	18.75	12.5	18.75	6.25	18.75	6.25	12.5
% in 2-0		6.25			6.25	6.25	12.5	6.25	6.25	6.25
% in 1-1				12.5	6.25					6.25
% in 1-0			6.25	6.25	6.25					
% in 0-0										

We observe a clear pattern in the responses of the managers. The initial rounds of the games are characterized by many pairs over-abating (i.e. playing 3,3 or 3,2). In round C3, the percentage of pairs over-abating even reaches 50%. The share of pairs that over-abate remains stable at 25% after round C4 and falls toward the end of the game with random penalties. Also, with the exception of rounds C2, C4 and C5, the share of pairs in non-compliance ranges from 19 to 25% throughout both games. Finally, the share of pairs that played the Nash outcome steadily increases after round C2 going from 12.5% in C3 to 68.75% in R4. We observe a slight drop in the end round. Finally, apart from round C2 and C3, based on the simple majority rule we do not reject the conjecture that both games are efficient in reducing pollution optimally in all rounds of the games. Furthermore, for the random game, which in this case was played second, we do not reject the conjecture even under the absolute majority rule. Round 2 of the collective penalties game has similarly high shares of individuals under and over-abating; whereas round C3, as we mentioned earlier, has a very high (50%) share of pairs that over-abate, possibly reflecting a cautious behavior by the managers.

Table 5: Structure of responses in games with collective penalties (C1-C5) and random fines (R1-R5): Sample of Students-1

Strategy	C1	C2	C3	C4	C5	R1	R2	R3	R4	R5
% in 3-3	4.76				4.76					
% in 3-2	4.76	9.52	14.29	4.76	4.76	9.52	19.05	9.52	4.76	9.52
% in 3-1	4.76	9.52			9.52			4.76		
% in 2-2	**38.10**	**33.33**	**33.33**	**33.33**	**38.10**	**57.14**	**42.86**	**47.62**	**52.38**	**33.33**
% in 3-0		4.76		9.52				4.76	4.76	4.76
% in 2-1	28.57	28.57	23.81	19.05	14.29	4.76	19.05	4.76	4.76	4.76
% in 2-0			4.76	4.76		19.05	4.76	14.29	14.29	19.05
% in 1-1	14.29	9.52	14.29	9.52	23.81	4.76			9.52	4.76
% in 1-0	4.76	4.76	4.76	19.05	4.76	4.76	4.76	14.29		14.29
% in 0-0			4.76				9.52		9.52	9.52

A different pattern of responses can be established for the student sample. Most striking is the observation that the share of pairs over-abating is very low throughout all rounds of both games. On the other hand, the share of pairs in non-compliance is very high, reaching even 52% and 62% in rounds C3 and C4, respectively. Moreover, with the exception of the end round, R5, the game with random penalties exhibits smaller shares of non-compliance. Finally, based on the simple majority rule we do not reject the conjecture that both games achieve optimal pollution reduction. Once again, in the case that the game was played second, we do not reject this conjecture in 3 out of 5 rounds, even under the absolute majority rule.

Table 6: Structure of responses in games with collective penalties (C1-C5) and random fines (R1-R5): Sample of Students-2

Strategy	R1	R2	R3	R4	R5	C1	C2	C3	C4	C5
% in 3-3				6.25						
% in 3-2	18.75	18.75	6.25		6.25	6.25	12.50			
% in 3-1	6.25	6.25			6.25					
% in 2-2	**37.50**	**37.50**	**37.50**	**31.25**	**37.50**	**62.50**	**56.25**	**50.00**	**62.50**	**62.50**
% in 3-0		6.25	6.25	6.25						
% in 2-1	12.50	18.75	37.50	31.25		25.00	18.75	31.25	12.50	12.50
% in 2-0	25.00	6.25	6.25	12.50	31.25			6.25	12.50	12.50
% in 1-1			6.25		6.25	6.25	12.50	6.25	12.50	6.25
% in 1-0		6.25		6.25				6.25		6.25
% in 0-0				6.25	12.50					

Finally, we turn our attention to the second sample of students, who received the games in switched order. These results seem to indicate that order effects, rather than game effects, are responsible for the changes from one game to the other that we observe in Tables 4 and 5. This result is consistent with our hypothesis, since both games provide the same incentives if players are risk neutral, and theoretically their expected results are equivalent. Disregarding the type of game we are discussing, we observe the same structure of strategies in Tables 5 and 6 for both student samples. Again, the share of pairs over-abating is very low, compared to the share of pairs that under-abate. This consistent pattern seems to be in contrast to the behavior of the managers, who tend to over and under-abate in approximately equal proportions. Once again, under the simple majority rule for the game that was played first, and even under the absolute majority rule for the game played second, we do not reject the conjecture that the games achieve optimal abatement in all rounds.

In summary, based on the simple majority rule outlined at the beginning of this section, we find that both games in all three samples achieve the optimal outcome. Furthermore, if the game was played second, i.e. after possible learning effects, the game even passed the more stringent condition of absolute majority. This holds irrespective of the type of game played second. We also obtain some indication of possible differences between students and managers, and possible order/learning effects rather than game effects.

4.2 "Between samples-between games" analysis

In order to further explore the possible differences between students and managers, as well as the presence of order versus game effects, we would need to do between sample and between game comparisons. Since our data consists of frequencies in discrete categories, we can use a Chi-square test to statistically determine any presence of differences between the two independent groups from which those frequencies originated. If the groups are different, then we would expect that the frequency with which the subjects in each group enter each category would also differ more than we would expect from usual random deviations (Siegel and Castellan, 1988; Gibbons and Chakraborti, 1992). The Chi-square test is valid for independent samples and is likely to fail if the number of occurrences in each category is too low. Therefore, we cannot use the categories presented in Tables 4 to 6, and we need to pool similar categories until we meet the requirement. We decided to have three categories, namely: (i) suboptimal compliers (3,1 and 1,3) and over-abaters (3,3; 3,2; and 2,3), (ii) optimal abaters (2,2) and (iii) under-abaters (3,0; 2,1; ...;0,0) (see Annex 1). Additionally, the observations in each category are pooled across the rounds of each game. This effectively assumes that the rounds of a game are independent, which is granted by the use of the *random problem selection procedure* described in Section 3.3.[15] The test will then be based on the following Table 7, where each column can be intuitively perceived as a summary of how the game was played.

<u>Table 7: Pooled frequencies in all three categories</u>

	Managers		**Students-1**		**Students-2**	
	Collective game	*Random Game*	*Collective Game*	*Random Game*	*Random Game*	*Collective Game*
Subopt. Compl. and over-abating	33	14	15	14	15	3
Optimal Abating	20	48	37	49	37	47
Under-abating	28	18	53	44	53	30

We start our analysis based on the Chi-square test by continuing to explore the hypothesis that both games, i.e. the game with collective penalties and the one with random fines, are efficient in the sense that they achieve optimal abatement as the Nash equilibrium. In the previous analysis we do not reject the conjecture that both games achieve the Nash outcome at least in a simple majority of cases in each round. Still, the original hypothesis was stricter since it originated from the fact that both games are theoretically equivalent. However, Tables 4 and 5 for the manager and student-1 sample seem to indicate that the random fines games, which was played second, is more efficient in achieving the optimal Nash outcome, since this outcome was chosen even in the absolute majority of cases. Nevertheless, Table 6 shows exactly the opposite result, i.e. the game with collective penalties, which was played second this time, seems to be more efficient. Apparently, the observed differences between games are due to the sequence of playing rather than due to actual differences. Table 8 contains the formal statistical analysis based on the Chi-square. The critical value in all cases (2 degrees of

[15] The alternative is to do tests for each round and aggregate the results into one overall conclusion, possibly using the Bonferroni adjustment of the significance. Still we believe that not much can be gained by opting for this alternative, which in turn precludes the use of the Chi-square test due to small frequencies.

freedom[16]) is 5.99 (9.21) at a 5% (1%) confidence level. The p-values are provided in parentheses in all tables.

Table 8: Chi-square test for testing order versus game effects

Hypothesis	Statistic
Managers: Collective fines game is equivalent to random fines game	21.37 (0.0002)
Stud-1: Collective fines game is equivalent to random fines game	2.84 (0.417)
Stud-2: Collective fines game is equivalent to random fines game	10.84 (0.0042)
Collective fines game (first, as in Stud-1) is equivalent to random fine game (first, as in Stud-2)	0 (1.000)
Random fines game (second, as in Stud-1) is equivalent to collective fines game (second, as in Stud-2)	4.79 (0.0911)

The first three rows of the table contain a test that compares both games within a given sample. This is granted by the fact that the subjects were told that they would have a new partner in each of the two games and, once again, by the *random problem selection procedure*. Based on this test, we conclude that for the sample of managers and students-2, the pattern of responses in one game is significantly different than the pattern in the other game. However, the last two rows of Table 8 contain a comparison based on the order in which the games where played. Hence, we compare the games that were played first and detect no significant difference in the pattern of responses. The same is done for the games that were played second, and we obtain a similar result. In summary, the set of tests presented in Table 8 contribute to not rejecting the hypothesis that both types of games are equivalent, and that all perceived differences are due to ordering effects. We could also test for order effects within games by comparing both samples of students. This is done in Table 9.

Table 9: Chi-square test for testing within game order/learning effects

Hypothesis: No order/learning effects	Statistic
Game with collective fines when played first (Stud-1) is equivalent to the collective fines game when played second (Stud.2)	12.41 (0.0020)
Game with random fines when played first (Stud-2) is equivalent to the random fines game when played second (Stud-1)	2.08 (0.3534)

[16] The degrees of freedom is calculated as $df = (r-1)*(c-1)$, where r is the number of categories (3 in our case) and c is the number of columns to be compared (2 in our case).

We detect a significant difference in the pattern of responses only for the game with collective penalties. Certainly, the game with random fines is more complex than the one with collective penalties, which might account for the weak learning effect in the first compared to the second.

Next we turn our attention to subject pool effects. As we recall from Section 3.2, the hypothesis now is that there are no observable differences between the subject pool of managers and that of the more usual, convenience sample of students. If this is the case, then the Chi-square test should not detect any significant difference in the pattern of responses between the two groups. Table 10 contains the test statistics for the hypothesis of no subject pool effects.

Table 10: Chi-square test for subject pool effects

Hypothesis: No subject pool effects	Statistic
Collective fines game: The behavior of managers does not differ from that of students-1	16.72 (0.0002)
Random fines game: The behavior of managers does not differ from that of students-1	7.83 (0.0199)

In both cases, we reject the hypothesis that there are no differences in the pattern of responses arising from subject pool differences. In previous paragraphs, we already hinted at the source of this difference; in Table 7 we observe that in the game with collective penalties the managers have similar shares in all three categories. This is particularly so for the suboptimal categories of under and over-abatement. In the sample of students, we definitely observe a different pattern of responses characterized by a very high share of under-abatement. A similar structure holds for the game with random fines. Again the managers have similarly low frequencies in the sub-optimal categories, which is in clear contrast to the sample of students who have a large frequency of pairs that under-abate.

Another interesting difference between managers and students, related to the previous discussion on differences in their patterns of responses, lies in their reaction to the fine in the game with random fines. In the sample of managers, 100% of those who received the fine in any given round subsequently increased their stated abatement in the rest of rounds. Even more, 67% of them reduced pollution optimally in all of the rounds left of the game. The picture is very different for both student samples. More or less 50% of the fined players increased abatement in the next period, but for approximately half of them, this increase was still sub-optimal and, most interestingly, only temporary (for one round). This points out two potential differences between the two subject pools. The first one is a possibly higher degree of risk aversion in the sample of managers. The second relates to the strong learning effects observed in the experiments. We believe that receiving the fine in the game with random fines could be seen as a revelation of information regarding the true consequences of cheating in that game. From this perspective, it seems that 100% of the managers successfully incorporated that information into their decision-making process, whereas a much smaller share of students did so, and most of them only temporarily. In this sense, one could hypothesize about differences in the ability of both subject pools to incorporate new information into their decision-making process.

5. Conclusions

We presented an experimental study to test different mechanisms to regulate non-point pollution. In particular, we tested two efficient mechanisms proposed by Xepapadeas (1991) to deal with this type of pollution. One of them involves the combination of collective fines and subsidies for pollution abatement, whereas the other combines the same subsidies with random fining.

Regarding our first hypothesis, we do not find significant differences in the games. Most of the observed differences in the strategies employed by the players can be linked to order/learning effects. Further, based on the simple and absolute majority decision rules described in Section 4.1, we conclude that both methods are efficient in achieving the optimal pollution reduction. The decision about which to use can then, in principle, be guided by which method is more politically feasible, or in accordance with the previously existing regulatory framework. For example, it might be perceived as unfair that a firm receives the full fine given that its individual pollution was in line with the desired target. Some participants in the experiments privately expressed this opinion about the game with random fines. In such a case, a system of collective fines would possibly be perceived as ethically preferred.

Another interesting conclusion is the importance of learning effects. In those games that were played second, compliance with the desired pollution reduction was significantly higher, and this outcome dominated all others by an absolute majority in each round. This result indicates the need for providing a suitable learning or maturing period for a newly implemented policy, before evaluating its performance. The importance of learning behavior is discussed by Starmer (1999), who argues that, if the proper incentives are provided to encourage learning, then "…actual behavior might gravitate towards optimal solutions over time" (p.F12).

In addition, our results indicate that firms can understand and adapt their behavior to elaborate regulatory contracts after a suitable learning period, leading to efficient outcomes. This result is particularly relevant for the developing country context from which our sample was drawn, where skepticism to any, not to say an elaborate, regulatory contract is widespread.

A second issue in this paper was to test for behavioral differences among "real" decision makers and convenience samples such as students. In all cases, our analysis allows us to reject the hypothesis that managers and students performed similarly in this experiment. The main difference lies in the observation that the managers tend to over and under-abate in similarly low proportions, whereas both samples of students show a clear pattern of high under-abatement in all rounds of the games. The observed differences between subject pools are in line with similar exercises in the experimental economics literature. This evidence seems to suggest that professional experience might create a behavioral gap between the managers and the students (Ball and Cech, 1996; Binmore, 1999). Additionally, a related argument is that experienced managers have different attitudes towards risk than students. Although this study was not designed to test for risk aversion, the fact that the frequency of students "cheating" is high compared to the managers could be an indication of different degrees of risk aversion.

Ball and Cech (1996) argue, based on an extensive review of the literature, that the validity of using a convenience sample of students as surrogates for a more realistic sample of individuals, will depend on the nature of the study. Our experiment studies environmental regulation under non-point pollution, and is hence a case in which "information subtleties or behavioral nuances" are relevant for the results. Very few students, if any, have any experience with environmental regulators and regulation, i.e. taxes, command and control measures, etc. Therefore, the evidence presented in this paper is in line with Ball and Cech, and confirms that students cannot immediately jump into the shoes of firm managers.

Appendix 1

Table A1: Calculation of net social benefits

Possible strategies by each pair		Social Benefits	Total costs	Net Social Benefits	Categories
3,3		300	240	60	Overcompliers and suboptimal compliers
3,2	2,3	250	180	70	
3,1	1,3	200	140	60	
2,2		**200**	**120**	**80**	**Optimal compliers**
3,0	0,3	150	120	30	undercompliers
2,1	1,2	150	80	70	
2,0	0,2	100	60	40	
1,1		100	40	60	
1,0	0,1	50	20	30	
0,0		0	0	0	

Acknowledgements

We are greatly indebted to John List, Albert Schram and Sharon Kühlman for their valuable contributions to this paper. In addition, the advice from Peter Martinsson, Fredrik Carlsson, Olof Johansson, Christoph Schmidt and Thomas Sterner significantly improved our work

References

Ball, S.B and P-A. Cech (1996), "Subject Pool Choice and Treatment Effects in Economic Laboratory Research", *Research in Experimental Economics*, **6,** pp. 239-292.

Beattie, J. and G. Loomes (1997), "The Impact of Incentives Upon Risky Choice Experiments", *Journal Of Risk and Uncertainty*, **14,** pp.155-168.

Binmore, K. (1999), "Why Experiment in Economics", *The Economic Journal*, **109**, pp. F16-F24.

Coller, M. and G. Harrison (1996), "Methods in Experimental Economics: A Review", *Journal of International and Comparative Economics*,

Cunningham, W., W. T. Anderson and J.H. Murphy (1974), "Are Students Real People?", *Journal of Business*, **47:**3, pp. 399-409.

Friedman, D and S. Sunder (1994), *Experimental Methods, A Primer for Economists*. Cambridge University Press.

Fudenberg, D. and J. Tirole (1991), *Game Theory*, The MIT Press, Cambridge.

Gibbons J. D. and S. Chakraborti (1992), *Nonparametric Statistical Inference*, 3rd edition, Marcel Dekker, New York

Herriges, J., R. Govindasamy, and J.F. Shogren (1994), “Budget-Balancing Incentive Mechanisms”, *Journal of Environmental Economics and Management*, 27:3, pp. 275-285.

Holmström, Bengt (1982), “Moral Hazard in Teams”, *Bell Journal of Economics*, 13, pp. 324-340.

Kritikos, A. (1993), “Environmental Policy under Imperfect Information: Comment”, *Journal of Environmental Economics and Management*, 25:1, pp. 89-92.

Meran G. and U. Schwalbe (1987), “Pollution Control and Collective Penalties”, *Journal of Institutional and Theoretical Economics*, 143, pp.616-629.

Nalbantian, H.R. and A. Schotter (1997), “Productivity under Group Incentives: An Experimental Study”, *American Economic Review*, 87:3, pp.314-341.

Plot, C. (1982), “Industrial Organization Theory and Experimental Economics”, *Journal of Economic Literature*, **20:4,** pp.1485.1527.

Segerson, K. (1988), “Uncertainty and the Incentives for Nonpoint Pollution Control”, *Journal of Environmental Economics and Management*, 15:1, pp. 87-98.

Siegel, S. and N.J. Castellan (1988), *Nonparametric Statistics for the Behavioral Sciences.* McGraw-Hill.

Smith, V. (1982), “Microeconomic Systems as an Experimental Science”, *American Economic Review*, 72:5, pp. 923-955.

Spraggon, J. (2002), “Exogenous Targeting Instruments as a Solution to Group Moral Hazards”, *Journal of Public Economics,* forthcoming.

Starmer, C. (1999), “Experimental Economics: Hard Science or Wasteful Tinkering”, *The Economic Journal*, **109**, pp. F5-F15

Starmer, C. and R. Sugden (1991), “Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation”, *American Economic Review*, **81**, pp. 971-978.

Xepapadeas, A. (1991), “Environmental Policy under Imperfect Information: Incentives and Moral Hazard”, *Journal of Environmental Economics and Management*, **20**, pp.113-126.

The Pricing of Protected Areas in Nature-Based Tourism: A Local Perspective

Francisco Alpízar

Abstract

This paper extends the literature on optimal pricing of recreation in protected areas by introducing price discrimination between groups of visitors and, given that the agency charges different prices to subsets of visitors, by including a distributional dimension that is particularly relevant for a national park agency receiving visitors from different origins. Other issues related to optimal entrance fees, including congestion costs and positive spillover effects on local communities resulting from changes in visitation, are also discussed. Based on the theoretical model, the paper provides an estimation of optimal entrance fees and revenues for the Costa Rican system of protected areas.

Keywords: Optimal prices, price discrimination, recreation, protected areas, Costa Rica

JEL-classification: D62, H23, Q26

1. Introduction

Most countries have reserved a portion of their territories to the protection and preservation of natural environments. There are different categories and levels of conservation, but in general all protected areas are created in order to "…promote the persistence of species, communities or ecosystems that would otherwise decline or become extinct in the wild" (Lunney et al. 1997, p.138). A second main purpose in the creation of protected natural environments is the preservation of landscapes of outstanding beauty that provide opportunities for recreation or scientific study (USNPS, 1996).

In this paper, I seek to provide the theoretical underpinnings for the optimal pricing of protected areas used in recreational activities, from the perspective of a local park agency interested in maximizing welfare. The approach in this paper extends the existing literature on pricing of protected areas by introducing price discrimination between different groups of visitors. This allows for a more efficient adjustment to the distortions imposed by, for example, a cost recovery restriction. Another innovative element in the theoretical model is that, given that the agency charges different prices to different subsets of visitors, the model also includes a distributional dimension that is particularly relevant for a local park agency receiving visitors from different origins or of different nationalities. Finally, I discuss several other issues related to optimal entrance fees, including congestion costs on fellow visitors, the spillover effects of changes in entrance fees on local communities and the treatment of fixed investment costs.

Based on this theoretical foundation, the paper provides an estimation of the optimal entrance fees and associated revenues for the Costa Rican National Park System. Costa Rica is one of the few countries in the world where entrance fees to protected areas have changed several times, allowing for the estimation of the demand for recreational-day-visits using actual visitation data. This input is then used to calculate optimal prices. Although other studies have explored the willingness to pay to enter specific protected areas in Costa Rica and elsewhere, to my knowledge none have attempted to estimate optimal prices based on welfare maximization, and only one (Chase et al., 1997) has calculated entrance fees that at least maximize revenues from foreign visits.

Given the reduced availability of public funds, user fees for recreation in protected areas are an increasingly relevant source of funds to the park agency. A well-designed system of fees can make these areas more financially self-sufficient, thereby sending a positive signal to the government authorities regarding the private value of lands dedicated to conservation. If the required information is available, user fees can also be used to manage visitation in order to avoid imposing an excessive burden on the natural environment, reduce congestion in some parks and smooth the seasonal pattern of visitation. A related issue is that if the protected area is located in a relatively poor host country and entrance fees are small and below the amounts that foreign visitors are willing to pay for enjoying the resource, then the perverse consequence is that the host country will subsidize recreation for visitors from richer countries (Laarman and Gregersen, 1996).

As in any other typical developing country, the Costa Rican system of national parks is under constant threat from the lack of funds to manage the resources properly, and the pressure to exploit them in alternative commercial activities. Still, the case of Costa Rica is widely cited as an ecotourism success story. In the last 15 years, Costa Rica has experienced a large increase in international tourism. Most of this increase is due to Costa Rica's system of national parks, which encompasses approximately 24% of the national territory and a vast array of different types of marine and land ecosystems. According to a survey of the local tourism board (ICT, 2000), 71% of all holiday visitors to Costa Rica visit at least one national park, and most of the activities of foreign visitors are highly related to natural attractions in the country.

Despite the importance of nationally protected areas for tourism in Costa Rica, the increase in tourism activities has not caused an increase in the government budget dedicated to managing and protecting this resource. On the contrary, this budget, together with international support, has been steadily decreasing. In 1994, the government decided to recover the costs of managing the parks for recreation, and started to price discriminate based on nationality. Since then, there have been several changes in price to foreign visitors, but little or no formal criteria have been used in the determination of price changes (Bermúdez, 1996). An improvement in the criteria for setting the entrance fee is important not only for Costa Rica, but for the rest of the countries in the region, and potentially for other regions of the world attempting to develop an ecotourism strategy for their protected areas.

2. The economic model

Assume there exists a local park agency in charge of managing the system of protected areas. In particular, it sets the price for recreational visits to the park system,[1] pursuing the objective of maximizing welfare from the consumption of recreational services. For simplicity, assume that the park agency treats all parks as a single composite good, i.e. the price is the same for all parks. This is currently the situation in Costa Rica and is therefore consistent with the available data.[2]

---

[1] An obvious assumption underling the analysis of this paper is that it is possible to charge an entrance fee to enter the parks and, furthermore, that the costs of managing this fee are not prohibitive.

[2] A drawback of this assumption is that no differential pricing of parks is possible. Entrance fees could be used to drive visitors away from highly visited parks, if information about cross-price elasticities and carrying capacities were available. A further extension of this model to include cross-price effects among several parks in a system is an obvious next step. Wilman (1988) develops a model where the park agency is treated as a multiproduct monopoly, and obtains the optimal Ramsey prices.

In the absence of the cost recovery restriction, the traditional first best optimal pricing strategy requires the price to be set equal to marginal cost plus a correction for external effects. If the park agency faces a binding budget balancing requirement, the price should be equal to the average cost.

If the government is concerned with distributional issues, particularly regarding national and foreign visitors, and can exercise price discrimination between the two groups, then the maximization of weighted social welfare from visitation will result in pricing rules that take into account the differences between the two groups of visitors, and will therefore depart from the first best rule mentioned above. Accordingly, the park agency sells the same good to two types of consumers: national (*n*) and foreign (*f*) visitors. Hence, it can identify the separate demands (x_n and x_f) and charge different prices to each group. Also, the park agency follows a weighted utilitarian social welfare function, and assigns a weight of $\alpha \in [0,1]$ to the consumer surplus of foreign tourists. The optimal prices for national and foreign visitors to the protected areas, p_n and p_f, respectively, are obtained from the solution to the following welfare maximization problem:

$$Max..S = \alpha \int_{p_f}^{\infty} x_f(v)dv + p_f x_f(p_f) + \int_{p_n}^{\infty} x_n(v)dv + p_n x_n(p_n)$$
$$- C(x_f, x_n) - I - g(I, x_f, x_n) + T(x_f, x_n)$$

subject to: (1)

i. $p_f x_f + p_n x_n - C(x_f, x_n) - I - R \geq 0$

ii. x_f, $x_n > 0$.

The first four terms in S are the consumer and producer surplus of foreign and national visitors. The next four terms correspond to the park agency's total costs of recreation, $C(x_f, x_n) + I$, the external congestion costs, $g = g(I, x_f, x_n)$, and the positive spillover effect from visitation, $T = T(x_f, x_n)$.

The demand of group i ($i = f,n$) for recreation-visitor-days is given by $x_i = x_i(p_i)$, where p_i is the full or generalized price for a day-visit, i.e. it is the entrance fee including congestion costs. In addition, since income effects are expected to be small, then the Marshallian surplus is an exact measure of welfare change.

Throughout this paper I assume constant marginal costs of recreation, i.e. $C(x_f, x_n) = cx_f + cx_n$. This assumption seems quite realistic for nature-based tourism in the Costa Rican system of national parks (Bermúdez, 2001), and probably for ecotourism in protected areas in general. The term *I* captures all costs of recreation regarded by the park authority as not related to the number of visits (like trail clearing, signs, fire prevention, park rangers). This term also includes any expenditure aimed at increasing the quality of the recreational experience. In this model, investment in quality, *I,* is not treated as a decision variable for several reasons. First, the aim is to obtain optimal pricing rules for foreign and national visitors for any given level of *I*; in this sense, changes in *I* will affect the level of the optimal prices, but not the optimal pricing rule. Second, the park authorities can perceive investment in quality as a lumpy decision variable and therefore not freely adjustable, at least in the short run. Finally, there are natural (and legal) limitations to investment in a protected natural area. For example, although new trails can be opened to deal with increased visitation, these trails have an impact on the natural ecosystem that might in turn reduce the quality of the experience.

In most countries, the park agency faces a cost recovery condition like the one included in (1), which requires the agency's revenues from recreation to cover at least the fixed and variable costs of this activity. Furthermore, on many occasions the government imposes a minimal profit constraint, captured by *R,* which reflects the government political consensus regarding how much of the costs of managing the park for conservation purposes should be contributed by visitors. Obviously, recreational visitors not only enjoy the trails they walk on, but also some of the public services provided by the natural environment. Still, it could be regarded as unfair to charge visitors the full costs of creating and managing a protected area for conservation and recreational purposes (Beal, 1996; Dixon and Sherman, 1991; Anas, 1988).

In addition to the park agency's costs, visitors also face congestion costs, which are treated here as external costs; for a large number of visitors, each will disregard his/her own effect on congestion. Congestion plays an important role in nature-based tourism since solitude is often considered an important determinant of the quality of recreation

(Fisher and Krutilla, 1972). Still, the capacity of a natural area to sustain visitation without experiencing a decrease in the quality of the visit due to congestion is not exogenously determined, rather it is linked to investment in infrastructure. For example, the frequency of encounters might be reduced if the park agency increases the number of trails available to the visitors, thereby keeping up the quality of the visit. This relation is explicitly incorporated into the model, where congestion costs, $g = g(I, x_f, x_n)$, are a function of visitation and investment in quality. In addition, a crucial assumption underlies the model, namely that the protected area has a positive ecological carrying capacity of visitation, and that this capacity is not exceeded.

Finally, the term $T = T(x_f, x_n)$ captures the positive spillover effect of visitation on areas surrounding the protected areas, with $T'_{x_i} > 0$. In most cases, given that a protected area has been established, the protection of natural resources within it requires the creation of buffer zones that preclude the practice of intensive agriculture and livestock grazing. This creates an uneven distribution of income-generating activities in the region, commonly located in a rural area. Nature-based tourism is, in many cases, the only source of income for the population living in the areas surrounding the parks, and the decision regarding optimal management of recreational activities cannot be made in isolation from the impact of those management practices on rural development.[3]

For simplicity in the presentation of results, denote *A* the total external effects of visitation from any of the two groups, such that $A = -g(I,(x_f + x_n)) + T(x_f + x_n)$. In the absence of additional information and for the sake of simplicity, I have assumed that national and foreign visitors have the same impact on congestion and spillover effects. This assumption, although questionable regarding the spillover effects,[4] greatly simplifies the analysis of the optimal pricing rules, by concentrating the attention on the relative price sensitivity of both groups. Still, I will resort to a more general formulation whenever it provides interesting insights, and Appendix 1 provides the optimal pricing rules under this general formulation.

[3] Note that rural development might be crucial in the long-term success of a protected area, by reducing the pressure on the natural environment (see Leitman, 1998, for further discussion).
[4] Foreign visitors have generally a higher income than national ones, and spend more time close to the protected areas.

Accordingly, the net marginal external effect of a change in visitation from any of the two groups is given by:

$$\frac{\partial A}{\partial x_i} = A' = A'_f = -\frac{\partial g}{\partial x_f} + \frac{\partial T}{\partial x_f} = -\frac{\partial g}{\partial x_n} + \frac{\partial T}{\partial x_n} = A'_n . \tag{2}$$

Hence, *A'* will be larger than zero if the positive marginal spillover effect on the surrounding communities prevails over the congestion costs associated with a marginal increase in visitation.

The first order conditions are given by:[5]

$$(1-\alpha)x_f + (p_f - c + A')*x'_f + \lambda\left[x_f + (p_f - c)*x'_f\right] = 0 , \tag{3}$$

$$(p_n - c + A')*x'_n + \lambda\left[x_n + (p_n - c)*x'_n\right] = 0 , \tag{4}$$

$$p_f * x_f - c(x_f) + p_n * x_n - c(x_n) - I - R \geq 0 , \tag{5}$$

$$\lambda \geq 0 , \tag{6}$$

$$\lambda * Restriction = 0. \tag{7}$$

Beginning with the case of $\lambda = 0$, i.e. no binding cost recovery condition, it follows immediately from (4) that $p_n = c + A'$, and from (3) we have the following optimality condition for p_f:

$$\frac{p_f - c + A'}{p_f} = (1-\alpha)\frac{1}{\varepsilon_f} , \tag{8}$$

where $\varepsilon_i = -x_i'\frac{p_i}{x_i}; i = f, n$. Then, if the consumer surplus of national and foreign visitors has the same weight, i.e. $\alpha = 1$, then the usual marginal cost pricing holds for both groups, which is in line with the traditional results provided before. If $0 \leq \alpha < 1$,

[5] From now on, I will use $\frac{\partial x_i}{\partial p_i} = x'_i < 0$ and $\frac{\partial c}{\partial x_i} = c > 0$.

optimality requires a larger deviation from the marginal cost-pricing rule the closer α is to 0, ultimately resulting in a simple monopoly-pricing rule for foreign visitors. The park agency is able to recover more than the variable costs, and potentially make a profit.

Returning to the case of a binding cost recovery restriction,[6] in order to provide interesting, interpretable results, the rest of the solution will be divided into two cases. In Case 1, the consumer surplus of both groups is weighted equally, i.e. $\alpha = 1$, and in Case 2, a weight of zero is assigned to the consumer surplus of foreign visitors, i.e. $\alpha = 0$. If the park agency is located in a poor country and receives visitors from richer nations, this might be regarded by the agency as a fairer perspective. The results of the empirical section of this paper will be based on the optimality conditions obtained under this second case.

2.1 Case 1: $\alpha = 1$

The park agency exercises third degree price discrimination based on nationality, but still assigns equal weights to the consumer surplus of both groups.[7] A first intuitive result can be obtained by setting *A' = 0*, i.e., the case of no external effects. Solving (4) for λ and rearranging (3), we reach the following condition for the optimal relative price margin for both consumers groups:

$$\frac{\frac{p_f - c}{p_f}}{\frac{p_n - c}{p_n}} = \frac{\varepsilon_n}{\varepsilon_f}, \qquad (9)$$

The relative price margin is relatively higher in the market with relatively lower elasticity, which is a special case of the usual Ramsey rule. Given that nationals generally have a lower income and, most importantly, a higher availability of time and substitutes for recreation, I hypothesize that their price elasticity is higher, and hence the

---

[6] In what follows, I shall refer to $\hat{I} = I + R$ as a fixed cost, and refer to cost recovery restriction even though *R* includes costs not related to recreation.
[7] Wilman (1988) provides a numerical example of this case, given no external effects, (*A'=0*).

relative price margin will be lower for that consumer group, compared to the foreign visitors.

Going back to the general case of positive or negative external effects, we can use the first order conditions to solve for the optimal pricing rules for both foreign and national consumers, which are given by:[8]

$$(p_f - c) = \left[\frac{x_f x_n' - x_f' x_n' A'}{x_n^2 x_f' - x_f' x_n' A'(x_n + x_f) + x_f^2 x_n'} \right] \hat{I} + \\ + \left(\frac{x_n (x_f x_n' - x_f' x_n)}{x_n^2 x_f' - x_f' x_n' A'(x_n + x_f) + x_f^2 x_n'} \right) A' , \qquad (10)$$

$$(p_n - c) = \left[\frac{x_n x_f' - x_f' x_n' A'}{x_n^2 x_f' - x_f' x_n' A'(x_n + x_f) + x_f^2 x_n'} \right] \hat{I} + \\ + \left(\frac{x_f (x_n x_f' - x_n' x_f)}{x_n^2 x_f' - x_f' x_n' A'(x_n + x_f) + x_f^2 x_n'} \right) A' . \qquad (11)$$

Prices both to foreign and to national visitors have to deviate from the first best marginal cost pricing rule to account for the cost recovery requirement and the external effects generated by the visits. Price discrimination allows the park agency to minimize the welfare cost of this adjustment by taking into account the price sensitivity of both groups, relative to their total visitation. For example, in correcting for the distortion imposed by the cost recovery condition, the aim is to raise revenues while minimizing the impact on visitation. This is achieved by increasing the price charged to the group with the lowest price sensitivity, provided there are enough visits from this group to raise the required revenues. On the other hand, in correcting for negative marginal

---

[8] I assume that the following conditions holds: i. $A' > \frac{x_f}{x'_f}$; ii. $A' > \frac{x_n}{x'_n}$; or iii. $A' > \frac{x_n}{x'_n}(\frac{x_n}{x'_n + x_f}) + \frac{x_f}{x'_f}(\frac{x_f}{x'_n + x_f})$. Hence the denominators in all four terms in equations (10) and (11) are always negative, and the numerators in the square brackets are also always negative. All three conditions restrict the range of the net marginal external cost, *A'*, at a very large negative number, such that no practical implications are expected from this assumption.

external effects of visitation, the pricing rule should decrease visitation in such a manner that profits are reduced as little as possible. This can be achieved with an increase in the price charged to the group with the highest relative price sensitivity. As mentioned in the introduction, I expect national visitors to have a higher sensitivity to price than foreign visitors. Therefore, if condition 1 $\left(-x'_n / x_n\right) > \left(-x'_f / x_f\right)$ holds, the previous discussion indicates that foreign visitors should be targeted in order to raise most of the funds required by the cost recovery restriction, while visitation should be managed mostly via changes in the price to national visitors. Accordingly, if condition (1) holds, the terms in the square brackets on the RHS of equations (10) and (11), denoted γ_i, will be such that $\gamma_f > \gamma_n$.

The term in round brackets, Λ_i, in the RHS of equations (10) and (11) is a further correction for the external effects generated by visitors to the park. Again, if "condition 1" holds, then we would have $\Lambda_f > 0$ and $\Lambda_n < 0$. Assume that the negative marginal congestion costs prevail, such that the overall net marginal external effect is negative *(A' < 0)*, then equations (10) and (11) show that we have to increase the price to nationals to correct for the external effects, while even decreasing the price to foreigners in order to balance the budget. This result is driven by the first order condition that $\lambda > 0$, i.e. the cost recovery restriction must hold with equality. Note that a minimum profit requirement of *R* is still included in that condition. The optimal pricing rules provided here are the typical second best solutions, which are a balance of the two distortions. The Appendix provides the optimal pricing rules for the case of different marginal external costs for the two types of visitors, i.e. $A'_f \neq A'_n$.

2.2 Case 2: $\alpha = 0$

The park authority maximizes **only** national welfare, assigning a weight of zero to the consumer surplus of foreign visitors. By rearranging equation (3), one obtains

$$p_f = \frac{x_f - x'_f\, c}{-x'_f} - \frac{A'}{(1+\lambda)}, \tag{12}$$

and given $\lambda \geq 0$, we have that $\frac{1}{1+\lambda} \leq 1$. The second term in the RHS of equation (12) is a correction from the usual monopoly price, due to the externality. Hence we have that $p_f\big|_{A'<0} > p_f\big|_{A'=0} > p_f\big|_{A'>0}$, where $A'<0$ if the negative marginal congestion externality prevails over the positive spillover effect on surrounding areas.

I will first derive the case for $A' = 0$, i.e. the case with no external effects. Since the empirical section of this paper does not attempt to deal with the estimation of externalities due to the lack of information about these effects, this case will be the basis for estimation of optimal prices. The optimal pricing rule for foreign visitors is given by

$$p_f - c = \frac{x_f}{-x'_f}, \tag{13}$$

which is the usual inverse elasticity rule, namely $\frac{p_f - c}{p_f} = \frac{1}{\varepsilon_f}$. From now on, denote this pricing rule as $(p_f - c)\big|_{A'=0}$, i.e. this is the optimal deviation from marginal costs for foreign tourists under the cost recovery restriction and no external effects. Inserting (13) into the first order conditions and rearranging, one obtains:

$$(p_n - c)\big|_{A'=0} = \frac{-\left[\frac{x_f^{\,2}}{-x'_f} - \hat{I}\right]}{x_n}. \tag{14}$$

Note that the numerator in the RHS is negative when the revenues from the foreign market more than cover the variable costs in that market plus all fixed costs. In this case, these revenues will subsidize the price charged to national visitors. When revenues from foreign visitors do not fully cover the fixed costs, the price charged to nationals has to recover all remaining fixed costs as well as its own variable costs. This result arises from the fact that the price to foreigners is the profit maximizing one, and any deviation from it will result in less fixed costs covered. Note that if the government increases the revenue constraint, *R*, and given that prices were previously at their optimal levels, this will not affect the foreign visitors, but will rather result in a higher entrance fee for nationals. In summary, the profits obtained from charging a “monopoly” price to foreign visitors allow the public park agency to move closer to the first best marginal cost pricing of national visitors. Under $\alpha = 0$, a comparison to a uniform average cost price is straightforward.[9] By exercising price discrimination, the price to nationals will definitely be lower than the uniform average price, thereby increasing national visitation and welfare.

Using (12) and the first order conditions one can obtain the optimal pricing rules for national and foreign visitors for the case of non-zero external effects, which are given by

$$p_n - c = \left\{ \frac{(x_n)^2 - A'x'_n x_n}{(x_n)^2 - A'x'_n (x_n + x_f)} \right\} \left(\frac{-\left(\frac{x_f^2}{-x'_f} - \hat{I}\right)}{x_n} \right) + \left[\frac{x_n x_f}{(x_n)^2 - A'x'_n (x_n + x_f)} \right] A' \qquad (15)$$

and

$$p_f - c = \left\{ \frac{(x_n)^2 - A'x'_n x_n}{(x_n)^2 - A'x'_n (x_n + x_f)} \right\} \left(\frac{x_f}{-x'_f} \right) + \left[\frac{-(x_n^2 + x'_n \hat{I})}{(x_n)^2 - A'x'_n (x_n + x_f)} \right] A'. \qquad (16)$$

Note that the terms in round brackets in equations (15) and (16) correspond to the rules for optimal deviations from marginal cost, given no marginal external effects, namely $(p_f - c)\big|_{A'=0}$ and $(p_n - c)\big|_{A'=0}$. These basic rules are augmented to correct for the presence of external effects. The coefficients in front of $(p_f - c)\big|_{A'=0}$ and $(p_n - c)\big|_{A'=0}$ in equations (15) and (16) are the same, i.e. both baseline cases are adjusted multiplicatively by the same factor, henceforth called ϕ. If $A'>0$, i.e. the positive marginal externality prevails over the negative marginal congestion effect of visitation,

[9] Note also that, given the assumption of constant marginal costs, the average cost price is equivalent to a two-part tariff composed of a fixed “membership” payment to cover fixed costs and entrance fees equal to marginal costs.

then we have that $0 < \phi < 1$. Hence, both baseline-pricing rules should be scaled down. Alternatively, if *A'<0*, then $\phi > 1$,[10] and both baseline prices should be scaled up. The coefficient in square brackets in equation (15), henceforth called φ_{p_n}, is always positive.[11] To summarize the results for the optimal pricing rule for national visitors, if the negative marginal external effect is larger than the positive (*A'<0*), then the optimal pricing rule calls for a scaling up ($\phi > 1$) of the baseline pricing rule given by $(p_n - c)\big|_{A'=0}$. In addition, the pricing rule also calls for an additive correction, which slightly lowers the price to nationals ($\varphi_{p_n} > 0$). Again, both corrections arise due to the second best nature of the analysis, and are a balance between the distortion imposed by the cost recovery restriction and the presence of externalities. Intuitively, the multiplicative increase in the price to nationals and foreigners under *A'<0* raises further revenues for the park agency. In order to balance the budget, the park agency can slightly lower p_n.

Lastly, the coefficient in square brackets in equation (16), henceforth called φ_{p_f}, will be negative if $x_n^2 > -x_n' I$. If this is the case, then a negative net marginal effect will call for an increase in price, and vice versa. If the condition does not hold, for example if fixed costs are very large, then $\varphi_{p_f} > 0$, and a negative marginal effect will call for a decrease in price, and vice versa. This result appears intuitive given that $(p_f - c)\big|_{A'=0}$ is the pricing rule that achieves the highest revenues for the park agency and any deviation from that price will reduce funds to cover the fixed costs. Therefore, if fixed costs are large such that $\varphi_{p_f} > 0$ holds, then the scaling up of $(p_f - c)\big|_{A'=0}$ due to a negative net marginal externality must be counteracted by the additive reduction. The opposite holds for a positive net marginal effect; the scaling down of $(p_f - c)\big|_{A'=0}$ must be counteracted by an additive increase in the price. In summary, if fixed costs are large, the park agency will attempt to set the price as close to

---

[10] As *A'* tends to $\frac{x_n^2}{(x_n + x_f)x_n'}$ from the right, ϕ tends to infinity, i.e. both entrance fees become infinitely high. I therefore impose the restriction that *A'* is larger than that negative threshold level.

[11] The numerator is always positive and given the previous footnote, the denominator is also positive.

$(p_f - c)\big|_{A'=0}$ as possible. If fixed costs are not large, i.e. $\varphi_{p_f} < 0$, then the additive and multiplicative correction to the optimal pricing rule will reflect the sign of the net marginal external effect, such that a negative net marginal external effect will cause an increase in the baseline price.

Finally, an interesting result can be obtained by exploring a more general formulation of the marginal external effect than the one provided in (2). Although both types of visitors most likely generate similar marginal congestion costs, it is possible that they differ in terms of their impact on the areas surrounding the parks; in particular, the spending of foreign visitors in accommodation and lodging is likely to be larger than that of national visitors, even if they come from distant origins. If we allow for different marginal external effects, i.e. $A'_f \neq A'_n$, the optimal pricing rules for the case of $\alpha = 0$ would be given by:

$$p_n - c = \left\{ \frac{(x_n)^2 - A'_n x'_n x_n}{(x_n)^2 - x'_n (x_n A'_n + x_f A'_f)} \right\} \left(\frac{-\left(\frac{x_f^2}{-x'_f} - \hat{I}\right)}{x_n} \right) + \left[\frac{x_n x_f}{(x_n)^2 - x'_n (x_n A'_n + x_f A'_f)} \right] A'_f \qquad (15')$$

and

$$p_f - c = \left\{ \frac{(x_n)^2 - A'_n x'_n x_n}{(x_n)^2 - x'_n (x_n A'_n + x_f A'_f)} \right\} \left(\frac{x_f}{-x'_f} \right) + \left[\frac{-(x_n^2 + x'_n \hat{I}))}{(x_n)^2 - x'_n (x_n A'_n + x_f A'_f)} \right] A'_f . \qquad (16')$$

Equations (15') and (16'), although more complicated, have a similar interpretation to that of equations (15) and (16) above. Hence, I will only provide an example of the effect of relaxing the assumption of equal marginal external costs. A realistic scenario is $A'_f = 0$ and $A'_n < 0$, i.e. one where the marginal congestion and the spillover effect of foreign visitation cancel out, and national visitors only generate marginal congestion costs. In this case, the optimal pricing rules are given by $p_f - c = \frac{x_f}{-x'_f}$ and

$(p_n - c) = \frac{-1}{x_n}\left[\frac{x_f^{\ 2}}{-x'_f} - \hat{I}\right]$, which are equal to (13) and (14) for the case of no external effects. This result is a typical second best in the sense that, if foreign visitation generates no marginal external effects, then the park agency's best policy is to maximize revenues from foreign visitation by charging the usual monopoly price. Given the above, the price to nationals can only be used to balance the budget, disregarding the external effects generated by national visitors.

3. Application to international tourism in Costa Rica

This section contains an empirical application of the theoretical model presented above. The objective is to compute optimal prices for foreign and national visitors to the Costa Rican system of protected areas. Given the absence of information about external effects, this application will be based on the optimal pricing rules provided in equations (13) and (14). The main input into the computation is the estimation of the demand of foreign visitors for recreation-day-visits as a function of the entrance fee. In addition, an estimate of marginal (*c*) and fixed (*I*) costs is also required. Only if the profits from the foreign tourist group do not cover the fixed costs, will the price to nationals deviate from the marginal cost-pricing rule, in which case demand information will also be needed for that group.

The empirical part of this paper is an application using actual changes in the price to enter the Costa Rican system of national parks. As mentioned in the introduction, Costa Rica is one of the few countries were entrance fees have changed several times, which provides enough information to estimate the demand for protected areas. In those cases where direct demand information is not available as input for the theoretical model, the most common method of estimating the demand for a recreational site, as a function of visitation costs, is the Travel Cost Method. As in Riera (2000), I partition the vacation planning decision into two, and study the demand for day-visits given that the tourists have chosen Costa Rica as their vacation destination. Intrinsic in the empirical analysis is the assumption that changes in the entrance fee to national parks do not affect the decision to visit Costa Rica. This assumption seems reasonable given that entrance fees are expected to represent a relatively small share of the total vacation budget, and the fact that the country offers a myriad of other possibilities for doing nature-based tourism outside the protected areas.[12] Still, entrance fees are relevant in determining how many parks to visit and how many days to spend there, and as such might not be a negligible share of the on site vacation budget.

The crucial information about price changes and visitation was obtained from Bermúdez (1996, 2001). From 1991 to 1999 the Costa Rican system of national parks (SINAC) changed entrance fees to foreigners seven times by decree, setting it sometimes in USD and other times in colones. The price was always the same for all parks with the exception of a few months, and was changed at the same time; it ranged from USD 0.8 to USD 15. Visitors could pay in dollars or colones, and the average monthly exchange rate was used. All prices are in constant 1995 dollars.[13] Park visitation data contained monthly information on foreigner visits to the most important national parks, from the January 1991 to December 1999 period. Only 12 parks contained complete information for that period. Visitation for these parks was aggregated into a variable that reflected monthly foreign visits to the national parks system. This was required since there is no information about substitution effects in the data, given the governmental policy of a single price for the whole system. In addition, data on the monthly international flow of tourists was obtained from the Costa Rican Tourism Board (ICT, 2000).

After several specification tests, the following econometric model was selected:

$$\log(x_t) = b_1 P_t + b_2 \log(x_{t-1}) + b_3 \log(x_{t-12}) + b_4 \log(V_t) + d_1 SHI + d_2 S + d_3 Aug94\,, \quad (17)$$

which is also a reflection of the informational requirements of the theoretical model. The visits of foreigners to the park system, x_t, and to the country as a whole, V_t, are introduced as logarithms into the model, following the results of a Box-Cox specification. I expect that changes in international conditions will affect the demand for recreation in the protected areas through changes in the arrival of international tourists to Costa Rica, V_t. In addition, using a Box-Cox formulation to test for a suitable

[12] Alberto Salas, the chief of promotion at the Costa Rican Tourism Board, confirmed the validity of this assumption based on marketing studies by the board.
[13] I used information from the Costa Rican Central Bank and the IMF.

transformation of the price variable, P_t, the hypothesized logarithmic transformation can be rejected at a significance level of 6%. Since the objective is to obtain an input into the optimal pricing rules provided in the previous section, I keep the linear specification of price, in order to avoid cumbersome calculations. In addition, the demand equation includes three dummy variables, where SHI = 1 for the high season, which is between December and April; S = 1 for July and August, the medium-high season; and Aug94 = 1 for August 1994, i.e. the month immediately after the highest, unannounced increase in price. Lags are introduced to capture observed autocorrelation in the data, and possible stochastic seasonal effects. Notice that the demand equation does not include the price of substitute goods. As mentioned above, Costa Rica offers several other possibilities for doing nature-based tourism. Still, most of these options are openly available (beaches are a clear example) and there are very few privately owned natural reserves, such that no significant substitution effects are expected. [14]

Several other formulations of this demand equation were estimated and rejected in favor of the one presented here. More important is the fact that the estimated prices varied little from one econometric specification to another.

By rearranging equation (17) one obtains the long-run demand function, given by:

$$x_t = (e^{\Psi(d_1 SHI + d_2 S + d_3 Aug94)} V_t^{\Psi b_4}) e^{\Psi b_1 P} = W e^{\Psi b_1 P}, \tag{18}$$

where $\Psi = (\frac{1}{1 - b_2 - b_3})$. From equation (18) it is easy to derive the following expressions for the long run elasticity of visitation, $\varepsilon_{LR} = \Psi b_1 P$, as well as the optimal price rule, $p = \frac{-x}{x'} + c = \frac{-1}{\Psi b_1} + c$.

Table 1 presents the results of the estimation of equation (17), using the available information, and some fitness statistics.

[14] In any case, the Monteverde Cloud Forest is Costa Rica's most popular private reserve, and its entrance fee has systematically been above those of national protected areas.

Table 1: Results for equation (17)

Variable	Coefficients	P-values
Price	-0.0479	0.000
Lag-1 of log(X_t)	0.0696	0.050
Lag-12 of log(X_t)	0.5359	0.000
Log(visits to CR)	0.3745	0.000
Season Hi	0.1710	0.000
Season Med	0. 1068	0.016
Aug94	0.6907	0.000

Fitness statistics

Adjusted R-square	0.92
Breusch-Pagan (Ho:Homoskedasticity)	$\chi^2_{calc} = 4.363$, with 6 df
Test for Normality (Ho: Residuals are normal)	$\chi^2_{calc} = 1.74$, with 2 df
Godfrey's Test for group autocorrelation (Ho: no autocorr)	$\chi^2_{calc} = 21.09$ with 18 df

The econometric analysis contains the required information about the demand of foreigners for day-visits to the Costa Rican system of national parks. The average long-run price-elasticity of foreign visits is $\varepsilon_{LR} = 0.68$, which reflects the high substitutability between the protected areas and other natural beauties in the country.

I focus next on the other required information for the estimation of optimal entrance fees. Unfortunately, SINAC has no accurate information about the fixed and variable cost per visitor to the national parks, and there is no clear way to separate expenses related to recreation from those related to the provision of the park's public services. According to Bermúdez (2001), the constant recreational costs are between USD1 and USD3 per visitor. Therefore Table 2 provides different estimates of the optimal price charged to foreigners given different marginal costs, as well as their 95% confidence intervals estimated using the Delta method. Optimal prices range from USD8 per day-visit for the case of zero marginal costs, to USD12 for the case of c = 4. Since May 2002, the entrance fee to the national parks has been USD7 (5.80 in 1995 USD), which is below the 95% confidence interval for the optimal prices calculated using the model in this paper, even for the case of zero marginal costs. However, according to the model of Section 2, the optimal prices might be smaller than the ones presented in Table 2 if the positive marginal spillover effect dominates the congestion externality,[15] or if the park authority assigns a positive weight, $\alpha > 0$, to the consumer surplus of foreign visitors (see equation (8)). In addition past events indicate that the lobbying efforts of tour operators are successful in keeping the entrance fees low. The last row of Table 2 therefore provides the implied weights (α) calculated using equation (8), under the assumption of zero net externalities and for different marginal costs. For a marginal cost of two, the current price implies a value of $\alpha = 0.5$, i.e. the revenues relinquished due to the lower price imply the acceptance of a welfare transfer from Costa Rica to the international visitors at a rate of 1 to 2.

Table 2: Optimal prices and average monthly profit, in 1995 USD[16]

	c = 0	**c = 1**	**c = 2**	**c = 3**	**c = 4**
Optimal prices to foreign visitors	8.23 (6.6-9.9)	9.23 (7.6-10.9)	10.23 (8.6-11.9)	11.23 (9.6-12.9)	12.23 (10.6-13.9)
Profit	141,939	125,693	111,306	98,566	87,285
Profit under today's prices	134,403	111,230	88,057	64,884	41,711
Implied α - value under today's prices	0.29	0.41	0.53	0.65	0.78

Equation (14) shows that the price to nationals has to be set equal to marginal costs plus any fixed costs not recovered via the profits from foreign customers. If fixed costs are recovered (or in the absence of a budget restriction), then the price to nationals can be set equal to the marginal costs of recreation. Table 2 also contains the expected profit under optimal pricing to foreign customers after covering all variable costs. This can then be compared to budget information from SINAC in order to approximately determine whether the price to nationals should deviate from the marginal costs. Finally, Table 2 also contains the econometric model's predicted profit for the current entrance fee of USD7 (5.8 in 1995USD).

As mentioned before, unfortunately there is no information about the fixed costs of the recreational activities within the Costa Rican System of national parks. Still, there is

[15] There is some evidence that the government is particularly sensitive to the well-being of the people living immediately around the parks

[16] Multiply by 1.2 to obtain the price in 2002 USD.

information for some parks, particularly for the Irazú Volcano National Park, one of the three most visited and most developed parks in the system. The total budget in 1999 for this park was USD130,000, including all recreational and conservation costs during that year.[17] Hence, the budget balancing condition will be approximately satisfied with the profits of one and a half months using the optimal prices provided above for a marginal cost of two. This is an indication that the price to nationals could follow the first best optimal marginal cost-pricing rule. Currently, the price to national visitors is slightly lower than USD2 per visit; that is, very close to the marginal cost suggested by the experts of SINAC.

There are other previous studies that inquire into the willingness to pay of foreign visitors for entering a national park in Costa Rica. Using contingent behavior analysis, Chase et al. (1998) find that the WTP for the three most popular parks are between USD21 and USD25. In addition, they calculate revenue-maximizing fees that range from USD7 to USD13 for those three parks, all in 1995 dollars. Our estimates lie in that range. Finally, Schultz et al. (1998) conducted a CVM study aimed at establishing optimal differentiated entrance fees to Costa Rican national parks. They found that WTP for entering the most popular parks ranges from USD14 to USD23.

4. Conclusions

This paper suggests a theoretical model for the optimal pricing of a system of protected areas used for recreation, by stressing the possibility of third degree price discrimination based on the visitor's nationality, and by stressing the distributional fairness of assigning different welfare weights to the consumer surplus of different groups of visitors. Price discrimination allows for a more optimal adjustment to the distortions created by a cost recovery requirement, and to possible external effects from changes in visitation.

If the park authority assigns zero welfare weights to the consumer surplus of foreign visitors, the optimal price to foreign visitors follows a monopoly-pricing rule. When revenues from foreign visitors do not fully cover the fixed costs, the price charged to

[17] This budget excludes the initial costs of establishing the area for conservation. Most of these costs are borne by residents in terms of the social costs of tax revenues dedicated to expropriation and population resettlement, if necessary, but also in terms of the opportunity cost of dedicating the land to agriculture or livestock production when no expropriation is involved.

nationals has to recover the remaining fixed costs as well as its own variable costs. If externalities are introduced, then the pricing rules described before have to be adjusted. A negative marginal externality, for example, will require an upward multiplicative adjustment in the pricing rules provided before. In line with the welfare weights assumed, the additional revenues will allow a decrease in the price to nationals in order to balance the budget. Furthermore, in Section 2 I allow for different marginal external effects from the two types of visitors. In particular, I present what I believe is a realistic scenario, i.e. one in which the marginal external effects from foreign visitation cancel out ($A'_f = 0$) due to their positive marginal spillover effects, whereas local visitors generate marginal congestions costs and small spillover effects on the surrounding areas. In this case, the optimal pricing rules are reduced to the case of no external effects discussed above.

Since May 2002, the Costa Rican park agency has set the prices for entering a national park at USD7 per foreign visitor and approximately USD2 per national visitor. Our analysis indicates that the practice of price discrimination can successfully raise revenues and achieve a more optimal pricing policy. If the local authorities aimed at maximizing only local welfare from recreational activities, then the empirical application of our model indicates that prices to foreign visitors could be raised to USD8-12 depending on the approximate estimate of marginal costs. On the other hand, prices to nationals seem to follow the optimal pricing rule already, given that the increase in revenues from foreign visitors under the proposed optimal prices will most likely cover the fixed costs of recreation.

An additional issue is the estimation of the external costs of congestion and the positive spillover benefits on the surrounding areas. There is some evidence that the local communities surrounding the protected areas depend more and more on tourism, and one can expect that a strong increase in price might have a negative impact on the surrounding areas. Regarding congestion costs, there is no clear indication of the likely importance of it, although some parks have very high visitation rates during the high season. I acknowledge that the inclusion of external effects might have an impact on the estimated optimal prices.

Finally, the assumption that the park agency treats all parks as a single commodity is certainly a limitation of the theoretical analysis, although it closely reflects the likely price setting behavior of a park agency with no available information on cross price elasticities. Furthermore, this is currently the case in Costa Rica. This assumption might be fairly realistic for a country with several areas that are close substitutes in recreation, or a large country or area with only one park. A further extension of the model could treat the park agency as a multi-product firm.

Acknowledgments

I am thankful to Fernando Bermúdez, Olof Johansson-Stenman, Fredrik Carlsson and Jorge García for interesting discussions and comments.

Appendix 1: Optimal prices with differing marginal external effects

Assume the following general formulation for the external effects:

$$A = -g(I, \mathrm{x_f}, x_n) + T(\mathrm{x_f}, x_n),$$

with

$$A'_f = -\frac{\partial g}{\partial x_f} + \frac{\partial T}{\partial x_f} \neq -\frac{\partial g}{\partial x_n} + \frac{\partial T}{\partial x_n} = A'_n .$$

Then the optimal pricing rules provided in equations (10) and (11), for the case that the park authorities set equal weight to both consumer groups ($\alpha = 1$), have to be corrected and are provided below:

$$(p_f - c) = \left[\frac{x_f x_n' - x_f' x_n' A'_f}{x_n^2 x_f' - x_f' x_n' (x_n A'_n + x_f A'_f) + x_f^2 x_n'} \right] \hat{I} + \left(\frac{x_n (x_f x_n' A'_n - x_n x_f' A'_f)}{x_n^2 x_f' - x_f' x_n' (x_n A'_n + x_f A'_f) + x_f^2 x_n'} \right), \qquad (10')$$

$$(p_n - c) = \left[\frac{x_n x_f' - x_f' x_n' A'}{x_n^2 x_f' - x_f' x_n' (x_n A'_n + x_f A'_f) + x_f^2 x_n''} \right] \hat{I} + \left(\frac{x_f (x_n x_f' A'_f - x_f x_n' A'_n)}{x_n^2 x_f' - x_f' x_n' (x_n A'_n + x_f A'_f) + x_f^2 x_n'} \right). \qquad (11')$$

References

Anas, A. (1988). Optimal Preservation and Pricing of Natural Public Lands in General Equilibrium. *Journal of Environmental Economics and Management*, **15:** 158-172.

Baumol, W.J. and D.F Bradford (1970). Optimal Departures from Marginal Cost Pricing. *American Economic Review*, **60:** 3, pp.265-283.

Beal, D. (1996). A Theoretical Market Model for National Parks of Different Ecological Character Managed by One Agency. *Review of Marketing and Agricultural Studies*, **64:2**, pp.177-189.

Bermúdez, F. (1996). Propuesta de Variables para Definir Tarifas de Ingreso a las Áreas Silvestres Protegidas para Visitantes no Residentes. Universidad de las Americas.

Bermúdez, F (2001). Personal communication. Expert consultant in management of protected areas, SINAC.

Chase, L.C., D.R. Lee, W.D.Schulze and D.J. Anderson (1998). Ecotourism Demand and Differential Pricing of National Park Access in Costa Rica. *Land Economics,* **74(4):** 466-82.

Dixon, J.A. and P.B. Sherman (1991). Economics of Protected Areas. *AMBIO* **20:**68-74.

Feldstein, M.S. (1972). Distributional Equity and the Optimal Structure of Public Prices. *American Economic Review*, **62:** 1, pp. 32-36.

Fisher, A.C. and J.V Krutilla (1972). Determination of Optimal Capacity of Resource-Based Recreation Facilities. *Natural Resources Journal*, **12:** 417-444.

ICT(2000). Encuesta Aerea de Extranjeros. Instituto Costarricense de Turismo, Costa Rica

ICT(2000). Anuario Estadístico de Turismo, Costa Rica.

Laarman, J.G. and Gregersen, H.M. (1996). Pricing Policy in Nature-Based Tourism. *Tourism Management*. **17**: 4, pp. 247-254.

Leitman, J. (1998). Options for Managing Protected Areas: Lessons from International Experience. *Journal of Environmental Planning and Management*, **41:** 1, pp.129-144.

Lunney, D., B. Pressey, M. Archer, S. Hand, H.Godthelp and A. Curtin (1997). Integrating Ecology and Economics. *Ecological Economics*, **23:** 135-143.

Menkhaus, S. and D.J. Lober (1996). International Ecotourism and the Valuation of Tropical Rainforest in Costa Rica, *Journal of Environmental Management*, **47:**1-10

Riera, A. (2000). Mass Tourism and the Demand For Protected Areas: A Travel Cost Approach. *Journal of Environmental Economics and Management,* **39:** 97-116.

Ross, T.W. (1984). Uncovering Regulators' Social Welfare Weights. *The Rand Journal of Economics*, **15:** 1, pp.152-155.

Salas, A. (2002) Personal communication. Jefe de Promoción, Instituto Costarricense de Turismo

Schultz, S., J. Pinazzo and M. Cifuentes (1998). Opportunities and Limitations of Contingent Valuation Surveys to Determine National Park Entrance Fees: Evidence from Costa Rica. *Environment and Development Economics* **3:** 131-149.

USNPS (1996). "Criteria for Parklands" http://www.nps.gov/legacy/citeria.html.

Wilman, E.A. (1988). Pricing Policies for Outdoor Recreation. *Land Economics*, **64:3,** pp.234-241.