Valuation of a Moose-Vehicle Accident Mitigation Policy in Newfoundland[1]

Roberto Martínez-Espiñeira

Nikita Lyssenko

Department of Economics, Memorial University of Newfoundland[2]

July 15, 2016

Abstract

While moose are an iconic Canadian symbol, they are also considered by many in Newfoundland and Labrador as an environmental disamenity, due to the rising number of negative moose-people interactions. There are about 600-800 moose-vehicle collisions each year, which result on average in two fatalities per year. There is also an ongoing public debate regarding moose control options in the province. Economic theory suggests that only policies that balance benefits and costs result in a maximization of net social benefits. When it comes to the removal or abatement of an environmental bad (such as wildlife-vehicle accidents), it is more often than not the case that, while it is relatively easy to compute the cost of a policy, its benefits are difficult to monetize, because such policies often deal with goods and services without a market price. Economists have, however, developed non-market valuation techniques that aim at estimating the economic benefits of these types of policies. In this case, we apply the Contingent Valuation Method to estimate the benefits that Newfoundlanders would derive from a moose control program, including the estimation of a key ingredient of the policy decision in this instance, the value of a statistical life. The valuation exercise yields estimates of the benefits that respondents would derive from the mitigation of moose related road accidents in Newfoundland. These values could then be used to inform cost-benefit analyses of risk reduction policies related to collisions with wildlife in this province but also, with appropriate adjustments, of policies related to the sacrifice of resources in any other way in order to protect the public.

Contents

IV Methodology 79

List of Tables

List of Figures

Part I

Introduction

Moose is a species that brings many benefits to the province of Newfoundland and Labrador but it is also seen by many as an environmental disamenity, due to the rising number of negative moose-people interactions, particularly in the form of moose-vehicle collisions (MVCs). There is also an, often emotional and at times even bitter, ongoing public debate about how to address this issue. As it is the case with many other public projects aimed at improving road safety, the obvious solutions would impose costs on society in exchange for reducing the risk of death and injuries. Economic theory suggests that, in order to maximize net social benefits, any initiative aimed at reducing the risk of MVCs should balance the relevant benefits and costs. In this case, while it may be relatively easy to estimate the cost of a risk reduction policy, its benefits are difficult to monetize, because there is no easily observable market price for reductions in the risk of mortality or morbidity.

Similarly, in the case of many environmental and safety policies and regulations, some or all of the benefits are associated with the prevention of human fatalities and human injury. Implementing the policies or complying with the regulations typically imposes costs on taxpayers, producers and/or consumers. It is then necessary to decide whether the lives saved (or injuries avoided) are worth the costs when deciding whether a certain policy is desirable from a social point of view or to determine the scope of that policy. Economists have, therefore, developed non-market valuation techniques that aim at estimating the economic benefits of these types of risk reduction policies. The currently favoured approach that economists adopt to address this task, that is, the eco-

nomic approach to valuing life-saving was initially proposed by Schelling (1968) and it hinges on the notion that one cannot know whose life will be "saved" by a policy. Therefore, the question is not how to value the avoidance (or more rigorously, the delay) of a specific death but, instead, how to value small changes in mortality risk across a given population.

And, following the principle of "consumer sovereignty", which assumes that individuals are the best judges of their own best interest, economists have focused on estimating the rate at which individuals would trade their own money for small changes in their own mortality risk within a defined time period (Hammitt, 2000b). This trade-off rate helps derive a measure commonly used in the cost-benefit analysis (CBA) of safety policies, the *value of a statistical life* (VSL henceforth). In simple terms, which we will elaborate further in Section III, if an individual is willing to pay, say $3 for a reduction in the risk of his dying of one in a million, then his VSL is equal to $3 million. This neither implies, as unfortunately the concept is often misinterpreted, that the individual would himself be willing to pay $3 million to avoid certain death this year nor that he would accept certain death in exchange for $3 million. It simply means that 1,000,000 people similar to that individual would, together, be willing to pay $3 million to eliminate a risk that would be expected to randomly result in the death of one among them this year. Additionally, one should be aware that the VSL is not supposed to be a universal constant but will vary instead by individual and by circumstance (Hammitt, 2000b).

In this study, we apply the Contingent Valuation Method (CVM henceforth)

to estimate the benefits from a reduction in the risk of death (and injury) associated with moose vehicle collisions in the insular part of Newfoundland and Labrador. This valuation exercise yields an approximate measure of the benefits of potential policies aimed at the mitigation of moose related road accidents in this part of Canada, where moose densities are particularly high. This measure could then be transferred to any CBA that involves the sacrifice of resources in any other way in order to protect the public from the risk of collisions. It is, to our knowledge, the first type of study that values death risk reductions in this province.

In principle, one could follow a benefit-transfer approach and make use of VSLs obtained from earlier work based on road traffic risk reduction policies elsewhere. Indeed cost-benefit analysts have access to a wealth of studies[1] that provide a large range of estimates of the VSL. However, choosing the right estimate is, of course, a very difficult task (Dionne and Lanoie, 2004; Hauer, 2011), if only because there is an "embarrassment of riches" situation given by the plethora of different values that have been proposed.

A benefit-transfer approach would call for the adjustment of estimates previously obtained in other jurisdictions according to a series of observable characteristics of the target population. This is because, as we explain in our literature review in Part III, there is considerable evidence to support the notion that the the estimates of the VSL will depend on both individual characteristics and characteristics of the policy context, such as the income (or wealth), age, cul-

---

[1] Part III mentions several meta-analyses and surveys of studies that estimate VSL in road safety in different regions of the world.

ture, and health status of individuals or the baseline risk level and the individual degrees of control over the risk considered when estimating the VSL.

Also, as Dionne and Lanoie (2004) point out, even among societies sharing the same type of traffic-risk parameters and similar insurance plans, there may exist fundamental differences in terms of the personal preferences of individuals across jurisdictions. The individual's utility index would capture these differences, whether they are related to religious, cultural, or demographic factors and they would be the justification for conducting independent studies in individual countries. Furthermore, these determinants surely evolve throughout time even for a given individual, so it is advisable to update these estimates relatively frequently, even if an earlier estimate already exists for the target jurisdiction.[2]

We are not aware of any previous study that estimated the VSL for Newfoundland and Labrador, let alone in the specific context of moose-related accidents or even road safety in general. The empirical evidence in terms of VSL is also relatively scarce in the wider Canadian context. Therefore, we expect that our analysis will contribute to policy-making in this province by providing a point of reference for those cases in which the value of prolonging life is an ingredient in the CBA of a policy.

The report is organised as follows. Part II includes a description of the situation in Newfoundland regarding the abundance of moose and the human dimensions problems that it causes, together with the benefits it generates for this

---

[2] Unfortunately, as we comment on in Section 13.4.,often public agencies end up simply adopting as operating VSLs old estimates that are the product of "recycling" process and both perpetuating and legitimizing what can represent very outdated estimates (e.g. Hauer, 2011).

province, as well as a description of the measures hitherto considered to reduce the frequency of moose-vehicle accidents both in Newfoundland and elsewhere. Part III contains a summary of the theoretical approach we adopt to consider the economic benefits of reductions in the risk of suffering MVCs, based on the concept of the VSL, as well as a literature review of empirical studies aimed at providing estimates of the VSL, with particular attention to the context of road safety. Part IV contains a description of the main methodology we apply in this study to elicit the individual preferences about MVC risk reductions. We place the CVM within the context of stated preference methods and within the broader context of non-market valuation techniques designed to estimate individual WTP for small risk reductions, an approach that has all but replaced the traditional human capital approach to the valuation of risks to life. Additionally, we close this part of the report with a chapter on the issues involved in the analysis of the double-bounded dichotomous choice payment questions used in this study. In Part VI, we provide details of the data collection exercise. We describe the different versions of the questionnaire we used in our phone survey and we provide a brief descriptive analysis of the information we got on each resulting variable. We also explain in detail different manipulations applied to the raw data before they were used in the statistical analysis of WTP. The results of the econometric exercises aimed at the estimation of measures of WTP are presented in Part VII, first in terms of individual mean WTP for risk reductions that apply to our sample (Chapter 12) and then in terms of aggregate welfare measures for the population of the insular part of Newfoundland and Labrador

and in terms of values of a statistical life (Chapter 13). Our estimate of average WTP for the average reduction in mortality risk associated with MVCs we proposed depends on our choice of how to econometrically model the information obtained from the double-bounded dichotomous choice payment questions but a lower bound close to $100 per person per year can be quite reasonably derived from our analysis. From the mean WTP obtained from the analysis of the data in our sample we can extrapolate that a risk reduction policy that delivered this average change in risk proposed to our respondents (a reduction of 4.46 in 100,000 in the 10-year mortality risk rate) would be worth a total of either a bit less or a bit more than 20 million dollars per year, depending on how we choose to deal with those respondents expected to have a negative WTP by the estimated model. We also show how the WTP, and thus the value of the benefits of a risk reduction policy, is quite sensitive to how the (hypothetical) risk reduction is supposed to effected and who is supposed to benefit. We focus on the distinction between a policy protecting individual drivers only in the event of a MVC and a policy reducing the general risk of MVCs in the province. Although the mean WTP is lower in the former case, when aggregated over individual drivers, the benefit of that policy reaches close to 40 million dollars per year, while the public policy (aggregated, conservatively, only over households) would yield some 26 million dollars a year.

We also find estimates of the VSL to be dependent not only on the type of policy proposed but also on the scope of the risk reduction involved[3] and

---

[3]Because, as it is common in these type of studies, we find that the sensitivity of WTP to

averaging between 2 and 3.5 million dollars (again depending on the treatment of respondents predicted to be hurt by the policy). These VSLs fall perhaps a bit on the lower side but still reasonably close to the central ranges most commonly found in the literature. All of these calculations and choices are explained in detail in Part VII.

The main report closes with a discussion of the main results and their implications (Part 14) before Part VIII presents the main conclusions and limitations of this research, as well as suggestions for further work. The main body of report is followed by a list of references cited and a series of appendices dealing with some preliminary analysis of different particular aspects of the econometric analysis unrelated to the estimation itself of the benefits of a risk reduction policy. The last appendix (Appendix E) contains the text verbatim of the survey instrument used to obtain the data.

---

the size of the risk reduction is not sufficient for near-proportionality.

Part II

Background

Chapter 1

Moose in Newfoundland

and Labrador

Moose (*Alces alces*) is considered as the world's largest living deer. It can reach 2 metres at the shoulders, with long legs supporting its big barrel-shaped body. Males can carry about 20 kilograms of antlers and weigh up to 600 kg. Moose can be found throughout most of the northern forests of Canada.

Although they have now become a staple in Newfoundland's image and part of the seasonal hunting patterns of the province, moose are not native to Newfoundland. The first attempt to introduce moose took place in Gander Bay in 1878, although biologists cannot determine if this was a successful incorporation (Byrne, 2012). The population today is a reflection of the moose who were delivered to Howley in 1904. This resilient cargo spread throughout the

island in only three years from the initial drop site. Moose were not delivered for hunting purposes and are a result of an effort in the economic growth of Newfoundland. At the time, the Newfoundland government wanted to foster Newfoundland as a tourism destination, based on nature-related activities that included sport hunting and fishing. Efforts to promote hunting by non-residents was considered as a way to pay for new transportation systems arose in the late 1800s (McLaren, 2002). Some even referred to this successful introduction as a "new age in Newfoundland's development" (DOECNL, 2011; Byrne, 2012).

Much of the Newfoundland landscape is covered in forest, which provides a habitat that allowed for the newly introduced moose to grow successfully. These forests are less threatened by forest fire because of the climate of the province, as well as being large and old to be relatively immune to the most common diseases that may threaten the moose habitat. This allowed them sustainable and assured increases in population for a long period of time and contributing to their high density today (Thompson et al., 2003; McLaren et al., 2004). The natural predator of *Alces Alces* is the wolf. However, this species was deemed virtually nonexistent here since 1922. The main natural predator remaining in the area is the Black bear, whose density is much lower than moose density[1]. These predators attack only the calves of the species and do not thin the flock as effectively as hunting, so the moose population has exploded during the last century and is now estimated to be within the 125,000-150,000 range. However, the population of moose is no longer growing. Due to recent mitigation efforts

---

[1]The provincial population of black bears has been roughly estimated at 6,000 to 10,000 animals and is considered to be stable (Pelton et al., 1994)

and changes in the moose hunt per year, the moose population in Newfoundland is stated to be on the decline (DOECNL, 2013).

The high density of moose over time has also had a moderate effect of shaping the landscape of the province. Some of Newfoundland's previously flourishing balsam fir tree population has been converted to white spruce overgrowth. This can be attributed to the overpopulation of herbivores in the region. These moose have also successfully removed many hardwood species from national parks. This will relocate the densities and allow for less amiable living conditions for moose, causing them to move off reserves and closer to areas with danger of MVCs. The effect is a loss of natural growth patterns and hardwood diversity of forests. This effect is attributed to the over-grazing nature of high density populations and the "no hunting" restrictions based in these national parks. It is also speculated that small insect outbreaks could be attributed to the rapidly changing biodiversity due to overgrazing. This has caused the lumber industry in Newfoundland to lose its ability to flourish and one paper even refers to the industry being non-existent since 1950 (McLaren et al., 2004; Byrne, 2012).

Moose management and control is effected at the level of the Moose Management Area (MMA). On the Island of Newfoundland, there are 50 Moose Management Areas (MMA) where hunting takes place, including three enclaves. Those areas are captured in Figure 1.1. Moose densities range from 1 to 7 moose/km^2, with an average density island-wide around 1.7 moose/km^2 (Clevenger, 2011). The highest density of moose in the world has been registered in the Western part of the Island. Stantec (2010) showed a 6.82 moose/km^2 in MMA 43 on the
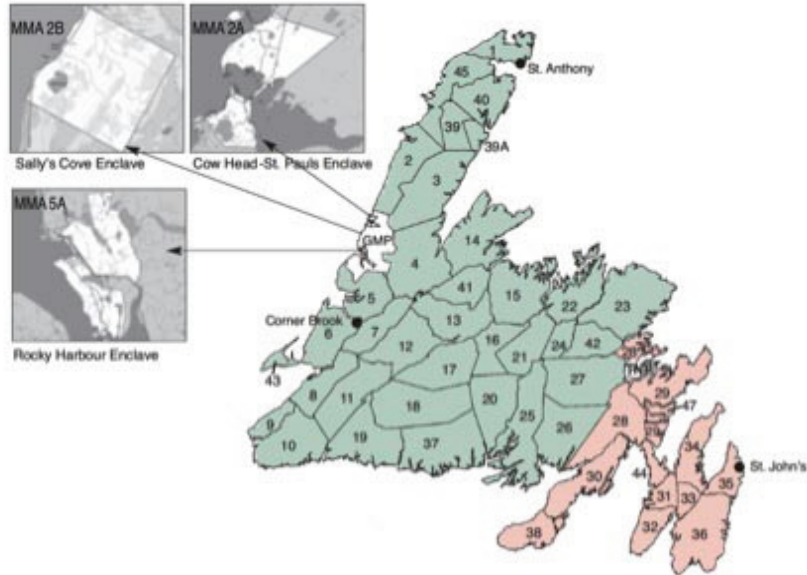
Figure 1.1: Moose Management Areas in Newfoundland.

Port-au-Port Peninsula, while a density close to 7 moose/ km^2 was registered in Gros Morne National Park (McLaren et al., 2000).

Moose provide significant benefits to the local residents and businesses. Hunting is a popular activity in Newfoundland and provides not only sport and recreational opportunities but also a source of food to many Newfoundlanders (Condon and Adamowicz, 1995). Big game hunting is an important and valued tradition in Newfoundland and Labrador. Hunting wild game also enhances a lifestyle of self-reliance and a tradition of living from the land. In Newfoundland and Labrador, the big-game hunting species are moose, caribou, and black bear. Moose hunting in Newfoundland began in 1935, just over thirty years after the species was introduced to the Island. More than 20,000 are now

25

being harvested annually. Moreover, the hunting success rate attracts visitors from the other Canadian provinces and from abroad, so moose hunting is likely responsible for spillovers onto other sectors of the tourism industry.

However, the high moose density in the Island generates some significant problems. For example, under high moose density levels, substantial costs to the forest industry could emerge (McLaren et al., 2000; Hörnberg, 2001; Timmermann and Rodgers, 2005). Additionally, high moose density is one key factor among several others (landscape, traffic volume, road design) that contribute to register a high number of moose-vehicle collisions in the area. It is unsurprising that such high density has led to conflicts between humans and moose, mainly because of moose-vehicle collisions: there are about 600-800 moose-vehicle collisions (MVC) each year, with an average of two fatalities per year (Clevenger, 2011). While moose is certainly an iconic mammal and one of the most recognizable species in this province, many Newfoundlanders are now considering moose as environmental bad in the province.

Chapter 2

Moose-vehicle collisions in Newfoundland

2.1 The issue of moose-vehicle collisions

A motor-vehicle accident is often a very serious issue that involves substantial economics costs, both human and non-human. One type of accident we are particularly interested in are collisions with large mammals, more specifically, Moose. These particular types of collisions cause more concern, because they are associated with a relatively high risk of mortality and morbidity and because they occur quite frequently in some areas (Clevenger et al., 2001; Christie and Nason, 2003; Rea, 2003).

We are particularly concerned with the collisions that occur in Newfound-

land, because this island component of the Canadian province of Newfoundland and Labrador boasts the highest moose population density in North America and possibly even the world (Clevenger, 2011). The number of collisions with moose can be correlated with economic costs in different ways. Most obviously, some of the relevant costs are directly associated with the loss of human life (human mortality) and with human injury (morbidity), as well as the repair costs of vehicles. Other, less immediately obvious, costs involve the cost of the precautions drivers take in order to reduce their risk of encountering a moose on the road and the impact of a potential encounter (driving at lower speeds if at all, driving at different times of the day, using different roads, improving the safety features of their vehicle, and so on), the costs faced by taxpayers when public measures are taken to reduce the risk of collisions (signalling, education campaigns, vegetation removal etc.)[1], and the loss of wildlife, which can result in overgrowth and loss of hunting profits (Weir, 2002; Olaussen and Skonhoft, 2011). These costs are placed on the individual, the economy in general, and on the natural environment in which the accident takes place.

2.1.1 Other jusrisdictions

Newfoundland is our area of study but motor vehicle accidents have drawn much attention in several other places as well. For example, official statistics show that there were 1,482 moose-vehicle accidents in New Brunswick in the five year period between 1995 and 2000 (Christie and Nason, 2003). These accidents

---

[1] Some of these measures would entail further social costs. For example, herbicide used for vegetation removal could have negative environmental impacts.

cost on average $1000 each to the drivers,[2] according to the Maintenance and Traffic Department of the New Brunswick Department of Transportation. The moose population in New Brunswick was estimated as 25,000 animals. The size of the moose population in Nova Scotia was estimated around 6,000 and in conservation status. However, Nova Scotia still registers close to 500 accidents annually (Christie and Nason, 2003; Duncan, 2004). A single Quebec national park averaged 251 wildlife vehicle collisions annually over a three year period, half of of which involved moose-vehicle accidents (Dussault et al., 2006).

The Atlantic Provinces are close and most easy to relate to the case of Newfoundland. However, the problem stretches much further. Yearly, there are an estimated 1.5 million collisions with ungulates in the US, costing 200 lives and over a billion dollars in personal and property damage. Studies of these types of expensive collisions were conducted, for example, in Nebraska and Colorado, highlighting the effectiveness of different mitigation techniques that may help to reduce the number of MVCs in a cost-effective manner (Reed et al., 1982; Mastro et al., 2008). The US Department of Transportation issued a detailed document of the cost and effectiveness of multiple mitigation measures, recognizing the high cost of such accidents and attempting to share their information with these states and other countries on how to prevent them (Huijser et al., 2008).

There are also several studies from Scandinavia. They focus on the specific effects of MVCs, with some studies also considering the effect of different types of mitigation efforts. One study (Olaussen and Skonhoft, 2011) estmates

---

[2]This estimate is considered to be an underestimate, considering that accidents below a certain threshold value, that varies among regions, are not reported.

alternative costs to property and vehicle damage, claiming that an astounding 270 million NOKs are lost annually to hunting losses and forest damage. It also estimates 500 million NOKs in vehicle and property damage. A few other studies have been conducted in countries such as Japan and India, making this a worldwide phenomenon (Elvik, 1995; Bashir and Abu-Zidan, 2006; Eldegard et al., 2012).

2.1.2 The issue in Newfoundland

Newfoundland lends itself to a higher rate of MVCs for several reasons. One of the foremost is that the areas with the highest densities of human populations are surrounded by areas with wet boreal forests, dominated by coniferous trees that moose use as shelter and nutrition in the winter months of the year. Newfoundland's roads also cross all the different types of its landscape. The Trans-Canada Highway, which is the main method of cross province travel, passes through boreal forests, mountains, wetlands, and marshlands, all of which lend themselves to ideal habitats for the high density of moose in this province.

An average of two people die every year from MVCs in Newfoundland, which is high, considering most areas do not have any fatal accidents involving ungulates annually (Geehan, 2011). This cost of life is considerable but there are many other costs to be considered. Moose fatalities are also an important consideration. Studies show that 4,800 moose are killed in these collisions each year, resulting in close to $600,000 in lost moose meat and $200,000 lost in tourism and natural resources. These studies state the total economic loss per year from

anywhere between $1,000,000 Canadian to $3,860,000, which represents a sub-stantial cost to society (Clevenger et al., 2001; Joyce and Mahoney, 2001; Rea, 2003; Clevenger, 2011). Table 2.1 shows some statistics related to moose related vehicle collisions in the province.

Year	Total moose related collisions	Collisions with at least one injury or fatality	Total number injured[1]	Number with major injury[2]	Number killed
2010	586	91	112	13	2
2009	629	82	99	7	6
2008	475	47	63	4	0
2007	430	77	114	2	0
2006	422	62	93	7	1
2005	376	61	80	6	0
2004	427	80	119	8	2
2003	363	61	87	5	2
2002	402	62	86	6	2
2001	352	47	64	6	1
2000	378	58	76	5	3
1999	356	n.a.	99	n.a.	0
1998	310	n.a.	61	n.a	2

1- Includes Major Injury 2 - Major injury means being hospitalized

Table 2.1: Moose-Vehicle-Collisions in Newfoundland and Labrador. Source: Newfoundland and Labrador Department of Environment and Transportation.

2.1.3 Temporal and spatial distribution of accidents

An important part of understanding wildlife-vehicle collisions is to analyze where and when these types of accidents occur most often. We will examine the temporal and spatial distribution of MVCs in many parts of the world as to determine at what times and in which places they are most likely to occur. This will help us understand Newfoundland's best way of mitigating accidents most cost effectively.

First, we consider the temporal distribution of MVCs in a broad sense. Seasonally, there is an ambiguous effect on MVCs. Study areas and other variables constantly change, so it is difficult to find a single season of the year during which drivers are most at risk of an MVC. The closest range of months that all MVC studies have found in common are the months between April and November, with June, July, and August normally exhibiting the highest rates of collisions. These months encompass the main growth period for ungulates and the period in which summer vacations and the largest number of motor-vehicle activity occurs. It is also common that harvesting or cutting of some certain roadside vegetation during this period could actually stimulate the growth of more attractive and nutritional foraging for ungulates, which may have an adverse effect on collision rates (Farrell et al., 1996; Joyce and Mahoney, 2001; Weir, 2002; Christie and Nason, 2003; Gunson et al., 2003; Rea, 2003; Mastro et al., 2008; Danks and Porter, 2010; Olaussen and Skonhoft, 2011). The combination of driving activity and moose activity also results in a typical distribution of accidents throughout the year, which is summarized in Figure 2.1.[3]

Although it is difficult to find a definite month of highest risk, the time of day in which vehicle collisions occur most is unambiguous. Most accidents occur on clear nights. Dusk and dawn are the most common times for ungulate activity to increase and for driver visibility to decrease. The increase in ungulate activities is due to increases in ungulate foraging activities. They are more protected in darkness, as they are much less detectable (Rea, 2003). Close to or over half

---

[3]See http://www.env.gov.nl.ca/env/wildlife/moose_vehicle_awareness.html
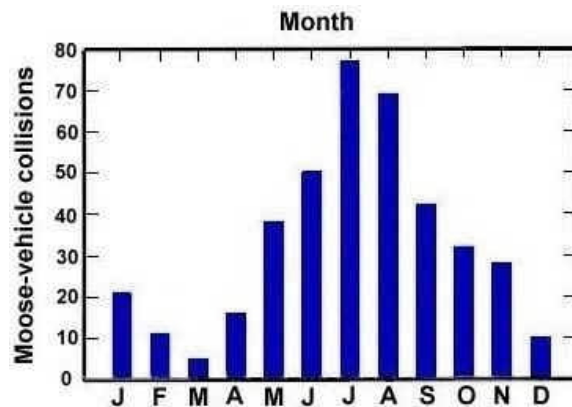
Figure 2.1: Typical seasonal distribution of moose-vehicle-collisions in Newfoundland. Source: http://www.env.gov.nl.ca/env/wildlife/collision.jpg

of all collisions are thought to occur during dark periods of the day (Dussault et al., 2006; Bashir and Abu-Zidan, 2006; Rattey and Turner, 1991). This is due to the fact that the drivers have low visibility during these times of day. There is an estimated 160 meter range of vision at night and, because ungulates are often dark and matte coloured animals they do not provide drivers with an adequate signal for stopping (Farrell et al., 1996; Joyce and Mahoney, 2001; Gunson et al., 2003; Rea, 2003; Danks and Porter, 2010).

MVCs are unequally distributed through time but also through space. There are certain places that are more prone to MVCs. Changes in speed limits and traffic volumes can be correlated with accident rates, and the foraging and migrating patterns of the moose may also be a factor. Increasing rates of traffic and higher speed limits have also been found to contribute to a higher frequency of accidents (Farrell et al., 1996; Christie and Nason, 2003; Clevenger, 2011;

Ford et al., 2011). For example, Danks and Porter (2010) suggest that for each additional 500 vehicles/day passing through an area prone to MVCs, there is an approximate increase of 57% chance of collision. They also suggest that for each incremental increase of 8 km/hr in speed MVC rates will increase by 35% (Danks and Porter, 2010). The spatial clustering of these accidents is most likely due to displacement of habitats, and foraging patterns of Moose. In particular, Newfoundland's main roads cut through most of its rural area, and it is more likely that higher traffic volumes in areas that are choice habitats for moose will cause a spatial cluster of high MVC rates (Weir, 2002; Clevenger, 2011).[4]

2.2   Measures to mitigate MVCs

Many forms of mitigation have been tried by management agencies in the rest of the world and also by provinces and states troubled with high MVC rates here in North America. Each approach has been extensively studied and provides certain benefits and costs, which will be considered in the scope of Newfoundland MVCs. There are a myriad of prevention techniques, all of which have their own costs and associated benefits. The most common practices to be analyzed include: fencing and crossing methods, GPS tracking and surveying, road signs and warning mechanisms, preventative education and primary warnings, vegetation removal and the effects of population thinning, and a yearly harvest quota.(Huijser et al., 2009a) In the particular case of Newfoundland,

---

[4]Salt pools along the roads have been shown to attract moose too (Grosman et al., 2011).

the *Strategic Plan of Newfoundland and Labrador Department of Environment and Conservation* for 2011-2014, shows different moose management strategies aimed at reducing the number of MVCs. First, the 2005-2006 strategy to deliver an educational awareness program consisting of radio-ads, billboards and bumper stickers continues. This effort encourage people to be mindful and observant of moose when traveling on the roads or highways of Newfoundland and Labrador.

2.2.1 Hunting as a density reduction measure

We should first consider the effects of hunting, often considered the most effective measure against MVCs. Some refer to hunting, however, as just providing a negligible effect through hunting quotas and downward pressure on ungulate removal (Clevenger, 2011). These disagreements may be on account of ungulate density and biological diversity (Mastro et al., 2008). We will consider specific studies on certain demographics and their effects on population reduction. Many jurisdictions, like Newfoundland, offer licensing for hunting moose. Hunters who purchase or qualify for these licenses accumulate to an overall quota that is taken into effect when determining changes in the overall moose population. The percentage in Newfoundland for a successful kill has been as high as 70% some years (Duncan, 2004). Hunting can indeed prove to be an effective measure in reducing moose populations, which in turn will thin moose density in some areas and result in fewer MVCs. AMEC (2004) mentions that, in a year during which 950 licenses were sold, 10% of the estimated moose popula-

tion was successfully hunted. Olaussen and Skonhoft (2011) suggest, after their study involving Norway's moose population, that a reduction of 13.5% of the entire population would have negligible effects on the overall hunting bag and save the environment and farmers from browsing damage and costs in MVCs. Similarly, when commenting on the particular effectiveness of hunting quotas, Clevenger (2011) suggests that only dramatic decreases in moose density will have an effect on the overall rate of MVCs. Licenses may be too wide spread to result in an efficient decrease in the population. In sum, just as when dealing with deer-vehicle collisions herd management could be considered a somewhat effective but controversial strategy (Hedlund et al., 2004). Therefore, we may have to look elsewhere for an effective mitigation technique.

2.2.2   Road signs and detection systems

The effectiveness of public awareness programs and road signing is relatively limited, although these are strategies very commonly used to deal with the problem of wildlife-related road accidents (e. g. Bond and Jones 2013). Road signs, marquees, and signal lights have been installed in many places to prevent MVCs. Simple signs that indicate a reduction in speed or that there are more moose present, while easy and relatively cheap to install, have not proven to be an effective measure. For example, it has been suggested that, at least in the case of deer, since passive road signs are used so frequently where these animals are present only occasionally, drivers probably ignore them (Putman, 1997; Sullivan and Messmer, 2003). In the United States, simple signs showed a net-

loss of \$18/km, with a 0% known effectiveness for reducing MVCs (Huijser et al., 2009b). Seasonal signs, although more expensive, provided a small improvement in MVC rates in some areas (Huijser et al., 2008; Huijser et al., 2009b).

One type of road sign that has proven effective is the method of advanced detection systems. When the ungulate crosses an infrared beam, or triggers a pressure panel, a road sign is programmed to inform drivers that there are moose approaching or will inhibit their passing. These technologies can now also transmit warnings within the vehicle, allowing for an on screen display or an audible warning. Huijser et al. (2008) cited these measurements as being up to 95% effective, with a net benefit of roughly \$3,000 (\$US2008) (Mastro et al., 2008; Clevenger, 2011). All of these methods are aimed at enhancing drivers' awareness of danger and, although some are not proven to be as effective as others, they all increase the effectiveness of other mitigation strategies (Mastro et al., 2008; Huijser et al., 2008; Ford et al., 2011; Clevenger, 2011).

2.2.3 Vegetation removal

Vegetation removal has also been suggested as one of the best management options involving MVC rates. It can achieve positive results in two ways. Moose, as herbivores, will browse the areas with the highest density of the appropriate vegetation. Removal will therefore discourage them from staying in these areas and possibly prevent MVCs. Roadside vegetation removal is also an aid to drivers who would otherwise have trouble seeing past the edge of the road and spotting coming moose. It has also been observed that the removal of vegetation

along the road will make moose more sceptical about approaching the road, because it makes the traffic more visible. It is a debated method, though, because, in some circumstances, removing vegetation to increase visibility at the wrong time of year can yield attractive, nutrient rich food for moose and other ungulates.

Most road sections will pass through rural areas where human population densities are highest. This will make the initial operation and further maintenance of a vegetation removal project particularly expensive. Rea (2003) claims that, with the appropriate timing structure and a targeted program, vegetation removal can be more cost-effective than expensive fencing methods (although it is also very effective). For example, cutting vegetation at popular maintenance times such as the summer, around July, will yield large-shooted plants closer to roadsides in the following Winter. Their growth will be stimulated from the run-off of early falling snow. They are rich in proteins and do not carry many defensive plant hormones, making them more attractive to ungulates (Rea, 2003; Danks and Porter, 2010). Having plants that are high in potassium and water further away from the roadside will also cause moose to interrupt their browsing periodically to search for salt rich water like the water appearing closer to or on a road. This reinforces that choosing the appropriate vegetation to cut at the right times will yield the most effective results (Rea, 2003; Danks and Porter, 2010).

The degree of success for programs based on roadside vegetation removal is variable with percent reductions in MVCs ranging from 20% to 56%, be-

ing considered a medium-cost option (Wiseth, 1991; Rea, 2003; Huijser et al., 2009b).

2.2.4 Fencing strategies

There has been abundant research into the effectiveness of strategic fencing measures as a mitigation tool for MVCs. An effective fence will reduce or eliminate the movement of ungulates into the right of way of drivers. These fences are generally considered the most effective tool for mitigating MVCs (Seiler, 2005; Leblond et al., 2007), as well as collisions with other members of the deer family (Romin and Bissonette, 1996; Bissonette and Rosa, 2012),[5] though their installation and maintenance costs are relatively high (McLaren et al., 2004; Danks and Porter, 2010).

Fences can be made of wire, chain-links, or electric material. Huijser et al. (2008) suggest that these fences should be 2.4 to 2.7 meters in height to successful prevent a moose from crossing above. They must be attached to pressed wood posts or metal posts, the latter being more expensive and more durable than the former. Fences should be accompanied by a crossing method for wildlife, which will prevent the animal traffic from being bottle-necked into a certain area.[6] It is also a good practice to have fences closed tightly against areas where they may end, such as rock formations or a crossing opportunity (Duncan, 2004; Huijser et al., 2009b).

---

[5] In the context of deer-related collisions, "the only widely accepted method with solid evidence of effectiveness is well-designed and maintained fencing, combined with underpasses or overpasses as appropriate" (Hedlund et al., 2004, p. 122).

[6] See Section 2.2.5 on details for the effectiveness of crossing applications.

This method of moose-vehicle accident reduction is most effective when executed in large portions given that the browsing range of large mammals tends to be massive and encompassing. It is also a necessary condition that the fences cover both sides of a given road, given that one sided fences have proven to have a much weaker effect in reducing MVCs. The effectiveness of this mitigation measure is considered very high and ranges between 75% to 100% when combined with the use of a safe crossing opportunity for wildlife. For example, MVCs decreased by 80% after about 1,300 km of main roads were fenced in Sweden (Lavsund and Sandegren, 1991). Given this effectiveness, a given situation that requires this installment should have the benefits of installation outweigh its costs by at least 36% (Reed et al., 1982; Clevenger et al., 2001; McLaren et al., 2004; Duncan, 2004; Huijser et al., 2009b). This method, even in the absence of safe crossing opportunities has an effectiveness of 86% with a cost of $6,289 per percent reduction (in 2009 USD), giving it a much higher cost per percent effectiveness rating than many other mitigation efforts. However, it has the highest effectiveness of any when included with crossing and may be the only mitigation required in some road structures.

2.2.5   Safe crossing

Many studies include the use of overpasses and underpasses as a useful tool to increase the effectiveness of fencing in areas where mitigation is required. These safe crossing mechanisms are built as vegetation covered overpasses or road tunnels and bridges to provide safe passage for ungulates from one gap

of a fenced area directly into another. Underpasses should be spacious enough for ungulates to pass safely below without interacting with vehicles, on study suggests about three meters height and 30 meters in length was adequate for their purposes (Reed et al., 1982). These underpasses may be eight meters high and five meters wide to meet a minimum specification of a large mammal underpass (Huijser et al., 2008). Overpasses must only be long enough to cross the road which is in need of mitigation. The most cost-effective way of producing these overpasses may be by building them in hourglass shapes in more places as opposed to a large, single, free-floating platform (Duncan, 2004).

Overpasses may be the only truly effective measure of mitigating MVCs specifically. Huijser et al. (2008) state that overpasses alone have an effectiveness of reduction between 79 - 97%, depending on the location. The increase of effectiveness from these structures can be between 14% and 25% when combined with large mammal fencing, usually indicating a 100% overall reduction. Reed et al. (1982) suggest that the excess of cost from the construction of over or underpasses must mitigate at least 20 accidents per year over a 1.6 km length of road. The cost of this measure can range between 1 million in 1992 USD for a 15m wide overpass with an hourglass shape to 11.5 million in 2005 USD for a 45m wide bridge spanning over a four lane highway. Although they are expensive initially, a typical overpass has a lifespan of about 75 years and requires little maintenance once vegetation from the surrounding area has been integrated (Huijser et al., 2008; Huijser et al., 2009b).

41

2.2.6 What has been done to prevent MVCs in NL?

As mentioned in Section 2.1.2, MVCs cause around 4,800 moose deaths per year, and 2 human fatalities here in Newfoundland with an average of 9.5 hospitalized with a serious injury. According to the Department of Transportation and Works there were an average of 423.5 collisions between the years 1998 and 2010.[7] These collisions result in an economic cost between \$1 million and \$3.5 million per year, and should be considered a very serious problem for Newfoundland Drivers. It follows that there should be effective mitigation strategies in place around the province in order to prevent these accidents and save the lives and capital lost each year.

The policies that the government has taken in mitigating have been based almost entirely on public awareness or are reliant on changes to the hunting policy in the province. The government website for moose reduction states that "Care and attention when driving remains your best defense against a moose-vehicle accident" (DOECNL, 2013). Public Awareness campaigns must be strict and employed over long periods of time in order to be effective, and in most cases are not recommended as the most effective form of mitigation. Other researchers go on to critique this sole method of mitigation as a misinterpretation of other methods. Clevenger (2011) criticizes that the government assumes that a new or improved mitigation program to influence animal behaviour should be 100% effective (Clevenger, 2011; Huijser et al., 2008; DOECNL, 2013).

Another method the current Government of Newfoundland has employed is

---

[7]Elaborated by the authors from statistics shown in Geehan (2011).

the clearing of roadside vegetation. Clearing occurs along highways scheduled at times that will influence shrub growth and provide better foraging opportunities for moose in periods of higher mobility. As mentioned in Section 2.2.3 these programs are most effective when they are planned to be perfectly executed and maintained appropriately (Clevenger, 2011).

Probably the most effective measure taken to date has been the modification of yearly hunting licences. Dramatic decreases in moose population and density may help to provide safer driving conditions. For the 2011 - 2012 hunting season, an additional 5000 moose added to the yearly quota, providing a total of 5020 moose licences across the province and a yearly quote of 33,440 moose. The hunting season was also lengthened by a weeks time in order to allow more big game hunters and organizations to take advantage of the world-renown Newfoundland Moose Hunt (DOECNL, 2013).

Newfoundland's licence quota system was introduced in the 1960s, when moose management became more complicated (McLaren, 2002). The Wildlife Division at the Newfoundland and Labrador's Department of Environment and Conservation manages big game populations to ensure hunting tradition may continue indefinitely in a sustainable manner. The Wildlife Division determines the maximum number of moose that can be safely harvested each year to prevent population declines, while at the same time making sure populations are being kept at levels that will not degrade the habitat and possibly lead to a population collapse.

Figure 2.2 shows the annual moose quotas and harvest numbers for the
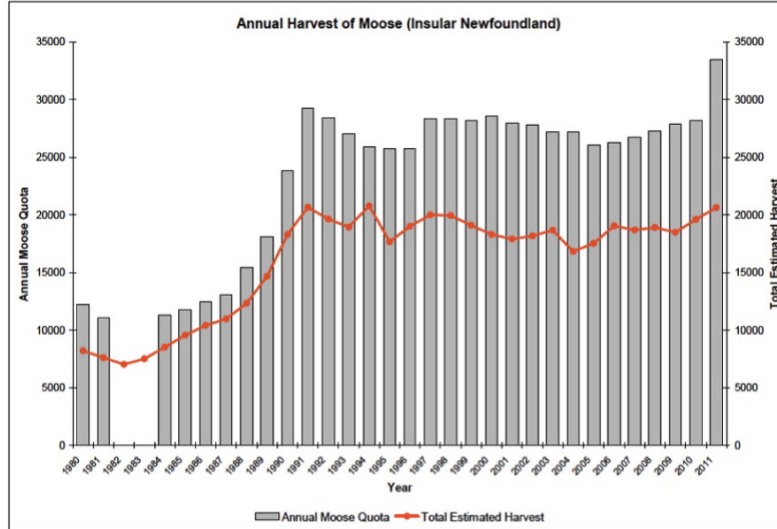
Figure 2.2: Annual moose quotas and estimated harvest. Source: Newfoundland and Labrador Department of Environment and Conservation

period 1980-2011 on the island of Newfoundland. The figure shows that during the mid-eighties the annual harvest matched the level of quotas more closely than in other periods of the series. At the end of eighties and the beginning of nineties both the quotas and the harvest experienced substantial increases. From 1992 to 2010 the trend is much more stable than in the previous period. However, the recent 2011-12 moose quota was increased by 5,020 over that of previous year for a total quota of 33,440 moose licences. This lethal control option of moose the population can be effective in terms of reducing the number of moose but it is unclear that it can be effective in reducing accidents on the road. Although the substantial reduction of MVCs in 1992 (decreasing from almost 900 collisions in 1991 to 670 in 1992) was directly linked to the increase

44

of hunting quotas, Clevenger (2011) claims that only sharp decreases in moose densities would affect the overall rate of MVCs. Furthermore, this policy could result in substantial costs to tourism operators, insofar as a reduction in the hunting success rate it is expected.

2.3   Options for a mitigation strategy in New-foundland

It has been shown that there are several ways of reducing MVCs, some for only hundreds of dollars of well planned execution. With the current influx of human population and government capital it should logically follow that more adequate accident mitigation efforts should be made. Furthermore, although public opinion appears to be divided around the issue, there has been some growing concern about the fact that the provincial government could be doing more to protect drivers from the risk of crashing with moose.

In particular, there was recently a court order that Newfoundland extend its class lawsuit cases to include moose accidents within a ten year period. In 2011 Ches Crosbie Barristers, a St. John's law firm dealing in personal injury cases, filed a highly publicized class action lawsuit against the province of Newfoundland and Labrador, claiming that the provincial government is responsible for the injuries suffered by those who crash into moose on the province's roads (Guo, 2011). Ches Crosbie Barristers requested a series of statistics (Geehan, 2011) about the issue using a freedom of information requests. After receiving

the statistics, some of which were mentioned in this paper, the law firm's appeal was granted and the time period for being able to file a class action suit was extended from just two years to ten years. In the end the class action lawsuit was unsuccessful but it highlighted some of the most extreme sentiments towards the oficial policies aimed at dealing with the issue of MVCs in the province.

The government also began extending hunting seasons and offering larger quotas as an effort to reduce MVCs after their fault had been brought to light. The population has surely noticed that mitigation is not appropriately applied here and the government of Newfoundland and Labrador government will have to adjust and provide a more thorough mitigation plan (Geehan, 2011).

Part III

The theory and praxis of valuing road traffic risks

Chapter 3

Valuing traffic risk reductions: The value of a statistical life

Since the pervasive scarcity of resources forces policy-makers to prioritize between policies, those in charge with designing policies dealing with building, improving, and maintaining transport infrastructure need to have a indication of the benefits of these activities in order to be able to compare them with their cost.[1] The use of a common metric for benefits and costs makes it easier to assess the merits of alternative policies, which enables decision makers to

---

[1] This section borrows extensively from the excellent review of the issues considered provided by Andersson and Treich (2011).

allocate taxpayers funds more efficiently. In this section, we present the theoretical basis for deriving that type of measure and we consider the theoretically expected relationships between its value and a series of influential factors. We also describe the main trends in terms of empirical results obtained in this field of the non-market valuation literature.

Most of the costs categories associated with transportation infrastructure can be easily monetized. For example, the expenditures needed to improve the width or the quality of a road in order to increase its safety and to reduce travel time are easily observable, since they involve the purchase of inputs that are traded in markets and have easily observable, if often distorted, market prices. On the other hand, many of the costs and benefits involved in the decisions about transport infrastructure are not so easily monetized, since they do not have market prices which could signal their marginal benefits. In particular, and this is the aspect of transportation policy that concerns us, while the expected benefits from avoided material damage caused by vehicle collisions can be, if imperfectly, calculated by resorting to the observation of available market prices, other, likely much more substantial in many cases, benefits that society derives from the reductions in the risk of collision are not observable.[2] Therefore, non-market valuation techniques (Hanley and Spash, 1995; Boardman et al., 1997; Weimer and Vining, 2005) must be applied in order to approximate the full

---

[2] In particular, it is quite difficult to place a monetary value on the "human" costs of road traffic accidents (such as pain, suffering, and bereavement). Schwab Christe, and Soguel (1996) constitute one of the few attempts to isolate those costs of road accidents by asking respondents about their willingness to pay to reduce the likelihood of an accident considering themselves either as potential victims of the accident or as relatives of potential victims.

value of improvements in road safety. A key benefit that requires this type of techniques is the protection of life.

In general, many policies and regulations, not only only those related to road safety, are intended to, directly or indirectly, protect human life. And just as in the case of road safety policies, in order to meaningfully compare their benefits and costs, society's willingness to sacrifice resources in exchange for further safety, that is, the social willingness to pay (WTP) for risk reductions must be considered. Economists usually report the estimates of this WTP in terms of a so-called "statistical life" (VSL). This concept has been defined as "an estimate of the monetary benefits of preventing the death of an unidentified person. It is the maximum amount government agencies will pay to save one life" (Brady, 2008, p. 541).

Sometimes, less rigorously, economists refer to the monetary value of reducing mortality risks as the "value of life", which is an unfortunate shortened version of the more common and more accurate "value of a statistical life". This jargon has a clear and precise meaning for economists and is quite uncontroversial among them.[3] However, it may be controversial for others, since it seems to imply that a finite monetary price can be attached human life, while human life should be instead seen as "priceless".[4] In fact, apart from the many technical problems associated with managing risk and valuing risk reductions, the

---

[3] Note, however, that there are alternative expressions to refer to the concept that the VSL tries to capture. For example, terms such as the value per statistical life (Hammitt, 2000b), value per life saved (Jones-Lee, 1976), and value of prevented fatality (Jones-Lee, 2004) have been used before (Hammitt and Treich, 2007), although they have not caught on as VSL itself.

[4] See, for example, Ackerman and Heinzerling (2004).

reluctance many feel to price human life is one of the key hurdles facing those in charge of policy design. "Of all the difficulties that surround the attempt to calculate the economic "value of a life," one of the thorniest is a moral one, namely, whether it is morally permissible to place any "price" on a human life" (MacKinnon, 1986, p. 29). This problem has led Cameron (2010) to propose the elimination of the term "value of life" even from the specialized literature and substitute it for something less confusing for non-economists (*willingness to swap*).

The reason why economists in general feel relatively comfortable with the use of the term VSL is that they readily understand that the use of this expression is based on the important distinction between an *identified life* and a *statistical life*, first considered by Schelling (1968):

> It is not the worth of human life that I shall discuss, but of 'life-saving,' of preventing death. And it is not a particular death, but a statistical death. What is it worth to reduce the probability of death - the statistical frequency of death - within an identifiable group of people none of whom expects to die except eventually? (Schelling, 1968, p. 113).

It is the value of preventing unidentified deaths *ex-ante*, once the risk is identified and a remedial policy is under consideration, not the value of preventing any specific identified individual's death *ex-post,* once the death has occurred,[5]

---

[5]The appropriate perspective when evaluating policies involving risks to life is the "tastes and preferences (and views of the future) as these exist at the time of the social decision."

that is relevant when trying to decide how much to spend on risk reduction policies (Brady, 2008). This difference between an identified and a statistical life has been illustrated by how easy it is to raise funds for the treatment of a young girl, who needs expensive care to prolong her life by a short time period relative to how difficult it can be to get support for, say, a tax-rise to finance health-care expenditures that would reduce the mortality risk for many, but unidentified, individuals like that identified girl (Schelling, 1968; Pratt and Zeckhauser, 1996). Attempting to value actual identified lives using the WTP approach is certainly difficult and probably also meaningless (Brady, 2008).[6] In that sense, "there is no 'value of a statistical life'; there are only values for the reduction of statistical risks" (Sunstein, 2004b, p. 392).

Yet another way to view the VSL is as society's equilibrium income-risk exchange rate (Brady, 2008). Indeed, the VSL is theoretically defined as the marginal rate of substitution between mortality risk and money. If we label the risk as r and the income as Y, the formula for the VSL is formally given by:[7]

$$VSL \ = \frac{\partial U/\partial r}{\partial U/\partial Y} = \frac{\partial Y}{\partial r} \qquad (3.1)$$

The VSL is then a ratio in which the numerator is the marginal utility of a small reduction in mortality risk (usually a very small value, $\partial U/\partial r$) and the

---

(Jones-Lee, 1979, p. 251). Under this *ex ante* perspective, it is not know whose lives will be lost or saved at the time of valuing the risk (Trumbull, 1990).

[6] The willingness to pay approach assumes that the best judge of the preferences towards risk is the very individual. This, more recent approach to the valuation of risks to life, supersedes the now outdated human capital approach, whereby an individual's life would be valued by what the individual was expected to contribute to society, through wages earned, etc. (Whittington and MacRae, 1986).

[7] See Section 3.1 for a fuller formal model to derive the formula for VSL.

denominator is the marginal utility of a small change in income (normally a much smaller value $\partial U/\partial Y$, which results in the ratio being very large).[8] In other words, the VSL is the monetary value of a (small and similar among the population) mortality risk reduction that would prevent, once aggregated, one statistical death. Therefore, it should not be interpreted as how much individuals are willing to pay to save an identified life:

In sum, the question is not how to value prevention of a specific death but how to value small changes in mortality risk across a population (Hammitt, 2000b, p. 1396).

The practical task of estimating the value of the small reductions in risk can be conducted using a variety of valuation techniques (see Chapter 5), with the dominant approach nowadays is based on directly eliciting individuals' willingness to pay (WTP) using the CVM. In practice, moreover, the VSL is not measured as a derivative but rather as the ratio of the WTP for a specific (small but non-marginal) risk reduction to the absolute level of that reduction.

A hypothetical example, adapted from Andersson and Treich (2011), of the calculation of the VSL illustrates the main notion behind the concept. Imagine that in a jurisdiction where 100,000 identical individuals live, a risk-reduction project is being considered consisting on improving road safety. The baseline risk is known to be given by the fact that on average 5 people die every year on the roads. The project is expected to reduce that risk from 5 to 2 expected

---

[8] The WTP-approach to value mortality risk reductions was initially introduced by Drèze (1962) in French and the concept was adopted more widely only after Schelling (1968)'s contribution, while Mishan (1971) and Jones-Lee (1974) further developed its theoretical basis within the expected utility framework.

fatalities per year. If each individual were willing to pay $150 a year for this risk reduction, the VSL would be $150×100,000/3 =$5 million. A total of $15 million could be collected to save 3 statistical lives, so the value of a statistical life would equal $5 million.

It is important to note that this approach to the valuation of risk reductions is valid only for small reductions in risk when the baseline risk is also small. Thus, the marginal trade-offs individuals make between income and risk at low levels or baseline risk cannot be linearly extrapolated to non-marginal changes in risk and higher levels of baseline risk. For example, that someone is willing pay $100 for a 0.01 percent risk reduction does not mean that she is also willing to pay $200 for a 0.02 percent reduction (Brady, 2008).

As pointed out by Viscusi (2013), calculations of WTP for risk reductions involve comparisons of individual utility when alive and when dead. Utility functions being unique only up to a monotonic transformation of each other, "adding a constant to the utility functions or multiplying the utility functions by a positive constant does not alter their structure, but it does affect their level. To give the difference in utility functions cardinal significance, the formula for VSL divides the difference in utilities by the expected marginal utility of income, which serves to normalize the units of the utility-difference expression. As a consequence of this mathematical structure, VSL serves as a cardinal measure of preferences with respect to fatality risks" (Viscusi, 2013, p. 1736).

If small changes in mortality risk were treated by individuals like any other consumption good, there should be no or little controversy to use the concept

of WTP to value a reduction in the risk of death (Mishan, 1982; Freeman, 2003). There are several reasons, however, why the valuation of reductions in mortality risks is a particularly complex component of a CBA. One important issue arises because of the commonly observed fact that individuals misperceive, sometimes grossly, mortality risks, which leads to inconsistent WTP estimates (Hammitt and Graham, 1999). Another problem is that, although the standard preference framework relies on the assumption of purely self-interested behavior, individuals may care about the risks to life of others, so altruistic concerns may matter in the WTP for reduction in mortality risks. The distributional effects of policies based on the results of applying the WTP approach, which may for instance give disproportionate weight to wealthier, older, or less healthy citizens, most of all in a context in which differences in the choices about exposure to risks often constitute a controversial issue too. Policy-makers are often reluctant to use different VSLs to account for differences in the types of risks and in the characteristics and choices of individuals, as noted by several authors who focus on the heterogeneity of VSL and challenge the notion that it is desirable to use a common value for the VSL in policy design (Alberini et al., 2004; Aldy and Viscusi, 2008; Aldy and Viscusi, 2007; Baker et al., 2008; Baker et al., 2009; Carlsson et al., 2010).[9]

As expressed by Zhang et al. (2005, p 154):

Estimating the value of life to use for policy purposes is an extremely

---

[9] See also Armantier and Treich (2004), for a critique of the usual procedure of aggregating unweighted WTP for risk reductions, which is only valid when individuals share the same marginal value of money.

difficult task. The evidence is extraordinarily uncertain and the estimated VSLs vary enormously. At the same time, it is important to emphasize that the VSL has probably been studied in more depth than any other "plug in" number.

The next sections summarize the main issues surrounding the use of the VSL as a measuring rod to inform the comparison among risk reduction policies. We first outline the standard VSL model to provide a brief theoretical derivation of the concept of VSL and explain the theoretical predictions about how it is affected by individual characteristics. The results of several empirical applications are discussed afterwards.

3.1 The standard VSL model

Consider an individual who maximizes her (state-dependent) expected indirect utility which is given by:[10]

$$V = pu(w) + (1-p)v(W) \tag{3.2}$$

where p is the probability of surviving the period, $u(w)$ is the utility she derives from her wealth w if she survives the period, and $v(w)$ is the utility she derives from her wealth w if she dies.[11] It is standard to assume that both u

---

[10] We adopt the notation used by Andersson and Treich (2011) and also follow Hammitt (2000). The original derivations can be found in the earlier contributions by Drèze (1962); Jones-Lee (1974) or Weinstein et al. (1980). A very useful reference for the theoretical models summarized in this section is also Pratt and Zeckhauser (1996).

[11] This utility would normally refer to the utility of a bequest, since the individual herself could not enjoy her wealth after death.
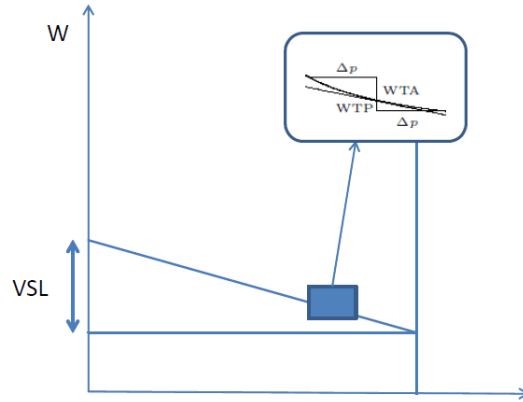
Figure 3.1: Geometric interpretation of the value of a statistical life (VSL).

and v are twice differentiable with

$$u > v, u\prime > v\prime \geq 0, u" \leq 0 \text{ and } v" \leq 0 \tag{3.3}$$

which means that state-dependent utilities are increasing and weakly concave. At any wealth level, both utility and marginal utility are larger if alive than dead. Under these standard assumptions, indifference curves in the wealth-probability of surviving space (w and p) are decreasing and strictly convex, as shown in Figure 3.1, adapted from Andersson and Treich (2011, p. 398). The VSL is the slope of the indifference curve between survival probability (p) and wealth (w). That is, the VSL is the marginal rate of substitution between p and w. The willingness to pay, henceforth denote WTP (willingness to pay, henceforth denoted WTA) represents the maximum amount that an individual is willing to pay (accept) for a mortality risk reduction (p) (increase).

The WTP for a given mortality small risk reduction of size $\Delta p \equiv \varepsilon$ is denoted C(ε) and is given by:

$$(p + \varepsilon)u(w - C(\varepsilon)) + (1 - p - \varepsilon)v(w - C(\varepsilon)) = V \qquad (3.4)$$

where V is the indirect utility defined in Equation 3.2. The WTA for a mortality risk increase $\Delta p \equiv \varepsilon$ is denoted P(ε) is given by:

$$(p - \varepsilon)u(w + P(\varepsilon)) + (1 - p + \varepsilon)v(w + P(\varepsilon)) = V \qquad (3.5)$$

The WTP and WTA are shown in Figure 3.1. Both WTP and WTA should be sensitive to the size of ε. For small ε WTP and WTA should also be expected to be nearly equal in size and near-proportional to ε (Hammitt, 2000a).

The VSL measures the WTP or the WTA for an infinitesimal change[12] in risk and is thus the marginal rate of substitution (MRS) between wealth and the probability of survival, that is, (the negative of) the slope of the indifference curve in Figure 3.1 at point (w; p). Mathematically the VSL is defined as:

$$VSL \equiv \frac{-dw}{dp} = \frac{u(w) - v(w)}{pu'(w) + (1 - p)v'(w)} \qquad (3.6)$$

Because of the properties assumed in Expression 3.3, the VSL is always strictly positive and it can be seen from Equation 3.6 that the size of the VSL

---

[12]For large changes in risk or in cases when the risk change has no close substitutes (Hanemann, 1991), there can be differences between WTP and WTA. We will continue our discussion in terms of WTP, which is used in the rest of the study as recommended by Arrow et al. (1993).

depends separately on the characteristics of the baseline risk (through p) and on the characteristics of the individual through u, v and w.

The value of a statistical life (VSL), as a measure of the trade-off between wealth and risk that individuals are willing to make, has now become a standard measure for valuing life-saving policies and policies that involve the reduction of risks to human lives in general. It should be again stressed that the VSL is not a measure of the worth or value of any particular life, nor is it a measure of the average value placed on a life. Rather, the VSL can be thought of as how much money a group of similar people would be willing to pay to reduce a risk to the extent that one fewer of them would be randomly killed by that risk.

3.1.1 The *dead-anyway* effect and the *wealth* effect

From Equation 3.6 two standard effects can be identified. The *dead-anyway* effect (Pratt and Zeckhauser, 1996) describes how VSL increases with baseline risk $(1-p)$, that is, how the size of the VSL decreases with survival probability p. Intuitively, someone facing a great risk of dying has little incentive to limit the spending of her wealth on trying to reduce her death risk, since she is unlikely to survive and be able to spend it on other ways. In Equation 3.6, the value of the numerator is independent of p and a decrease in p reduces the value of the denominator (because of the aforementioned assumption that $u\prime > v\prime$, which means that the marginal utility of wealth is larger when alive than when dead). Therefore, an increase in p will tend to decrease VSL. For example, in a game of Russian roulette, the standard analysis suggests that one should be willing

to pay more to reduce the number of bullets in a six-chamber revolver from 5 to 4 than from 2 to 1 (Zeckhauser and Viscusi, 1990).

Initial risk levels are typically ignored when conducting cost-benefit-analyses, however, which is acceptable if the policies considered involve remote mortality risks. For this reason, it has been suggested that the VSL approach should be reserved to estimate the benefits of small changes in minute death probabilities (Viscusi, 1992).

Additionally, the *wealth effect* predicts that the VSL will increase with wealth w. This effect can be explained by the fact that wealthier individuals have more to lose if they die (the numerator in Equation 3.6 is increasing in w, since $u\prime > v\prime$) and because the utility cost of spending is smaller for them due to the weakly diminishing marginal utility of wealth (equivalent to the notion of having risk aversion), which implies that the denominator in Equation 3.6 does not increase (because $u" \leq 0$ and $v" \leq 0$).[13]

Because of these two effects, the size of the VSL increases as one moves upward and leftward along an indifference curve such as the one in Figure 3.1.

3.1.2 Risk aversion and background risks

One of the criticisms of the approach that estimates VSL by comparing wage levels across occupations with different risks is based on the idea that compensating this analysis of wage differentials underestimates the average VSL in the

---

[13] The fact that the VSL is generally known to increase with wealth leads some agencies to consider increases in VSL throughout time to account for trending increases in income. Paradoxically, no agencies consider attaching different VSLs to individuals in the same cohort with different levels of wealth or income (Sunstein, 2004b).

population, because those individuals who choose to work in more hazardous industries are less risk averse than the rest. This suggestion, however, requires a more precise specification about what is meant by "less risk averse" (Andersson and Treich, 2011). For state-independent utility functions, it is usual to define risk aversion by the coefficient of curvature of the utility function (Pratt, 1964; Arrow, 1971, Ch. 3). However, since the above framework is based on state-dependent utility functions, because it assumes that the utility derived from wealth depends on the state of nature (whether the individual is alive or dead), the characterizion of risk aversion is less clear (Karni, 1983).

Eeckhoudt and Hammitt (2004) examine the relationship between aversion to financial risk and WTP to reduce mortality risk and find that this WTP is sensitive to other characteristics of the utility function. They show, using the standard model (Equation 3.2), that the effect of an increase in risk aversion increases the VSL when the marginal utility of bequest is zero and in a few other situations but in general, the effect of risk aversion on the VSL is ambiguous. Moreover, Eeckhoudt and Hammitt (2001) and Kaplow (2005) show that high coefficients of relative risk aversion $(\frac{-wu"(w)}{u'(w)})$ usually imply high values of the income elasticity of the VSL.

Eeckhoudt and Hammitt (2001) also show that, background mortality and financial risks decrease VSL under reasonable assumptions about risk preferences with respect to wealth in the event of survival and death. Andersson (2008) extends their analysis showing that, when individuals perceive the risks to be mutually exclusive, the background risk increases VSL.

The description of the VSL model we presented so far is based on a single-period approach, which suffices for our purpose of illustrating how the VSL informs the cost-benefit analyses involved in the reduction of death risks associated with road traffic. Andersson and Treich (2011) provide an excellent description of the more realistic multi-period models, whereby individuals have preferences over probability distributions of the length of life and over consumption levels at each period of their lives.

3.1.3 Altruism

Although our brief account of the VSL theoretical literature leaves out many interesting extensions and further arguments, we would like to briefly touch on one additional aspect of the theoretical modeling of VSL, namely the issue of altruism in individual preferences. This is because our survey instrument explicitly considers different versions of the payment scenario that will allow us to analyse, to some extent, the effects of altruistic concerns on WTP.

Although estimates of the VSL are used to value risk reductions provided as public goods, most studies derive WTP for reductions in private risks only, in practice allowing policymakers to value public risk-reduction measures using private WTP estimates (Brady, 2008; Alberini and Ščasný, 2013). This approach is valid assuming that there is no difference between WTP for reductions in private versus public risks or that any difference between them does not matter. The first assumption would be quite strong, since it is difficult to accept that individuals are not at all willing to pay to improve the safety of

others, particularly in the case of family members and friends. However, the second notion, that altruism could or should be ignored even if it exists, has been considered worthy of further analysis (Brady, 2008).

The relationship between altruism and the CBA of risk reduction policies has been addressed by some of the earliest contributors to the literature on VSL (Schelling, 1968; Mishan, 1971; Jones-Lee, 1976; Needleman, 1976). And indeed a key result raised by Bergstrom (1982) is that in the case of *pure altruism*, whereby an individual's utility increases in everyone's utilities, so that "Peter's welfare is affected by what Paul values, not by what he feels Paul *should* value" (Brady, 2008, p. 543), every Pareto optimum in the altruistic economy must also be a Pareto optimum in a selfish economy.[14] This means that the presence of pure altruism should not lead to any adjustment upward or downward of the VSL but should be kept the same as if individuals were selfish. Intuitively, a pure altruist benefits when someone else's risk is reduced but is also harmed when a financial cost if imposed on others. The sign of one person's altruistic valuation for another is then the same as the sign of the net private benefits to the other. Therefore, pure altruism does not alter the sign of the social net benefits (Andersson and Treich, 2011).[15] Jones-Lee (1991) further showed that people's WTP for others' safety should only be considered when their altruism is "exclusively focused upon other people's safety" (Jones-Lee, 1991, p. 91), that is when individuals are "safety paternalists" or "safety-focused" altruists, such that for example "Peter is willing to pay more for improvements in Paul's safety

---

[14]See also Milgrom (1993).
[15]This key argument was generalized by Bergstrom (2006).

than for improvements in other aspects of Paul's well-being" (Brady, 2008, p. 541).

Next, we will be considering some of the developments of the empirical literature dealing with the estimation of the VSL. The reader is directed to, for example, Andersson and Treich (2011) for further details on the extensive set of theoretical research questions surrounding the valuation of risk reductions and the CBA of policies that deal with life-saving measures.

Chapter 4

Valuation of road risk reductions and VSL in practice

4.1 *Empirical estimation of VSL*

As mentioned in the previous chapter, although both revealed and stated preference techniques can be used to estimate individuals' willingness to pay reduction of road traffic risks, the most common approach is now based on the analysis of stated preferences using, in particular, the CVM, which we cover in detail in Chapter 5. Indeed, many studies have elicited individuals' preferences for transport safety and in particular road safety has been the subject of many studies

focusing on the valuation of risk reductions.

As it is the case with the analysis of the VSL in general, it should be noted that studies that value risk reductions should not be expected to obtain a constant VSL. The value of a risk reduction may depend on the type of risk (health risk, accident risk, etc.). Additionally, the way in which people trade risk and money varies across individuals and also over time for given individuals as their age and as their economic circumstances change (Viscusi, 2010). The heterogeneity of VSL has become more prominent both in terms of economics research and risk policy. However, given the range of jurisdictions and the range of contexts in which risk reductions have been valued, the resulting sizes of the estimated VSL differ substantially. In general, empirical studies of the valuation of risk reductions have found estimates of VSL from around $100,000 to close to $40 million,[1] while Hammitt (2000a) suggested that most researchers would consider the range $3-7 million reasonable. De Blaeij et al. (2003); Wijnen et al. (2009); Andersson and Treich (2011); Lindhjem et al. (2011) contain surveys of previous studies and meta-analyses that help make sense, to some extent, of the disparities found in terms of the magnitude of the VSL. It should be stressed again that, since the VSL is supposed to reflect the WTP of a given population for a reduction in its death risk, values should be expected to differ across jurisdictions. This is because that WTP is associated with individuals' perceptions of the risk, individuals' attitudes, and individuals' preferences and there are "no a priori grounds for supposing these preferences, perceptions, and

---

[1] These indicative values (in USD 2005) are found in the survey by Andersson and Treich (2011).

66

attitudes need necessarily be the same" (Jones-Lee and Loomes, 1995, p. 184) across different populations.[2] For this reason, we provide in Section 13 a series of estimates of the VSL that apply more specifically in the Canadian context.[3]

However, it has been noted that the estimated VSL values are often more sensitive to the context, the affected population, the survey design, and other aspects of the study than what the theory would predict (Andersson and Treich, 2011). There is some quite concerning evidence, for example, of the effects of insensitivity to scope on VSL estimates (see Section 4.2.9), and also some examples of undue effects of presenting the payment scenario in terms of a public or a private good (e. g. Hultkrantz et al., 2006).[4]

4.2 What affects the size of the VSL

Analysing the factors that affect the magnitude of the VSL in empirical studies is important not only because it is important for policy decision making to have an idea of how VSL values differ among socio-economic or demographic groups

---

[2] Sjöberg (2000) found that the more control an individual feels they have over a risk the more risk denial that occurs. Risk denial is the tendency to believe one's risk is below the average and, since when driving one has considerable control over the situation, risk denial is an explanation for why most people perceive their own risks as being lower than the average. These findings were also supported by Matthews and Moran (1986) and Glendon et al. (1996). These studies confirmed the widely (and erroneously) held belief that most people think themselves to be better-than-average drivers, known as self-appraisal bias. Furthermore, Sjöberg (2004) identifies the two major determinants of risk perception as being "dread and novelty" of the risks. Dread tends to be associated with risks such as nuclear disasters and illnesses such as cancer, thus we can infer that traffic risk would be a low dread risk. Novelty is related to how familiar a risk is to an individual, a novel risk being health effects from nanoparticles. Moose are certainly a familiar risk to anyone who has lived in Newfoundland for any extended period of time, and we can conclude that this would also be a non-novel risk. See Chapter 8 for more details on issues of risk perception.

[3] See also Zhang et al. (2004)

[4] In their meta-analyses of VSL in the context of road safety, Miller (2000) and de Blaeij et al. (2003) also found that estimates from stated-preference studies were significantly higher than estimates from revealed-preference studies.

but also because it also plays an important role in validity testing. In this section we briefly consider which main factors have been shown to affect the size of the VSL.

4.2.1 Wealth/income

The size of the VSL is expected to increase in the individuals' wealth (as explained in Section 3.1.1). Indeed, although a significant relationship between income or wealth and VSL is not always found, most studies do find a positive effect of income on VSL (Miller, 2000; De Blaeij et al., 2003). When it comes to the degree of sensitivity of the WTP for risk reductions to income values, most studies find comparable results (Dolan et al., 2008), with the income elasticity of the VSL normally found between zero and one (Jones-Lee et al., 1985; Persson et al., 2001). As it is usually the case in contingent valuation, though, we also find some studies in which the effect of income is not found statistically significant (e.g. Andersson et al., 2013).[5]

4.2.2 Baseline risk

Also shown in Section 3.1.1 is that the VSL is predicted to increase with the size of the baseline risk. However, in this case, the empirical evidence is a bit less conclusive. Some studies find the relationship positive (Persson et al.,

---

[5]Evans and Smith (2010) explored the theoretical underpinnings of the income elasticity of VSL after works such as Eeckhoudt and Hammitt (2001) and by Kaplow (2005) used simple models to link the values of income elasticity of the VSL to the individuals coefficient of relative risk aversion. Theoretical inconsistencies have been found among some of these studies and the discrepancies between empirical estimates of income-elasticity and the values implied by theories linked to the coefficient of relative risk aversion were reconciled by the results in Kniesner et al. (2010).

2001; De Blaeij et al., 2003; Andersson, 2008) but some others find it negative (Andersson, 2007). Andersson and Treich (2011) point out that a reason why the baseline risk does not always affect VSL in the predicted way could be due to the difference between the perceived risk (on which the individual bases her decision) and the objective risk (the risk observed by the analyst).

4.2.3 Background risk

The VSL may be affected not only by the specific baseline risk but also by existing background risks, that is, the risks of adverse events more generally (Eeckhoudt and Hammitt, 2001). If the background risk of death is large, one's WTP for a risk reduction may be lower, because of the overall low probability of survival. However, the effect of an increase in both baseline risks and background risks is uncertain.[6] The effect of a physical background risk depends on how individuals relate it to the specific risk being valued. Andersson (2007) suggest that the risks are perceived to be independent, i.e. VSL decreases with the background risk. The effect of background risk was also found non-significant by Andersson (2008).

4.2.4 Age

The VSL is sometimes converted to the value per statistical life-year, which relies on the assumption that the VSL decreases with age (Hammitt, 2007).[7] It

---

[6] There are also some discrepancies when it comes to the theoretically expected effect of background risks (Eeckhoudt and Hammitt, 2001; Evans and Smith, 2006).

[7] It has been suggested (Sunstein, 2004a) that the value of a statistical life year (VSLY) rather than the VSL should be used in cost-benefit analyses, an issue that is beyond the scope

is indeed often believed that VSL should decline with age (e.g., European Commission, 2000), sometimes leading to so-called "senior discounts" that attach a lower VSL to older individuals during policy evaluation (Krupnick, 2007). However, there is no theoretical support for this belief and the expected theoretical relationship between age and VSL is indeterminate, since the relationship is determined by the optimal consumption path which depends on assumptions on discount factors, saving opportunities, etc. (Johansson, 2002; Hammitt, 2002).[8]

When it comes to the empirical evidence, the findings in most studies support that VSL follows an inverted U-shape (Jones-Lee et al., 1985;[9] Persson et al., 2001; Krupnick, 2007). Krupnick (2007) also suggests that a quadratic specification should always be tested, since conceptual models show WTP declining with age at an increasing rate. Other studies however find that VSL declines with age (Corso et al., 2001a), or is independent of age (Andersson, 2007; Johannesson et al., 1996; Andersson et al., 2013a).

As pointed out by Krupnick (2007), age is often just a proxy for many other variables. It is, for example, quite closely correlated with baseline death risk. It is likely to also be correlated with health status and it related, although perhaps in nonlinear ways, to income or wealth. These correlations limit the ability to make inferences about its independent effect on WTP for risk reductions and thus VSL and some forms of introducing the age variable in the WTP model

of our current study.

[8] Although some theoretical and some numerical results suggest an inverted-U relationship peaking around age 40 between VSL and age (Shepard and Zeckhauser, 1984; Aldy and Viscusi, 2007).

[9] Jones-Lee et al. (1985) found a fairly flat, hump-shaped relationship between VSL and age, peaking at about age 40.

may result in lower correlation than other forms.

Another, less often considered issue, is how the WTP for a risk reduction changes depending not on the age of the respondent but on the age of the beneficiary. The relative value of reducing the risk of death for old beneficiaries is found lower than for younger ones (Johansson-Stenman and Martinsson, 2008).

4.2.5 Gender

Often, valuation studies of traffic risk reductions find that males are willing to pay less for a risk reduction than women. For example, Andersson et al. (2013) find that women are willing to pay more than men for enhanced car safety in Sweden, in contrast to other Swedish studies that, somehow exceptionally, found no statistically significant relationship between gender and WTP to reduce transport related mortality risk (Johannesson et al., 1996; Hultkrantz et al., 2006; Andersson, 2007).[10]

This type of result would fall within a more general effect observed when other perception of risks are considered (Savage, 1993; Davidson and Freudenburg, 1996; Finucane et al., 2000). Gender and race have been shown to combine in such a way that white males tend to perceive risk lower than other sociodemographic groups (the so-called *white male effect*), as explained in further detail in Section 8.5.[11] Andersson (2011) is a recent example.

---

[10] Johannesson et al. (1996) found, however, WTP statistically significantly higher among females in the case of a risk reduction delivered through a public good.

[11] A recent contribution suggests that the strength of the white male effect would depend on the jurisdiction. This is because differences in the relative socioeconomic status of women and men of different ethnicities might matter more than other underlying (e.g. biological) causes (Olofsson and Rashid, 2011).

4.2.6 Education

Several studies (Nielsen et al., 2012; Andersson et al., 2013a) find a positive and significant relationship between education and WTP for car safety in Sweden. However, this variable has also been found to exert a negative effect on willingness to pay (Johannesson et al., 1996). This might be explained because education levels are inversely related with the perception of the level of risk (Andersson and Lundborg, 2007), since more educated respondents tend to perceive risks more accurately[12] (Hakes and Viscusi, 2004) rather than exaggerating them, which is the most common case.

4.2.7 Health status

Respondents with better (self-perceived) health are expected to have a higher WTP for death risk reductions, since they have more to lose if they die. On the other hand, health may also affect the marginal utility of income, which may potentially have some offsetting effects on that WTP (Hammitt, 2002; Strand, 2006). Moreover, through the dead anyway effect, health is theoretically expected to affect negatively the VSL through its positive effect on general survival probability. Additionally, it affects positively the VSL through its positive effect on the expected future flow of incomes and its negative effect on expected health care expenditures (the wealth effect). Therefore, from a theoretical point of view, it is difficult to determine the effect of health status on the VSL.

---

[12] As expected, more educated respondents also show more responsiveness to scope (Nielsen and Kjær, 2011).

The empirical evidence generally suggests that the VSL is independent of health status. For example, Alberini et al. (2004, 2006) found no effect on the WTP to reduce a general risk of dying, with a similar type of evidence found in terms of traffic risk reductions (Andersson, 2007; Andersson et al., 2013a).

Our survey instrument included a simple question about self-perceived health status, based on an index that could take continuous values from 1 to 100, similar to the one used by Andersson (2007).

4.2.8   Altruism (private versus public policies)

When it comes to the issue of altruism (that is, how individuals WTP for others' safety affects overall estimates of the value of road traffic risk reductions), the empirical evidence appears rather inconclusive. Some of the evidence implies that individuals are safety-paternalistic (Andersson and Lindberg, 2009). On the other hand, individuals have also been found not willing to pay as much for others' safety as they are for their own safety, since they are not prepared to pay as much for a public risk reduction policy as for a private one (Johannesson et al., 1996; De Blaeij et al., 2003; Hultkrantz et al., 2006). Crucially, it is rather difficult, as pointed out by Alberini and Ščasný (2013), in practice to distinguish between neutral altruism, which should not be included in cost-benefit analyses of safety-improving policies (Bergstrom, 1982; Milgrom, 1993) and safety-focused altruism, which adds to the uncertainties involved in CV studies (Krupnick, 2007).

Usually, CV studies of road traffic risk reductions define the good valued in

terms of a private defensive device or in terms of a public policy (for instance, seat belts or airbags, in the former case, or road improvements or signalling, in the latter). Economic theory suggests that there should be a difference between both types of goods, with the VSLs based on the valuation of private goods expected to be higher, because of the *free-rider* problem that affects the revelation of preferences for public goods (De Blaeij et al., 2003; Dekker et al., 2011). On the other hand, if no strategic behaviour is present, one should expect the WTP safety improvements that strictly apply to the individual (be it the driver, the passenger in a vehicle, a cyclist, or a pedestrian) to be valued less than equivalent safety improvements that cover the individual plus others (including family members and friends). Findings otherwise suggest that individuals are misanthropists, not altruists (Brady, 2008).

Brady (2008) suggests that flawed survey design appears to be behind the, in their view, counterintuitive results, and also show that when sources of bias are eliminated through improved survey design, more intuitive results arise. Evidence from the field of health risk studies supports the existence of altruism, for example, when the potential for strategic bias is explicitly addressed by including in the payment scenario a "provision rule stressing that the preventive campaign would be carried out only if everybody would agree to pay for the policy" (Araña and León, 2002). In our study we include a control for this type of provision rule reminder, based on a majority-based referendum. There are instances when the use of the provision point mechanism is not enough to find a higher WTP for the public good than for the private good, though (Hultkrantz

et al., 2006). Svensson and Vredin Johansson (2010) also show that the WTP for a private risk reduction is higher than its equivalent for a public risk reduction with a significant part of the difference due to respondents' attitudes towards privately and publicly provided goods in general.[13]

However, more recent works have found that when an effort is made to explicitly differentiate between risk reduction policies that affect only the respondent from those who also benefit others a larger WTP is found for the latter (Adamowicz et al., 2011; Whitehead et al., 2012; Zhang et al., 2013). In another recent study using Italian data partly related to reductions in traffic risks, Alberini and Ščasný (2013) explicitly considered differences in WTP between a privately provided and a publicly provided risk reduction, presenting them together with a reminder that the public policy would also benefit other individuals besides the respondent, whereas the respondent would be the sole beneficiary of the risk reduction under private action. Alberini and Ščasný (2013) acknowledge again that they could not identify the type of altruism behind the higher WTP they elicited for he risk reduction provided through the public policy, as compared with the private one.

4.2.9   Magnitude (*scope* or *scale*) of the risk reduction

When considering the validity of the estimates of VSL obtained through a contingent valuation survey, a recurring issue has to do with the insensitivity of

---

[13]Unfortunately, we have no information in our dataset about attitudes towards public good provision. Otherwise, we would have tried to analyse if attitudes of mistrust with the effectiveness of public measures of risk reduction affect the differences in WTP between the two types of policies.

WTP to the size of the risk reduction. This problem, which (as explained in Section 6.5) often affects other types of contingent valuation studies, although usually to a lesser extent or less often, is know as *insensitivity to scope* or *insensitivity to scale*.[14]

The standard theoretical model suggests that a necessary (although not sufficient) condition for validity is that the estimated values of small mortality risk reductions be near-proportional (increasing and slightly concave) to the size of the reductions (Hammitt, 2000a).[15] A worrying consequence of having a study in which the assumption of near-proportionality is rejected is that the VSL becomes sensitive to the proposed reduction in mortality risk (Andersson, 2007). That is, the VSL changes because of the theoretically undue influence of the scope of the risk reduction policy.

The notion that the size of the WTP should increase with the size of the propose risk reduction represents a weaker condition that the notion that the WTP should increase nearly proportionally with the size of the risk reduction. For example, Corso et al. (2001) thus distinguish between *weak* and *strong* scale sensitivity. Although there is usually support for weak sensitivity in the received empirical literature on VSL valuation, the assumption of strong sensitivity is often rejected (Hammitt and Graham, 1999). However, several recent

---

[14]The terms *scale* and *scope* are used interchangeably in the literature to define the size of the good, with *scale* being more common in the literature on risk reductions (Andersson and Svensson, 2013). In discrete choice models, *scale* can, unfortunately also refer to the spread of the latent variable underlying the model (Yatchew and Griliches, 1985; Alvarez and Brehm, 1995; Allison, 1999; Carlsson and Johansson-Stenman, 2010; Williams, 2009; Mood, 2010). We will only be carefully specific when using *scope* or *scale* in a sense different from the size of risk reduction. Otherwise, we will use them interchangeably.

[15]This means that the estimated VSL should be insensitive to small changes in baseline risk (Hammitt, 2000b).

works have found some relatively promising results when it comes to meeting the assumption of strong scope sensitivity after considering enhanced ways to describe the size of the risk reduction to the respondents and/or exploiting information about the the differences in cognitive skills of the respondents (Corso et al., 2001; Alberini et al., 2004; Andersson and Svensson, 2008).

In Corso et al. (2001) and Alberini et al. (2004), for example, using proper visual aids and training the respondents in the notion of trading at the margin their wealth for their safety yields a degree of scale sensitivity that matches theoretical predictions. This same type of aid to the understanding of risk measures and changes was less successful in other studies (Jones-Lee et al., 1985; Persson et al., 2001; Andersson, 2007; Andersson et al., 2013a). More recently, Andersson and Svensson (2008) examine the correlation between cognitive ability and scope sensitivity. They find that those respondents with better cognitive skills, likely because they can understand better the changes in small probabilities involved in the survey's payment scenario, are more scope-sensitive than the rest.

4.2.10 Other factors

Andersson et al. (2013) examine the effect of time framing on the estimates of WTP for car safety and find that the WTP per unit risk reduction depends on the time period over which respondents are supposed to pay. They compare an annual and a monthly scenario and, although their theoretical model predicts the effect from the time framing to be negligible, their empirical estimates from

the annual scenario are about 70% higher than the in estimates from the monthly scenario.

The stated WTP for road risk reductions can also be affected by other socio-economic characteristics of individuals to the extent that they reflect diverse preferences about risk (Chestnut and De Civita, 2009). However, the considering the heterogeneity of preferences with respect to risk reductions is complex. Easily observable individual characteristics (gender, education, race, marital status, etc.), rarely explain much of the differences observed in responses to WTP questions when income is held constant. The differences seem to reflect more a difference in attitude that is difficult to define (Chestnut and De Civita, 2009).

Part IV

Methodology

Chapter 5

Non-market valuation
methods

The valuation of the provision of public goods and services is more often than not impossible through the observation of transactions in real markets, since these rarely develop in the cases of nonexcludable goods or services. Several methods have therefore been developed to estimate the benefits derived from public goods and services. Two main categories can be considered: stated preference methods and revealed preference methods. Stated preference methods are based on data on individual preferences obtained by asking individuals directly about their preferences. A sample of the relevant population is contacted through some type of survey that describes and proposes a hypothetical market for the non-market good or service. The most often used stated preference technique is the

CVM. In this study, we use the CVM to estimate the benefits of policies based on decreasing the risk of moose-vehicle collisions.

Another main branch of environmental valuation tools involves the use of revealed preference techniques. These include the dose-response method (Barbier, 1994), which first quantifies the physical effects on an economic activity of a change in the ecosystem good or service and then values the effect on market activity brought about by the change; the perhaps more popular travel cost method, which estimates values based on the travel expenditures and time costs an individual incurs while visiting a recreational site (Parsons, 2003, p.269); and the hedonic valuation method, which involves observing monetary trade-offs made with respect to changes in the characteristics of a good (Taylor, 2003, p. 331).

In the risk valuation literature, revealed-preference studies most commonly exploit the notion of hedonic wages to value risk reductions by analysing compensating wage differentials, that is, by estimating the trade-off between wages and job-related risks. Examples and reviews of studies that follow this approach can be found in, for example, Gunderson and Hyatt (2001); Mrozek and Taylor (2002); De Blaeij et al. (2003); Aldy and Viscusi (2007); Bellavance et al. (2009); Kniesner et al. (2010); Cropper et al. (2011). Although less frequently, analyses of individual consumption decisions and residential property values have been used as well (Atkinson and Halvorsen, 1990; Dreyfus and Viscusi95; Viscusi and Aldy, 2003; Andersson, 2008).[1] However, hedonic wage studies could help

---

[1]As Cropper et al. (2011) point out, these studies are often seen as less suitable for

value road safety when dealing with individuals for whom a traffic-accident risk represents a work-accident risk (truckers, sales representatives, etc.) but not if the value placed by the average citizen/driver is of interest (Ludwig and Cook, 2001).

Additionally, and although revealed preference studies are often seen as more credible than stated-preference studies, the former have a crucial limitation compared to stated preferences methods: they can only be used to estimate *use* values (those that involve direct interaction with the resource), while stated preference methods can help elicit both use and non-use values, such as existence values (Krutilla, 1967), and also the values of goods and services that are only hypothetical. For example, revealed preference techniques cannot "estimate values for levels of quality that have not been experienced" (Boyle, 2003, p. 266), while stated preference methods can be used to elicit the willingness to pay for a good that is not yet available in real markets.

Studies based on the analysis of hedonic wages can be of limited use when valuing mortality risks associated with illnesses and nonoccupational exposures rather than with work injuries. They also fail to elicit the preferences of those who are not part of the labour force (particularly those who are under age and those who are retired, two collectives who are often among the most affected by risk reduction policies). Furthermore, compensating wage differentials must be inferred statistically, rather than being observed directly and the estimates

---

valuation than hedonic wage or stated-preference studies, because of individuals' difficulties in estimating actual or perceived risks, the need to make assumptions about key factors such as time costs (in some product studies), whether illnesses are likely to be fatal (in some hedonic property value studies), and other factors.

of the VSL based on hedonic wage equations assume that the measure of job risk used by the researcher matches workers' risk perceptions (Cropper et al., 2011). Also, since jobs are not allocated randomly, there might be a downward bias affecting the estimates of the value of risk reductions, because the workers who accept the most hazardous jobs are likely to be those who are the most risk-averse and thus demand the least compensation, that is, those workers who have the smallest VSLs (Cropper et al., 2011).[2]

On the other hand, stated-preference methods have also their shortcomings. For example, a *hypothetical bias* arises when individuals declare to be more willing to spend their money when asked inconsequential survey questions than when they answer consequential questions about it, so survey-based valuation studies tend to bias willingness to pay upwards.[3] Stated preference techniques are also susceptible to a number of other types of potential biases (Mitchell and Carson, 1989). Although it is impossible to avoid all possible biases, the literature dealing with non-market valuation suggests ways to identify and try and mitigate their effects. Chapter 6 describes how the empirical literature has dealt with the main issues surrounding contingent valuation studies.

Key advantages of the stated-preference approach in the context of valuations of risk reductions include the fact that they can target the general population, the fact that the scenario presented to respondents can be tailored to

---

[2]A study based on a Canadian sample (Lanoie et al., 1995) used both a stated-preference and a wage-risk analysis approach to find the WTP from the stated-preference study for reducing the risk of a motor vehicle accident fatality to be about an order of magnitude smaller than the WTP for reducing the risk of a fatal on-the-job accident, using stated preference and wage-risk analyses.

[3]List and Gallet (2001) and Murphy et al. (2005) are examples of meta-analyses of the problem of hypothetical bias in stated preference studies.

specific risks, which can even include risks that are purely hypothetical, and the fact that the scenario can provide detailed information about risk latency. These advantages have made stated-preference studies increasingly common in the risk valuation literature. Cropper et al. (2011) include a recent survey of this literature.

In spite of the issues that no doubt affect the method when applied to this type of task (e.g. Beattie et al., 1998), most stated-preference studies either adopt the contingent valuation (CV) format (Jones-Lee et al., 1995; Corso et al., 2001a; Carlsson et al., 2004; Alberini, 2005; Hultkrantz et al., 2006; Andersson, 2007; Andersson and Svensson, 2008; Leiter and Pruckner, 2009; Svensson and Vredin Johansson, 2010; Araña and León, 2012; Nielsen et al., 2012; Andersson et al., 2013a) or use choice experiments (Johansson-Stenman and Martinsson, 2008; Hensher et al., 2009; Cameron et al., 2010; Carlsson et al., 2010; Andersson et al., 2013b; Cameron and DeShazo, 2013; Veisten et al., 2013). In both cases, respondents are presented with a hypothetical scenario (often involving a device or product or a public policy program) that would decrease her mortality (and/or morbidity) risk. The scenario normally includes information about the respondent's or the general baseline risk (without the program), the size of the risk reduction that would be delivered by the program, the time period over which it will be delivered, the amount and form of payment, and other relevant information (Cropper et al., 2011). Respondents are asked to choose, in the case of CV, between enjoying the program and paying the stated cost (the *bid*) and the *status quo*. In a choice experiment, respondents choose among several

84

programs.

In our study, we use the CVM to elicit the willingness to pay for reductions in the risk of suffering a moose-vehicle collision. The next sections describe the methodology in general, with particular attention to the analysis of the type of question format we adopted, namely the double-bounded dichotomous-choice question.

Chapter 6

The contingent valuation method

The contingent valuation method (CVM) elicits willingness to pay most often using a close-ended, dichotomous choice, hypothetical market-type questions included in a direct survey, which can be administered via telephone, mail, or in person (Kanninen, 1993). When the dichotomous choice CVM is used, respondents are asked whether or not they would be willing to pay a particular amount (usually referred to as *the bid*) for a particular good or service in a hypothetical market.[1] Each individual is proposed a different value (allocated randomly) of the bid. The respondent must respond to the payment question choosing among a 'yes', a 'no', and usually some form of 'don't know/no answer'

---

[1] Often the payment scenario is phrased in terms of eliciting support for a public policy that will deliver a certain change in quality or quantity or a non-marketed public good.

options. The proportions of these responses can be then tracked in relationship with the size of the "bid" amounts offered to each individual.

The most commonly used bidding methods are the single-bounded and, to a lesser extent, the double-bounded dichotomous choice formats, which we adopt in this study and describe in detail in Chapter 7. The single-bounded model approach recovers the bid amount as a threshold by asking only one dichotomous choice question. The statistical efficiency of this approach can be improved by use of the double-bounded model, which engages in two bids by asking each respondent two dichotomous choice questions (Hanemann et al., 1991). Further generalizations (for example, based on further iterative bidding or a payment card method) are much less frequent.[2]

As a *stated preference* (as opposed to a *revealed preference*) method, the CVM can estimate the value of non-market public goods that generate non-use values (Mitchell and Carson, 1989; Bateman et al., 2002). The CVM is especially useful in those cases when some aspect of the good or service concerned does not currently have a market expression.

Some key issues in contingent valuation affecting our study and the ways in which they have been addressed in the valuation literature are considered below.

---

[2]Multiple bounded discrete-choice formats (Welsh and Poe, 1998), payment card formats, and other iterative designs beyond two questions are much less common but see Bateman et al. (1995); Loomis and Ekstrand (1997); Bateman et al. (2001); Cameron et al. (2002); Whitehead (2002); Alberini et al. (2003); Araña and Leon (2007).

6.1 Payment vehicle and elicitation format

The design of the survey instrument for use in a CVM study involves the adequate choice of the payment vehicle used within the hypothetical market scenario. Indeed the payment scenario should be as realistic and uncontroversial as possible. Therefore, if the payment mechanism is unrealistic or objectionable, some respondents may reject the whole valuation exercise, even if they value the environmental change, thus generating *protest responses* (Boyle, 2003),[3] or adjust their bids in such a way that they no longer reflect the respondent's underlying willingness to pay for the good (Morrison et al., 2000).

Willingness to pay can vary depending on how the payment is supposed to be collected or when, for how long, or how often, the payment is due (Stevens et al., 1997; Morrison et al., 2000). Mostly, CVM practitioners aim at choosing a payment vehicle as close as possible to the one that would be used if the actual policy came into effect, in trying to maintain a balance between realism and rejection of payment vehicle when designing the payment scenario (Mitchell and Carson, 1989). Oftentimes, CVM surveys propose an increase in taxes or voluntary donations to an agency or private entity entrusted with carrying out the proposed policy. However, depending on the type of payment scenario, fees and (for the case of goods already of privately provided or that could be potentially provided by the market) prices are also an option.

The undue sensitivity of the WTP estimate to the choice of payment vehicle leads to *payment vehicle bias* (Mitchell and Carson, 1989; Boardman et al.,

---

[3]We describe the issue of dealing with protest responses more fully in Section 11.1.

2001). Mitchell and Carson (1989) suggest that WTP is significantly affected by this choice. For example, one group of respondents could be assigned a tax as the payment vehicle, while the other group would be offered a utility bill increase. If the mean WTP estimates for the good show not to be statistically different, one can conclude that payment vehicle does not bias the estimate.

We use what is known as the *referendum format,* operationalized in our case by proposing an increase in either federal annual income taxes or on provincial vehicle-licensing fees for the next five years as the payment vehicle for one version of our questionnaire that involves a publicly provided good. The increase in annual income taxes is associated to a payment scenario that involved the Canadian federal government's funding of the risk reduction policy. A surcharge on the driver's licence fee is the payment suggested for the version that proposes a provincially funded policy instead.[4] For the version that deals with the private good, we use an annual rental fee for the hypothetical personal risk reduction device.

The main reason why we chose these payment vehicles is that they are likely the most plausible choice and because, in the Canadian context, coercive payments have proven in some of our earlier studies to be more robust than a voluntary donation (Lyssenko and Martínez-Espiñeira, 2012b). In the former case, it also made it possible to design a scenario based on a referendum, which has been very commonly used in previous contingent valuation studies.

---

[4]The choice of the agency in charge of the policy is another key issue in the design of the payment scenario, so we wanted to analyse the potential effects of this choice in the case of the public good version of the policy.

The payment schedule proposed for the public good should also be carefully considered. Stevens et al. (1997) examined the sensitivity of respondents to the payment schedules in their survey scenario, namely lump-sum and periodic payments. They noted that the insensitivity of the respondents to the payment schedule would imply the presence of a temporal form of embedding effect, related to the notion of insensitivity to scope.[5] This would threat the credibility of the estimate, since economic theory predicts that individuals should distinguish between payment schedules. Stevens et al. (1997) applied a split-sample approach, finding confirmation of that theoretical statement, although the implicit discount rates used by those respondents facing a payment series were very high.

In our case, two payment schedules are proposed to the respondent: a one-year rental cost for a private device in the case of the private good and an annual increase in either annual income federal taxes or provincial vehicle licensing fees during a five year period in the case of the public good version of the policy. Therefore, although the bid values that represent the proposed contribution to the provision of the good, both measured on a per year basis, are comparable, we have the opportunity to consider two payment schedules in addition to analysing the payment effects across three payment vehicles using a split-sample approach.

---

[5] See Section 6.5 on scope effects, where the related issue of insensitivity to scope effects is covered.

6.2 Respondent uncertainty and hypothetical bias

Most CVM studies implicitly assume that respondents respond to the valuation questions with full certainty. However, several solutions have been proposed to address and exploit potential respondent uncertainty at both a theoretical and empirical level (Shaikh et al., 2007; Blomquist et al., 2009; Moore et al., 2010; Loomis, 2011; Martínez-Espiñeira and Lyssenko, 2012).

Some empirical evidence (Champ et al., 1997; Johannesson et al., 1998; Johannesson et al., 1999; Blumenschein et al., 1998; Blumenschein et al., 2001; Ethier et al., 2000; Champ and Bishop, 2001; Vossler et al., 2003; Akter et al., 2008) suggests that information about response uncertainty can be used to ameliorate issues caused by hypothetical bias. This is because those who state lower levels of response certainty tend to exhibit most of the hypothetical bias, so recalibration or recoding of responses according to certainty levels can reduce the bias. The effectiveness of these calibration techniques seems to be commodity-, context-, and even individual-specific (Whitehead and Cherry, 2007).[6]

Two approaches, one based on exploiting information from a quantitative scale of respondent certainty and the other based on a qualitative scale (Whitehead and Cherry, 2007) have been followed when dealing with respondent uncertainty. Most often, follow-up questions about uncertainty are included as a debriefing question within the survey instrument. Examples of this strategy in-

---

[6]See also Norwood et al. (2008) for a critical analysis of several experimental analyses.

clude Li and Mattsson (1995), Champ et al. (1997), Blumenschein et al. (1998), Johannesson et al. (1998), Loomis and Ekstrand (1998), Ekstrand and Loomis (1998), van Kooten et al. (2001), Lyssenko and Martínez-Espiñeira (2012b); and Trenholm et al. (2013). On the other hand, instead of using a conventional dichotomous question, authors such as Ready et al. (1995), Wang (1997), Welsh and Poe (1998), Ready et al. (2001), Alberini et al. (2003), and Wand and He (2011) embedded information about respondent uncertainty directly in the response options offered to the respondents.

Two recent surveys of the literature by Shaikh et al. (2007) and Akter et al. (2008) and several recent empirical contributions by Chang et al. (2007), Groothuis et al. (2007), Broberg and Brannlund (2008); Hanley and Kriström (2008); Akter et al. (2009), Hung (2009), Blomquist et al. (2009), Champ et al. (2009), Li et al. (2009), Moore et al. (2010), Lyssenko and Martínez-Espiñeira (2012b), and Martínez-Espiñeira and Lyssenko (2012) confirm that the issue of accounting for respondent uncertainty is still preoccupying applied researchers in the area of contingent valuation.

Our survey instrument included a follow-up question whose answers were used to construct a numerical certainty scale with which to adjust the original responses obtained. Therefore, we used the first approach to dealing with uncertainty, obtaining a numerical certainty scale taking values from 1 to 10 from the follow-up question. As suggested by this section, there are different ways in which this type of information can be used, including using the numerical certainty scale to model heteroskedasticity in responses.

6.3 Treatment of 'don't know' responses

Another issue affecting the practice of CVM has to do with the recommendation by the U.S. NOAA Panel (Arrow et al., 1993) that referendum format[7] surveys offer respondents a "don't know/not sure/would not vote" option, since the Panel failed to provide sufficient guidance as how the 'don't know' responses should be interpreted empirically.

Indeed, the issue of 'don't know' responses in contingent valuation studies has been the subject of much research (e.g. Wang, 1997; Carson et al., 1998; Haener and Adamowicz, 1998; Groothuis and Whitehead, 2002; Alberini, et al., 2003, Balcombe and Fraser, 2009). However, as Groothuis and Whitehead (2002) point out, there is no clear answer in the literature yet about how the 'don't know' responses should be treated empirically or what kind of information they provide. They argue that often there is no information available for the analyst to justify the reassignment of 'don't know' responses into the 'yes' and 'no' categories. Of course, if such information were available, the recoding of 'don't know' responses would make it possible to obtain more precise estimates of willingness to pay and the coefficients of the variables affecting it. Although several existing studies provide a variety of recommendations, they are usually based on rules of thumb valid for individual studies and *ad hoc* suggestions after the analysis of a particular sample. How 'don't know' responses should be

---

[7]There seems to be some confusion (Champ et al., 2002) in the valuation literature about the term *referendum*. Some use it to refer to the dichotomous-choice format in general. Strictly, the term should be reserved to the posing of the payment question as a vote on a referendum with a dichotomous-choice response format (e.g., vote in favor/vote against). For our purposes, this distinction is not relevant, so we use the terms *referendum format* and *dichotomous-choice response format* interchangeably.

treated in empirical analyses is a question that remains open.

The simplest way to deal with 'don't know' responses is to just drop them from the sample. This strategy, though, implicitly assumes that the indecisive respondents are not significantly different in terms of their relevant socioeconomic and personal characteristics from those respondents that remain in the sample. Furthermore, this approach leads to the reduction of the sample size and may cause sample selection bias if the indecisive respondents are, instead, systematically different from those who choose either the 'yes' or 'no' options (Wang, 1997). Also, as Wang states, one cannot justify theoretically the elimination of 'don't know' responses. Wang (1997) observes that a respondent may state a true 'don't know' response, which Wang (1997) sees as simply different from a 'no' response, so recoding that 'don't know' as a 'no' response does not seem to be a robust approach. In fact, Wang (1997) was the first to argue that 'don't know' responses should be treated as a unique category of votes.[8]

The double-bounded payment questions in our survey offered a 'don't know' option at each stage. However, we did not find a substantial proportion of cases in which a 'don't know' response was offered to all the bids proposed. Moreover, we instructed the survey interviewers to restart the iterative double-bounded process using either a higher bid or a lower bid after a 'don't know' response to the initial bid, so we simply could make use without any adjustment of most of

---

[8]In particular, Wang (1997) suggested a utility threshold approach and proposed a maximum likelihood procedure to treat the 'don't know' responses in his contingent valuation study. According to this approach, individuals can make a choice only if a threshold is passed. This approach focuses only on non-strategic 'don't know' responses; in other words, on respondents who are uncertain about their preferences at the moment of the survey effort.

the cases in which the first response to the initial bid was inconclusive. Further details about how we dealt with the cases in which a 'don't know' was given at any stage of the WTP elicitation process are presented in Section 11.1.

6.4 Protest responses

Some respondents to 'contingent valuation surveys state a null[9] value of WTP or willingness to accept (open-ended format) or provide a 'no' response (dichotomous choice format) as an answer to the payment question (e. g. Dziegielewska and Mendelsohn, 2007). Answering 'no' to the proposed bid in a dichotomous-choice question, may mean that one's WTP is less than the bid (Halstead et al., 1992) or a lack of willingness to pay any amount at all. Therefore, 'no' responses can be classified either as legitimate (true) 'no' responses or as protest responses.

Genuine 'no' and zero responses indicate indifference or aversion to the policy (Strazzera et al., 2003a) or that the suggested contribution is not affordable for the respondent. In these cases the zero willingness to pay should be interpreted as the true (zero) value of the resource to the respondent (Strazzera et al., 2003a) and the welfare of these respondents is totally unaffected by the provision of the public good (Strazzera et al., 2003b). However, if a respondent states a zero willingness to pay or rejects a payment question for reasons other than lack of interest or non-affordability, such response is typically considered

---

[9]Or outrageously high.

a protest response.

According to Boyle (2003), there are at least three main reasons for protest responses. First, some respondents may not understand the survey fully but answer the questions anyway. Second, strategic behaviour may prompt respondents to refrain from revealing their true WTP with the intention to unduly influence the results of the valuation study and therefore the policy design. Third, respondents may protest against one or more components of the valuation exercise, because, for instance, they find some questions unethical or have an adverse reaction to the interview or the payment vehicle (Strazzera et al., 2003a). Jorgensen and Syme (2000) and Jones et al. (2008) found that most protest responses were related to the fact that the government was supposed to be involved in the provision of the good.

To some extent, the effect of protest responses can be mitigated by good survey design. However, even the best designed survey instrument will likely generate some protests (Calia and Strazzera, 2001; Strazzera et al., 2003a). The percentage of protesters, however, varies widely: around 5%, for example, in Boyle (1989), 33% in Halstead et al. (1992), 18% in Strazzera et al. (2003b), 31% in Cho et al. (2008), etc.

In order to identify protest responses, researchers most commonly ask a set of debriefing questions about reasons for a 'no' response. Two main general approaches can be found in the literature. The first involves asking an open-ended question about why a negative response to the payment question was given and then analyzing the verbatim answers before allocating them across

categories (if only just between the categories of protest response and non-protest response), as in Jorgensen et al. (1999). The second approach suggests a series of options and asks the respondent to choose one, a given number of them (usually three), or as many as deemed relevant (e. g. Morrison et al., 2000).

We combined both approaches by providing respondents with a set of standard reasons from which they could choose as many as they wanted and also offering them the option to choose the option 'other' and state their own reason for a negative response to the payment question. We explored the types of reasons provided for not being willing to pay any of the suggested bids and classified the negative responses as either legitimate 'no's', *clear protests*, or protest (the latter including a few more cases than the former, as described in Section 11.3).

As in most previous CVM studies, the follow-up debriefing question was asked only of those respondents who provided 'no' responses to all the dichotomous-choice question they were asked.

There is no agreement among contingent valuation practitioners regarding which explanations of the zero or 'no' responses should be classified as protest responses. Lindsey (1994) first offered a systematic view of the problem of protest votes, stressing that their identification depends on whether a market rationale or a referendum model is adopted. The market model assumes the existence of a 'true' value that is independent of the measurement process. Therefore, under the market model, protest responses are those that reflect

the undue influence of contextual elements of the valuation exercise (payment vehicle, information constraints, judgments of procedural fairness, etc.). On the other hand, the referendum model is more concerned with whether a zero response reflects intended behaviour. Since many of the protest reasons are considered to be legitimate influences upon actual behaviour (Lindsey, 1994), the referendum model is less stringent in terms of the classification of protest responses.

Apart from issues of definition and identification, there is no consensus about the treatment of protest responses either. Since protest bids do not reveal the true willingness to pay for the resource, the resulting willingness to pay estimate can be biased. Thus, it will be inappropriate to use such an estimate in CBA (Halstead et al., 1992). An inappropriate treatment of protest responses may threat the validity of willingness to pay estimate and lead to a suboptimal policy design.

Halstead et al. (1992) suggested three ways of treating protest responses. The first is to eliminate them from the sample, so protesters are implicitly assigned the mean willingness to pay. The second leaves protests as true zeros, which may lead to a downward bias of the willingness to pay estimate. The third approach treats the protest bids as missing values and assigns the most likely willingness to pay values to the protesters according to their socio-demographic characteristics. Some authors argue that the elimination of protest responses may lead to an invalid estimate of the value of the policy due to selective data removal (Strazzera et al., 2003a). Strazzera et al. (2003a) report

that a sample selection bias may arise if protesters are significantly different from non-protesters and are eliminated from the sample. To correct for the bias, the application of a sample selection model would be required, as in, for example, Calia and Strazzera (2001); Strazzera et al. (2003b); and Collins and Rosenberger (2007).

Jorgensen and Syme (2000) also argue that, if protest responses are censored off the sample, one needs to show that their likelihood is independent of the willingness to pay question format, values of exogenous variables (age, income etc.), and willingness to pay response, since deleting one type of response may result in biased estimates. For example, they found lower income individuals more likely to be protesters, so the elimination of the protesters may lead to the bias of willingness to pay estimates, since the preferences of higher income respondents would be given more weight. Gonzalez-Caban et al. (2007) analyze also the differential in the incidence of protest responses between Native Americans in the United States and the general population, suggesting that different subgroups within a given population could present remarkably different rates of protest responses.

Meyerhoff and Liebe (2006) and Jorgensen and Syme (2000) focus on protest motivations and argue that the censoring of protest responses is unjustified. Both articles argue that protest responses should not be eliminated from the sample, but considered as true zeros.

Morrison et al. (2000) suggested recoding some of the protest responses instead of deleting them. One of the debriefing questions in their study asked if

a respondent would support the project if this were not to require any payment. Those who responded positively to this question were coded as payment vehicle protesters and then asked further questions that allowed determining if they were accurately treated as supporters of the project.

In general, as suggested by the previous paragraphs, the literature agrees that there is no established procedure that has a solid theoretical basis for excluding protest responses (Boyle and Bergstrom, 1999). However, despite the ongoing attempts to examine different treatments of protest responses, it is still quite common to simply eliminate these responses from the sample (Morrison et al., 2000; Boardman et al., 2001).

In practice, the identification of protest responses seems to follow *ad-hoc* criteria and relies on substantive research judgement (Jorgensen et al., 1999). This is problematic, since two contingent valuation analyses conducted independently about the same policy using comparable population samples may lead to very different welfare estimates depending upon the identification of protest responses.

Jorgensen et al. (2001) suggested that protest responses likely arise more frequently when the scope of the public good change exceeds an individual's ideal level of service. Individuals may be willing to pay for small increases in a public good only, while they may consider large increases to be unfeasible or associated with negative consequences. In this sense, protest beliefs can also contribute to the absence of scope sensitivity (which, as described in Section 6.5 arises when respondents fail to value the part or the whole, but state willingness to pay

100

bids that reflect their general attitude toward paying for the policy instead).
These zero responses might appear an issue of individuals failing to distinguish
between different levels of the public good. However, this insensitivity to scope
may not be entirely due to an inability to appreciate the willingness to pay
question (Desvousges et al., 1993; Diamond and Hausman, 1994), the desire for
moral satisfaction (Kahneman and Knetsch, 1992) or to survey misspecification,
but to respondents' negative attitudes toward paying.

We treated *clear protest* responses in two different ways: as zeros and even-
tually eliminating them from the sample for the finally reported set of results.
The former is the most recommended approach and the one that leads to the
most conservative estimate of mean WTP. The latter, however, tended in our
case to provide more precise estimates even after accounting for the fact that
the sample size was reduced (see Section 11.3), so we reported only results based
on the elimination of *clear protest* responses. The observations we were more
in doubt about, the protest responses that were not *clear protest* responses we
left as zeros.

6.5   Scope effects

Economic theory predicts that the total willingness to pay for a good or service
should be increasing in its quantity, due to the property of *non-satiation* (or
free disposability). Individuals should be expected to value more of a good (or
a higher quality of a good) more highly than less of it (Boardman et al., 2001).

This theoretical expectation can help assess the validity of a result obtained through the use of the CVM. The implication for our valuation study is that the willingness to pay for a smaller quantity or *scope* of the good (a smaller size of the risk reduction and/or a reduction in only the risk of dying)[10] is expected to be less than its counterpart for the larger scope (a proposed policy with a larger scope of risk reduction and/or that reduces also morbidity risks).[11]

Scope insensitivity is in fact a major criticism of the CVM. It was identified by the NOAA Panel (Arrow et al. 1993) as one of the tests for a reliable CV survey. Carson et al. (1994) conducted a review of 27 contingent valuation surveys and found that all but two showed significant scope effects on WTP. He also noticed several issues with the methods used in these studies. The use of open-ended formats, the provision of information, and the lack of random sampling are a few (Hanemann, 1994). Also in one of the studies, once outliers were removed (as is standard practice), a significant scope effect appears. Whether or not these few studies that do not show scope effects are valid, most CVM studies do show significant scope effects. Another more recent review of 109 contingent valuation studies found that the majority of these passed scope tests (Desvousges et al., 2012a). A criticism from the authors on the majority of scope tests is, however, that they do not assess the adequacy of the response to scope effects.

---

[10] See Footnote 14 about the use of the terms *scope* and *scale*.

[11] Another way to refer to these two dimensions of the policy is to say that some proposed policies have a more comprehensive scope (they would reduce the risk of both dying and getting injured) than others and that some have a larger scaled than others (they offer a larger risk reduction from the baseline risk level).

If an estimated WTP function fails to pass the scope test, it may be considered suspect in terms of reliability, suggesting that respondents may not have adequately explored their preferences and budget constraints before answering the payment question or that they simply did not take the question seriously. Therefore, the inability to show empirically the scope effect (showing that the WTP increases in proportion or near proportion to the scope) remains one of the key points of the critique of the CVM (Goodstein, 2005; Heberlein et al., 2005) and its use in the risk valuation literature (Frederick and Fischoff, 1998; Hammitt and Graham, 1999; Leiter and Pruckner, 2009; Andersson et al., 2013b).

For instance, Kahneman (1986) showed that the respondents' willingness to pay to clean up all the lakes in Ontario did not significantly differ from their willingness to pay to clean up all lakes in one of the regions of that province. Desvousges et al. (1993) found that the mean estimates of saving 20,000 birds and 200,000 were not significantly different: $80 and $88 dollars, respectively, although in this case, the additional information respondents received in terms of the relative proportion of birds saved differed little among policies, which might explain the lack of sensitivity to absolute scope (Carson and Mitchell, 1993). Other contingent valuation studies that report scope insensitivity include Diamond and Hausman (1994), Schwartz (1997), Svedsäter (2000), and Shiell and Gold (2002). Only a minority of studies report that their results pass the scope test (Walsh et al., 1992; Smith and Osborne, 1996; Brouwer et al., 1999).

While initially this scope insensitivity was attributed to the survey design

and an *embedding effect*[12] (Carson, 1997), other reasons for scope insensitivity have been proposed (Hausman, 1993; Arrow and Leamer, 1994; Carson and Mitchell, 1993; Carson and Mitchell, 1995). Carson et al. (2001) conclude that failure to demonstrate a scope effect in a valuation study could be due to survey design issues, inability to detect the scope effect, or a violation of economic theory. The *embedding* and *warm glow*[13] effects are also often named as the reasons for scope insensitivity (Andreoni, 1990; Kahneman and Knetsch, 1992; Czajkowski and Hanley, 2008), suggesting that, if statements of WTP are considered by respondents as signals of their attitudes towards a public good, they exhibit low sensitivity to changes in scope.

Scope sensitivity tests can be categorised as either internal or external. An external test involves using different subsamples (which must be statistically equivalent) of respondents who are asked to place a value on different levels of a good (different scopes), while an internal test involves asking the same respondents to value the different levels of good. The latter, as expected, much more often yields results consistent with scope sensitivity (Smith and Osborne, 1996). This might be, however, simply because respondents internal test respondents may show "internal integrity" (Czajkowski and Hanley, 2008).[14] Indeed, most

---

[12] This term refers to the case when the respondents fail to recognize the difference between different quantities of the good, when one quantity is embedded into another. For example, respondents may think that the preservation of some particular species in a habitat means the preservation of all species in the habitat. Thus, the willingness to pay to preserve some species in the habitat will be close to the willingness to pay to preserve all species (Carson, 1997).

[13] Knetsch and Sinden (1984), describing the notion of a warm glow effect, argue that some respondents do not express their valuation of the resource, but rather "purchase moral satisfaction".

[14] Giraud et al. (1999) showed that, surprisingly, sometimes the internal test fails to show scope effect, while the external test (based on split samples) showed the expected scope

studies use a split-sample design, as the NOAA panel explicitly recommended, to test for scope sensitivity (Arrow et al., 1993).

Heberlein et al. (2005) went beyond the traditional split-sample approach in identifying the insensitivity to scope and tried, instead, to identify the conditions that may lead to scope insensitivity. The authors argue that the inability to pass the conventional scope test does not necessarily imply the invalidity of the estimate and that the reasons behind such inability may be consistent with economic and psychological theories. For example, respondents may treat the "part" and the "whole" as two different goods, so directly comparing the mean estimates of the "part" and the "whole" using a traditional scope test may mislead the judgment about the validity of the estimate. For instance, Boman and Bostedt (1999) found no significant change in willingness to pay according to the supply of wolves considered, suggesting that their respondents cared more about securing the wolf species as such than about the trade-off between money and wolves. Ojea and Loureiro (2008) also suggest that respondents to contingent valuation surveys who end up adopting a *citizen*'s approach[15] to the valuation exercise are likely less sensitive to scope, while those acting in a *consumer* mode are more sensitive to scope. To the extent that our exercise on the valuation of reductions in the general risk of colliding with a moose might trigger the activation of citizen-type preferences in respondents, we might have to expect some insensitivity to scope related to this issue.

---

sensitivity.

[15] For a detailed summary of the literature on the issue of distinguishing between citizen and consumer preferences in contingent valuation see Martínez-Espiñeira (2006).

Several studies used survey designs that allow testing of an incremental adding-up criterion. This test consists of assessing whether the sum of WTP for several incremental environmental goods (where each increment assumes that the previous increments is given) is equal to the WTP for all the environmental goods combined for each respondent. Diamond et al. (1993) performed this test and concluded that the responses to the survey did not vary consistently with economic preferences. A more recent review (Desvousges et al., 2012a) of several contingent valuation studies that apply a scope test found that (Chapman et al., 2009) is the only study among those that permits an adding-up test. Then, (Desvousges et al., 2012b) implemented the incremental adding up test expanding on (Chapman et al., 2009) and found a result similar to Diamond et al. (1993), that the sum of the estimated WTP for the increments being three times as great as the estimated WTP for the whole and they concluded that passing the standard scope test does not imply that the response is adequate. Hausman (2012) states that contingent valuation does not indicate stable individual preferences unless the survey can pass the Hausman-Diamond adding-up test.

The issue of sensitivity to scope is particularly relevant in the case of the valuation of risk reductions using stated preference methods. According to standard theory, WTP values for small mortality reductions, as typically valued in the literature, must be "near-proportional", increasing with the size of the risk reduction and strictly concave (Hammitt, 2000b). This would imply that respondents are expected to place just under double the value on a risk reduction

that is twice as large.This near-proportionality is referred to as the strong form of scale sensitivity.

Carson (2012) suggests the difficulty in understanding and valuing small probabilities (as evidenced in financial planning scenarios) as a possible reason for the lack of scope sensitivity in CV studies. The use of visual aids has been proven to help in abating this problem (Corso et al., 2001a), while some studies have found that the level of education influences the extent to which these aids help (Sund, 2009). Another reason could be that the expected utility model is not valid for the way individuals form valuations (Leiter and Pruckner, 2009). Yet another explanation for lack of scale sensitivity in WTP surveys is the exclusion of relevant qualities when carrying out sensitivity analysis (Heberlein et al., 2005). For example, when attitudinal factors were included in one recent study, the authors found that the results went from failing the strong scale test for near-proportionality to passing (Leiter and Pruckner, 2009). Finally, it has been noted that, in WTP-based studies, money is utilized as the scale that reflects individual's underlying utility functions but that this scale can vary across individuals because of different perceptions on what are the right or appropriate bounds for WTP. Using a correction based on the use of anchoring vignettes, Araña and León (2012) showed that the insensitivity to scale disappears once WTP responses are corrected for self-perception bias.

Finding willingness to pay estimates that are far from this theoretically expected near-proportionality to the scope of the risk reduction may lead to sometimes wildly differing estimates of the value of a statistical life (VSL) for a given

107

policy. It is then no surprise that the literature on the contingent valuation of risk reductions includes many studies that focus on this issue (Hammitt and Graham, 1999; Corso et al., 2001a; Leiter and Pruckner, 2009). Section 4.2.9 summarizes this component of the risk valuation literature.

In our survey, we constructed several variables to measure scope/scale of the policy (variable *comprehensive* itself indicating whether the proposed policy would reduce the general risk of crashing with moose in the province instead of just the risk of the respondent as a driver of the vehicle owned). The variable *diffM* measured the final absolute reduction in the mortality risk associated with MVCs, while *diffI* measured the equivalent injury (or morbidity risk). For a considerable subsample of respondents, *diffI* took the value of zero, since the policy they were asked to consider explicitly ruled out reductions in injury risks from MVCs.

Contrary to most previous studies dealing with the valuation of risk reductions, our scope variables took quasi-continuous values, since their construction was based on a combination of several discrete components. First, three randomised values for the magnitude of the proportional risk reductions[16] were proposed. Second, the baseline risk itself took one of five randomized values in the case of the public policy version of the survey,[17] while it was free to vary according to the respondents own perception (variables Q12 and Q13 for mortality and morbidity risk own perceived risk rates, respectively). Section 10.3

---

[16] The baseline risk was supposed to be either halved, divided by 3, or divided by 4, according to variable MULTI.

[17] Either 4, 6, 8, 10, or 12 in 100,000 for mortality risk (variable RM) and 30 times those values for RI, its morbidity counterpart.

describes the construction of these variables in more detail.

6.6   Analysis of the drivers of willingness to pay

We considered the potential effects of a series of covariates on the willingness to pay to reduce the risk of colliding with moose. In general, a key ingredient in CVM studies is the use of questions other than those directly related to the payment scenario and payment question. These include debriefing/follow-up questions after the main payment question,[18] questions about respondent attitudes, questions about respondent opinions about certain general policies or policy issues, and questions aimed at finding out whether the respondent is a direct or indirect user of the resource valued or similar resources. Finally, questions about the socio-demographic characteristics of the respondents are also usually asked.

The responses to these extra questions help test whether the responses obtained from the payment question are valid and reliable, which contributes to establish the credibility of the CV scenario. Attitudinal, opinion, knowledge, and use questions are used to assess "respondents' attitudes, perceptions, or feelings about the subject of interest" (Bateman et al., 2002, p.147). However, the functions of attitudinal questions in contingent valuation questionnaires are several. They 'warm-up' respondents, build up their trust and improve the

---

[18] Debriefing and follow-up questions serve two main purposes (Bateman et al. 2002, p. 145). First, they can be used "to explain why respondents were or were not willing to pay for the change presented" (as described more in detail in Section 11.1) and second "to explain respondents' views of the scenario presented."

rapport with the interviewer, set the tone for the rest of the interview, and get respondents involved in the questionnaire. They help respondents to think about the different aspects of the policy change being valued and encourage them to investigate their preferences about it. They help provide valuable qualitative and quantitative information that may help to validate the monetary valuations. Finally, these variables also often turn out to be good predictors of willingness to pay, so they can enter as explanatory variables in the willingness to pay functions estimated by most contingent valuation studies.

The questions about socio-demographic characteristics make it possible to complete a statistical analysis of how willingness to pay varies with respondent characteristics. This, again, helps build a willingness to pay function, which can also be used by other researchers in benefit-transfer studies. These questions also help to determine whether or not the sample is representative of the larger population of interest and, if not, to make adjustments in the estimated welfare measures before extrapolation to the population.[19] Additionally, they permit to find out how willingness to pay is related to household characteristics. This also helps establish the credibility of the valuation exercise, since the validity of the valuation measure is more credible when it can be shown to depend in ways predicted by economic theory on the characteristics of the household. For example, it is usually expected that willingness to pay should increase with income.

Socio-demographic characteristics usually suspected of influencing willing-

---

[19]In our case, we weighted the observations to make the sample more representative of the population of Newfoundland in terms of age and education.

ness to pay for risk reduction include age, gender, education, income, number of children, geographical location (rural versus urban), health status, and any measure of risk aversion.

More details about the specific covariates we considered can be found in Section 12.4.

6.7 Potential endogeneity of predictors of willingness to pay

As described above, in CV studies WTP functions are routinely estimated to identify the variables that affect WTP, which can help to test the theoretical validity of the estimated welfare measures when economic theory guides the empirical model. As part of these covariates in WTP functions, one can often find observed behavioural choices (membership in a conservation group, visits to certain recreational areas, experience whether negative or positive, with some aspect of environmental quality or with some type of environmental resource). However, these variables may be endogenously determined if the error term in the behavioural model is correlated with the error term in the willingness to pay model. If this is the case, including them in the explanatory model would lead to biased and inconsistent parameter estimates. On the other hand, since they often significantly affect willingness to pay and are usually correlated with other explanatory variables in the model, leaving them out could cause omitted variable bias.

Explicitly investigating the potential for endogeneity and, if detected, addressing the problem with the appropriate techniques is necessary when the potential for endogeneity is high. This issue has received very little attention in the literature. Cameron and Englin (1997), Whitehead (2005), Whitehead (2006), Garcia et al. (2008), Martínez-Espiñeira and Lyssenko (2011), and Lyssenko and Martínez-Espiñeira (2012a) would be examples of the few papers that explicitly account for the endogeneity of these types of variables. Their focus has been limited so far almost exclusively to variables that capture the respondents' previous experience with the resource.

In this study, we use several variables that could introduce endogeneity in the willingness to pay model. One of these variables, used to proxy underlying unobservable attitudes towards risk, is an indicator of the type of car most usually driven by the respondent. We also use an indicator of previous experience of MVCs (whether someone has suffered a collision or a near miss). It is also conceivable that variables like the indicators of the degree of risk perception are also endogenously determined with willingness to pay. Future extensions will deal with this issue but the basic results reported in this study do not consider this potential for endogeneity. Therefore, the conclusions drawn from the analysis of the effects on willingness to pay of covariates that could be endogenous should be taken with caution. For further details of the analytical complexities involved, see Lyssenko and Martínez-Espiñeira (2012a).

Chapter 7

Analysis of double-bounded dichotomous choice questions

7.1 Introduction

As explained in Chapter 6, there exist several different ways to elicit respondents' WTP for a public policy using the CVM. The discrete-choice format is strongly recommended by the NOAA Panel guidelines (Arrow et al., 1993). It involves proposing a *bid* (in terms of a contribution through a voluntary donation an increase in taxes, fees or prices, depending on the payment vehicle) to the respondent together with a question like "if the policy were to cost you $x,

would you be willing to pay that amount?" In this study we used a variant of the dichotomous choice format for the willingness to pay question, which was first introduced by Bishop and Heberlein (1979) and is the most widely used format nowadays.

The dichotomous choice payment format is also expected to be, if not fully incentive-compatible, the most incentive-compatible (that is, the most likely to elicit the truthful preferences of respondents) among the available payment formats. Another advantage is that it is relatively simple for respondents to answer, since it closely mimics a real market situation, whereby respondents only need to decide whether to accept or reject their payment of the proposed bid for a certain level of (hypothetical) risk reduction.

The main disadvantage of the dichotomous-choice payment format is that it is statistically inefficient. By asking whether the respondent is willing to pay or not certain amount, the researcher only finds out whether the individual's WTP is below or above that bid level. That is, their willingness to pay is "single bounded" and using parametric assumptions about the underlying WTP distribution to effectively overcome this sparse information can largely affect the resulting welfare estimates (Carson et al., 1999).

Not surprisingly, one of the first refinements, initially proposed by Hanemann (1985) and Carson (1985), of this payment format involved adding a follow-up question using a higher (lower) bid level after a positive (negative) initial response. This format allows the researcher to "doubly bind" the WTP for some

114

of the respondents[1] and to at least move the single bound closer to the WTP for the rest. This results in increased statistical efficiency, as first formally shown by Hanemann et al. (1991). Therefore, the extra information makes it possible to conduct CVM studies at a lower cost (with smaller sample sizes), while holding the precision of the WTP estimates constant or makes it possible to increase the precision of the estimates (lower variance) holding the sample size constant (for the same surveying cost). In fact, double-bounded dichotomous choice (DBDC) estimators have become very popular in the valuation literature, because they usually yield dramatically smaller confidence intervals around point estimates of statistics of the willingness to pay distribution (Carson et al., 1999).

We therefore used the DBDC format, in order to take advantage of the fact that, while maintaining the incentive compatibility of the single-bound dichotomous choice (SBDC) format, it collects more information from each respondent (Hanemann et al., 1991; Alberini, 1995a; Haab and McConnell, 2002, p. 114). In particular, our DBDC approach involved initially asking respondents whether they would be willing to pay x dollars for the proposed risk reduction (in some cases by renting a private safety device for the cars any other cases by paying additional taxes for a public risk reduction program) and, if respondents answered 'yes' ('no'), our asking them a similar question using as the original dollar amount (the *bid*) at a level twice (half) as large.

A disadvantage of the double-bounded dichotomous choice format is that the responses to the second bid may be unduly influenced by the initial bid or the

---

[1] These respondents are those who end up providing a YES-NO or a NO-YES response to the two questions.

response to the initial bid, resulting in potentially biased estimates. In practice, many empirical studies do find theoretically inconsistent results whereby the mean WTP differs significantly depending on whether it is calculated using information from the first question only or by the follow-up question. This issue is, of course of great relevance given that, once such undesirable response effects are controlled for, the efficiency gains obtained from the use of the double-bounded format may be lost. This tradeoff between the increased efficiency afforded by the double-bounded dichotomous choice format over its single-bounded counterpart and the potential for biases that it introduces will be considered in Section 7.3.

7.2   Estimating willingness to pay

Using the information from the DBDC responses, we estimated the individual mean WTP (the *location parameter* of the WTP distribution) for the average level of risk reduction presented to the respondents, as well as a measure of the spread (that is, the *variance*) of that distribution about its mean (also know as the *scale parameter*). Under certain (relatively restrictive assumptions) either a constrained bivariate probit model (Cameron and Quiggin, 1994) or an interval data model (Hanemann et al., 1991) can be used to analyse the data, two basic approaches compared by Alberini (1995). The interval-data model appears to be the most common technique (Bateman et al. 2002, p. 219, Boyle, 2003, p. 149) and was used as the benchmark to model willingness to pay in

this study. However, since this model is based on several restrictions about how the responses to the two questions are linked,[2] we also considered several alternatives, as further explained below, many of which are extensions of the basic interval model based on relaxing these restrictions or, equivalently, extensions based on using the bivariate probit model without constraints. Additionally, we ran random-effects models based on the interpretation of our data structure as a pseudo-panel, constructed by stacking the information from each question for each respondent (as in, for example, Alberini et al., 1997; Whitehead, 2002). This model corresponds, under a different parametrization, to a version of the bivariate probit with unconstrained error correlation (ρ) between the equations explaining the first and the second response and the coefficients of the the intercept and slope coefficients of these jointly estimated binary regression equations constrained to be the same across responses.[3]

In order to calculate mean WTP using the interval model, we consider four intervals[4] that correspond to the four possible response patterns obtained from the DBDC questions: 1) no-no (nn) 2) no-yes (ny) 3) yes-no (yn) and 4) yes-yes (yy). Each respondent was proposed only two bids (or, in the special case in which "don't know" was the response to the first question, usually three), an

---

[2] Or, equivalently, restrictions based on the notion that the two questions are prodding at a distribution of WTP values that is determined *a priori* and remains unchanged during the elicitation process.

[3] Conversely, the bivariate probit model can be seen as a generalization of the random effects probit model with the coefficients are allowed to vary across valuation questions (Haab, 1997; Whitehead, 2002).

[4] This interval model can also be re-parameterized considering only two bid values (the lowest one used across the two payment questions and the highest one) and defining only three intervals in which the respondent's WTP can fall. These intervals would be below the lowest bid, above the highest bid, or in between both. See, for example, Cameron and Quiggin (1994).
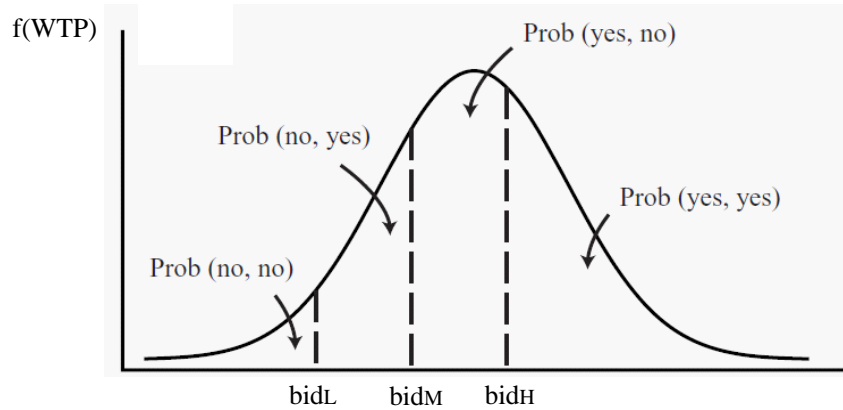
Figure 7.1: Probability density function of willingness to pay.

initial bid and a follow-up. However, three bids enter the relevant log-likelihood function, because *a priori* it is not known which follow-up bid amount will be proposed (the doubled bid or the halved bid in our case), as explained by Hanemann and Kanninen (1996, p. 64). The four intervals that build the log-likelihood function are bounded by these three bids. The *firstbid* (given by COST[5] and alternatively denoted bid_M) corresponds to the initial bid proposed, randomly assigned across respondents. The second bid was COST2 (bid_H), twice the size of COST, if the *firstresponse* was a 'yes' and COSTH (bid_L), half the size of COST, if it was a 'no'. A DK was followed by a randomly assigned COST2 (bid_H) or COSTH (bid_L).[6] The true value of maximum WTP held by

---

[5] Except for the case of a 'don't know' response, as explained in Footnote 6.

[6] If a DK response was given to this initial bid, COST, the whole process was begun anew by asking bid_L or bid_H, which would therefore become the new *firstbid*. If during that process the follow-up question became redundant our interviewers omitted it (and we later filled up ourselves the missing but obvious value for the corresponding variable). More details are

118

a respondent falls within one of the four intervals, which can be visualized in

Figure 7.1 (e. g. Train, 2003, p. 170).

Each interval is defined by a cumulative distribution function, as shown in

Equation 7.1:[7]

$$nn \quad : \quad Pr(W_i < bid_L) = F(bid_L; \alpha, \sigma^2) \tag{7.1a}$$

$$ny \quad : \quad Pr(bid_L \leq W_i \leq bid_M) = F(bid_M; \alpha, \sigma^2) - F(bid_L; \alpha, \sigma^2) \tag{7.1b}$$

$$yn \quad : \quad Pr(bid_M \leq W_i \leq bid_H) = F(bid_H; \alpha, \sigma^2) - F(bid_M; \alpha, \sigma^2) \tag{7.1c}$$

$$yy \quad : \quad Pr(W_i \geq bid_H) = 1 - F(bid_H; \alpha, \sigma^2) \tag{7.1d}$$

where $i = 1$ to n indexes respondents. Equations 7.1a to 7.1d are combined

to form the following log likelihood function:[8]

$$\ln L = weight_i \times \sum_i \begin{bmatrix} yy \times \ln[1 - F(bid_H; \alpha, \sigma^2)] + \\ yn \times \ln[F(bid_H; \alpha, \sigma^2) - F(bid_M; \alpha, \sigma^2)] + \\ ny \times \ln[F(bid_M; \alpha, \sigma^2) - F(bid_L; \alpha, \sigma^2)] \times + \\ nn \times \ln[F(bid_L; \alpha, \sigma^2)] \end{bmatrix} \tag{7.2}$$

This model can be motivated by assuming that there is a latent WTP given

---

available in Section 11.1.

[7] Since the latent WTP construct is assumed to be continuously distributed the probability of each punctual outcome is zero, so the placement of the equality signs in the inequalities is irrelevant.

[8] As shown in Equation 7.2, we allowed the possibility of using sampling weights $weight_i$. Indeed, the results reported in this report are based on sampling weights constructed to correct for the oversampling of individuals with certain age-education characteristics.

by a systematic component that is common to all respondents and an error whose size and sign depend on the individual i:

$$W^* = \mu + \varepsilon_i \qquad (7.3)$$

This unconditional prediction of WTP^* would be the same for all respondents. However, we can make the systematic component of the latent WTP a function (which we assume for simplicity to be linear) of the characteristics of the respondent and the payment scenario that respondent faced during the survey. That way the predicted WTP, that is, the systematic component μ of WTP^* will vary across respondents too depending on the observable characteristics on which it is conditioned:

$$W^* = x_i\beta + \varepsilon_i \qquad (7.4)$$

Using this willingness to pay or expenditure difference approach to modeling WTP, we just need now to realize that in Equation 7.2 the term α refers to the systematic component of the WTP^* construct and that σ corresponds to the standard deviation of the distribution of the stochastic component ε_i. It is necessary, though, to make an assumption about how that error is distributed. We experimented with three different cumulative distributions when modeling the log-likelihood in Equation 7.2: the normal (our base model which can also be readily compared with restricted versions of the bivariate probit and the random effects probit), the logistic, and the log-normal.

120

The cumulative distribution functions used in the analysis are shown below:

$$Normal \quad : \quad F(z, \alpha, \sigma^2) = \Phi\left(\frac{z - \alpha}{\sigma}\right) \tag{7.5}$$

$$Logistic \quad : \quad F(z, \alpha, \sigma^2) = 1 - \left[1 + \exp\left(\frac{z - \alpha}{\sigma}\right)\right]^{-1} \tag{7.6}$$

$$Log - normal \quad : \quad F(z, \alpha, \sigma^2) = \Phi\left(\frac{\ln z - \alpha}{\sigma}\right) \tag{7.7}$$

These equations are to be substituted into Equation 7.2 to construct the relevant likelihood function.

We estimated the WTP using the approach suggested by Cameron and James (1987) and Cameron (1988), which directly yields estimates values for the parameters of interest. The slope coefficients estimated through this approach can be interpreted roughly as those from an OLS regression on the level of WTP and the estimated indicator of the dispersion of the distribution of WTP can also be obtained directly as variable σ (Cameron, 1988).

The maximum likelihood routine in Stata 12 (as in e.g. Bosetti and Pearce, 2003, Gould et al., 2003) can be used to estimate the parameters α and σ resulting from each model. The first parameter (α) is considered the location parameter, while the second parameter (σ) measures the standard deviation of the WTP around the location parameter, so it is a measure of the scale of the distribution. In the case of the normal distribution, the contributed command *doubleb* (López-Feldman, 2010; López-Feldman, 2012) and the Stata-standard command for interval censored regression *intreg* were used equivalently.[9] Furher

---

[9]See Cameron and Trivedi (2010, 548–550) for a discussion of the differences among cen-

details about the mechanics of this type of estimation can be found, for example, in Hanemann and Kanninen (1996, p. 379-388); Train (2003, pp. 168-170); Haab and McConnell (2002, pp. 114-125); and Verbeek (2008, p.192-194).

In all cases, we first used a constant-only bid function to estimate the mean (and the median when different) WTP (Bateman et al., 2002, p. 197). This welfare measure was calculated, depending on the model, as shown in Table 7.1.

Table 7.1: Formulae for calculating mean and median WTP from results of maximum likelihood regression versions

Model	Mean	Median
Normal	α	α
Logistic	α	α
Log-normal	$\exp\left(\alpha\frac{\sigma^2}{2}\right)$	$\exp(\alpha)$

Additionally, a series of more flexible fully specified parametric models of willingness to pay that used additional covariates were estimated in order to observe how the willingness to pay for risk reductions changed with characteristics of the respondents and the particular version of the risk reduction policy proposed. This helped us assess the validity of our valuation exercise.

As opposed to modeling the WTP directly as just described, many CVM works the bid of the referendum is included only as one of the regressors of a conventional probit/logit analysis parameterized in such a way that the resulting estimates must be manipulated in order to obtain measures of mean (and median) WTP. Details of these manipulations as well as information on computing tools to make the task easier can be found in, among others, Cameron and

---

sored, truncated, and interval data.

James (1987); Cameron (1988, 1991); Haab and McConnell (2002, pp. 114-125), Jeanty (2007), Lopez-Feldman (2010, 2011, 2012). Following this approach, we estimated independent probits for the data stemming from each of the responses and then used bivariate probit models to jointly estimate the responses to both questions, first with no restrictions about the equality of the coefficients and the correlation term and then imposing increasingly restrictive assumptions on these parameters, so we would end up with a model equivalent to the interval model above.

The DBDC data can also be analysed using a random-effect approach, since the two responses constitute two observations of the choices of a given individual. We can, under this approach assume that the latent variable, WTP^*, consists of two components, a component that is permanent over repeated questioning for a given respondent, and a transitory component that is associated with the particular question, initial or follow-up (Alberini et al., 1997). This model can be written considering again that the latent WTP construct includes a systematic component

$$W^* = \mu = x_i\beta \tag{7.8}$$

given by the observable individual characteristics of respondent i and an unobservable component v_i:

$$W^* = \mu = x_i\beta + v_i \tag{7.9}$$

This error component varies across individuals but remains fixed across the

individual's responses, introducing a correlation across responses to the two payment questions. The random effects model accounts for this correlation in a manner akin to the random-effects models used for continuous dependent variables.

Furthermore, denoting the latent WTP informing the first and second responses of respondent i:

$$WTP_{i1}^* = \mu_i + \varepsilon_{i1}$$

$$WTP_{i2}^* = \mu_i + \varepsilon_{i2}$$

where ε_{i1} and ε_{i2} are assumed independent of each other and of the term v_i, so the variances of each of the latent WTP variables is given by:

$$V(WTP_{i1}^*) = \sigma_{vi} + \varepsilon_{i1}$$

$$V(WTP_{i2}^*) = \sigma_{vi} + \varepsilon_{i2}$$

The correlation coefficient ρ between WTP_{i1}^* and WTP_{i2}^*, which is equal (assuming that the variability of ε_{i1} and ε_{i2} is the same, so no within respondent heteroskedasticity exists) to:

$$\rho = \frac{\sigma_v}{\sigma_v + \sigma_\varepsilon}$$

This correlation will take a low value if the variance of the transitory component is large relative to variance of the permanent component and will be close to one when the variance of the permanent component is much larger than those of the

transitory components. When the two error components v and ε are assumed to be normally distributed, the random-effects is equivalent to a bivariate probit model that allows for a free correlation coefficient (Alberini et al., 1997). The larger this correlation ρ, the more likely the random effects model is appropriate (Greene, 1997, 896–899) relative to the pooled probit. In the extreme case of the double-bounded model, the first and second individual WTP^* amounts are assumed identical ($\rho = 1$), so the interval double-bounded model is obtained as a special case of the random effects model. Note also that the bivariate probit model with the bid and constant coefficients constrained to be equal across response equations and a free correlation coefficient ρ is equivalent to the random-effects probit, only with a different parametrization.

When ρ is zero, the panel-level variance component is unimportant, and the panel estimator is equivalent to the pooled estimator. A likelihood-ratio test of this null hypothesis can be used to assess the significance of the estimate of ρ and thus formally compare the pooled estimator with the panel estimator.

In the case of all three general estimation approaches described above, the Krinsky-Robb procedure (Krinsky and Robb, 1986; Krinsky and Robb, 1990) was used to estimate confidence intervals for the parametric models using random draws of 10,000 new values for each parameter (alpha and sigma) from a multivariate normal distribution (distributed according to the means, variances, or covariances of each parameter). The parameters from each random draw were then used to calculate new estimates of mean and median willingness to pay (accept). These estimates were then arranged from lowest to highest and

the 25th and 975th estimate represented the 95% percent confidence intervals. Jeanty (2007), Haab and McConnell (2002, pp. 110-113) and Hole (2007), for example, provide further details on this procedure.[10]

Our analysis of the data considered, additionally, distortions of the basic models identified during recent years and thus takes into account issues of anchoring, starting point bias, yea-saying, and framing (Herriges and Shogren, 1996; McLeod and Bergland, 1999; DeShazo, 2002; Whitehead, 2002; Chien et al., 2005; Flachaire and Hollard, 2007b; Watson and Ryan, 2007; Farmer and Belasco, 2011). These issues are described in detail in the next section.

7.3   Testing the Validity of the Survey Results

This part of the analysis examined whether the data conformed to a series of expected relationships between the independent variables and the estimated willingness to pay. For example, one would expect respondents with higher levels of household income to have a higher WTP for risk reductions, assuming road safety be a normal good. In our case, a key concern when it comes to assessing data validity is related to the sensitivity to scope (whether respondents are willing to pay more if they are provided with a more comprehensive good) and sensitivity to scale (whether respondents are willing to pay more for a greater quality or quantity of risk reduction). An additional issue, however, involved the potential for question effects in the double-bounded question format.

---

[10]Alternative methods of estimating confidence intervals have been suggested by the literature. See, for example, Cameron (1991).

7.3.1 Analysis of question effects with the DBDC format

As mentioned in Section 7.1, although the double-bounded format provides the advantage of increasing statistical efficiency, some concerns have been voiced about its validity (McFadden, 1994; Bateman et al., 2001; Carson and Groves, 2007). Double-bounded questions have often been found to result in lower estimates of WTP than comparable single-bounded formats. This suggests that it is possible that respondents may be motivated by different latent WTP values when answering the initial and follow-up questions, which contradicts the assumptions made in the basic types of analysis. In particular, a fundamental assumption behind the double-bounded dichotomous choice method is that respondents base both answers upon a "single untainted latent willingness-to-pay" (Czajkowski, 2009, p. 308), so their responses to both the first and the second questions are simply based on the comparison of this latent WTP value to the bid value proposed by the interviewer in each question. Statistically, this implies that, after appropriate conditioning, there should be perfect correlation between the WTP distributions implied by the responses to the two questions (Carson et al., 1999). This is related to the assumption that both response choices are made independently (Carson et al., 1999), with no undue influences by the bid faced in the first question or the answer given to the first question on the respondent's response to the follow-up question.

Soon after the DBDC format was proposed, Cameron and Quiggin (1994) formally examined the assumptions behind the basic treatments of double-bounded dichotomous choice data, first revealing several stylized facts now commonly ob-

served when it comes to the comparison of the WTP estimates based on the first binary discrete choice question (the single bounded approach that exploits the data from just the first question) and both binary discrete questions. First, it is commonly found that the WTP distributions implied by the first and second questions are not perfectly correlated. Second, the WTP estimate obtained from the single-bounded analysis of the first response is substantially higher than the one obtained by exploiting the information from both questions, which is precisely the point of collecting the double-bounded data in the first place. Third, the number of negative responses to the second question is higher than would be expected based upon the WTP distribution estimate from the first question alone (which means that the analysis of the data from just the second question yields a substantially lower WTP estimate than that from the first question).

Several alternative hypotheses have been suggested to explain these anomalies and incoherencies. These usually[11] involve some sort of behavioral response or Bayesian updating by the respondent, including those related to starting point bias, yea-saying, nay-saying, uncertain costs, random response shocks, structural shifts, heteroskedasticity, and framing (Cameron and Quiggin, 1994; Herriges and Shogren, 1996; Alberini et al. 1997; Bateman et al. 2001; DeShazo, 2002; Whitehead, 2002; Burton et al. 2003). In addition, the econometric models have been updated to test and, when possible, correct for these effects

---

[11]Although some explanations would instead not compromise so much the validity of the DBDC format. For example, Hanemann et al. (1991) consider the potential for respondent fatigue or weariness in DBDC formats but assert that it is unlikely to be a factor in models with only one follow-up bid. Bateman et al. (2001) counter that the mere expectation of weariness may be sufficient for question effects to arise. The "government wastage" explanation (Carson et al., 1994b; Alberini et al., 1997) described below would also be consistent with rational Hicksian preferences (Bateman et al., 2001).

(Chien et al. 2005; and Flachaire and Hollard, 2006, 2007b; Araña and León, 2007). We consider the literature dealing with these issues next.

7.3.2   Behavioral explanations for question effects

There are several explanations of behavioral inconsistency when responding to double-bounded dichotomous choice questions.[12]   These explanations tend to put in question the validity of the responses, since they suggest that, to a certain extent preferences might be constructed, rather than just elicited, by the elicitation process (McFadden, 1994).

One of the earliest explanations, suggested by Carson et al. (1994) and Alberini et al. (1997) is the so-called "government wastage model" (or "cost expectations model" as in DeShazo, 2002). According to this model, respondents who initially say 'yes' to the payment question may reject any increased subsequent amount, because they would see it as an attempt by the government to collect more money than what would be needed to finance the cost of provision. On the other hand, respondents who reject the first offered bid may interpret the subsequent lowered second bid as a sign of decreased quality of the good provided, therefore becoming artificially likely to reject it. As a consequence, respondents are more likely to vote against the second offered bid regardless of whether they accept or reject the first offered bid, so we would expect under the influence of this effect an overall downward shift in the second willingness-to-pay (WTP) response.[13]

---

[12]This section borrows substantially from Kang et al. (2013).

[13]See also Mitchell and Carson (1989); Alberini et al. (1997); Carson et al. (1994); Hane-

As an alternative, Mitchell and Carson (1989) proposed a "strategic behavior model" in which respondents answer the first question truthfully but respond to the second question strategically, because they feel they are now entering in some sort of bargaining situation with the interviewer. Therefore, respondents try to strategically lower the price by rejecting any additional higher bid offers. And respondents who rejected the first bid would tend to also reject a lower bid, hoping to be offered an even lower bid. Under this interpretation the respondents will become more likely to reject any follow-up question, regardless of whether their true WTP is higher or lower than the follow-up bid, leading to a downward shift in the second WTP and would explain why a lower WTP estimate is often obtained by the basic DBDC model than by the SBDC model applied to he initial bid only.

A related explanation has to do with two highly related effects, 'indignation' and 'guilt' (Cameron and Quiggin, 1994; Bateman et al., 1999) propose two highly related effects, 'indignation' and 'guilt'. The indignation effect would lower the probability of a positive follow-up response by respondents who agree to pay the initially proposed amount but then resent then being asked whether they would be willing to pay a higher amount in the follow-up question. Conversely, guilt would increase the probability of a positive follow-up response if respondents who initially refuse the first bid amount then feel an "elevated sense of social responsibility or simply embarrassment" (Bateman et al., 1999, p.195) when responding to the lower follow-up bid, so they feel more inhibited from

---

mann et al. (1991); Kanninen (1995); McFadden and Leonard (1995).

rejecting it. Alternatively, reducing the price might suggest to some respondents that, by reducing the bid, the interviewer is signalling a willingness to supply the good and that in turn makes the provision of the good seem more important than initially perceived. These effects would be likely heightened in the case of public goods (Bateman et al., 2001). As noted by Bateman et al. (2001), the "guilt/indignation" effects are identical to the effect of anchoring, discussed next, and may be considered as psychological interpretations of the latter.

Herriges and Shogren (1996) hypothesized that the unexpected question effects in the double-bounded model could be due to the influence of an "anchoring effect". Tversky and Kahneman's (1974) described the anchoring effect generally as the process in which "people make estimates by starting from an initial value that is adjusted to yield a final answer." In their Bayesian updating approach, respondents are assumed to answer the follow-up question based on their posterior expectations of WTP, which is the weighted sum of the first bid they faced and their prior expectations for WTP. The effect of the anchor on the mean of the second WTP response depends on the relative magnitude of the anchor and the prior expectation of WTP. Herriges and Shogren (1996) show how the magnitude of this relative effect can be estimated from the data using maximum likelihood. However, since respondents offered different initial bids have different anchors, the implications of this model on the initial and follow-up responses are unclear.

Finally, DeShazo (2002) suggested a "framing effects model" based on "prospect

theory" which predicts a downward bias in WTP in the ascending bid sequence (yes to the first response), since the second question is now negatively framed as a loss in the eyes of the respondent relative to the prospect of obtaining the good at the previously asked, lower bid. That is, respondents would have relied on having closed a deal whereby they secured some positive consumer surplus and then that deal seems compromised by the interviewer's asking about a higher bid. However, in a descending bid sequence (starting with a 'no' as the first response) no corresponding positive frame occurs in the second, lower bid, offer because the respondent answered 'no' to the first offer, so the "deal" was not struck. This means that one should expect under the influence of this effect an overall downward shift in WTP, but less than when the "government wastage model" or the "strategic behavior model" are dominant.

Carson et al. (2000) after considering several plausible hypotheses for how the second bid influences the respondent's second response, concluded that, on balance, one would expect that WTP estimates from a double-bounded format to be smaller than those from a single-bounded format.

One thing to keep in mind is that each explanation provided above relates only to the average dominant outcome in a given sample and that individual psychological inconsistency in responses cannot be identified. For example, some people might respond to the follow-up question with a strategic motive while others' responses are anchored to the initial bid, and yet the overall pattern in the data set might best fit a framing effects explanation.

Finally, although one might be tempted to discard the information provided

by the second response (at least in cases where significant question effects have been detected) it may be still desirable to use the double-bounded format, after accounting for the tradeoff between the likely downward bias generated by the question effects and the tighter confidence intervals afforded by the use of additional information from the second response (Alberini, 1995).

7.3.3 Analysis of question effects

As explained above, although several different models can be used to estimate the mean WTP from double-bounded questions (Haab and McConnell 2002: 115-125), the earliest specifications, such as the interval model (Hanemann et al., 1991), assumed that respondents consult the same latent WTP value when answering both the initial and the follow-up questions. This assumption is relaxed in the bivariate probit model, introduced by Cameron and Quiggin (1994), by allowing preferences to fully vary over both questions.

The flexible bivariate probit allows for differences between questions when it comes to the WTP coefficients building z, the associated regression error terms ε, and their dispersion (measured by σ) in Equations 7.5-7.7 (Section 7.2). The correlation between the errors of the two probits is usually expressed in terms of parameter ρ in the notation of the bivariate probit. The relatively restrictive, interval model assumes a correlation of 1.[14] The interval model is then a restrictive version of the bivariate probit model, which means that its

---

[14] If there is no correlation at all, the responses to the two payment questions could be analysed separately with two completely independent probits. Again, the null hypothesis that $\rho = 0$ can be tested using a likelihood-ratio test.

restrictions can in most cases be tested using likelihood-ratio tests.[15]

Other specifications, such as the random effects probit (Alberini et al. 1997), also allow preferences to vary but account for the pseudo-panel structure of the initial and follow-up questions. We estimated an interval model with a normal distribution, a general bivariate probit model, and a random effects probit model. Furthermore, we used other variants of these basic models to test for several question effects discussed in the literature.

Since most of the statistical testing associated with the analysis of question effects is based on the bivariate probit model, we abstract in this section from the logistic and log-normal distributions and focus on the results of our benchmark model, which is based on the normal distribution.

In our analysis, we considered several standard models to deal with question effects, all assuming that the second response is affected by the first bid. These models assume that a prior willingness-to-pay W_i is used to respond to the first bid, and that an updated willingness-to-pay $W_i\prime$ informs the response to the second bid.

First, a model with a shift effect[16], whose sign we did not restrict a priori, was modeled by introducing an additional parameter δ such that:

$$W_i\prime = W_i + \delta \tag{7.10}$$

---

[15] In some cases, however, we cannot easily discriminate between competing models using statistical testing (Alberini et al., 1997).

[16] As proposed under different theoretical justifications, by Carson et al. (1994), Alberini et al. (1997), and DeShazo, (2002).

The second alternative model we used was based on Herriges and Shogren (1996) hypothesis that an "anchoring effect" could affect the response pattern:

$$W_i\prime = (1 - \gamma)W_i + \gamma bid_{1i} \qquad (7.11)$$

so respondents are assumed to respond to the follow-up bid based on their posterior expectations of WTP, which is the weighted sum of the first bid proposed (bid_{1i}), which acts as an anchor, and their prior expectations for WTP. A value of γ equal to zero indicates that there is no anchoring and that the respondent's latent WTP is unchanged in the second question, so all of the efficiency gains from the second question remain intact. A value of one implies that the first bid completely replaces the respondent's WTP, so no new information is obtained from the follow-up question.

Next, we considered a model with both a shift effect and an anchoring effect (as in Whitehead, 2002):

$$W_i\prime = (1 - \gamma)W_i + \gamma bid_{1i} + \delta \qquad (7.12)$$

Finally, following more recent works, such as Watson and Ryan (2007), Aprahamian et al. (2008), Araña and León (2008), Schwarzinger et al. (2009), Jennings et al. (2010), we additionally considered the possibility of a heterogeneous anchoring effect:

$$W_i\prime = (1 - \gamma_i)W_i + \gamma_i bid_{1i} + \delta \qquad (7.13)$$

These alternative models were compared to the basic double-bounded model (the interval model suggested by Hanemann et al., 1991), which assumes that respondents uses their true and unchanged (W_i), to respond to both the first and second payment questions:

$$W_i = W_i\prime \qquad (7.14)$$

We report also a single-bounded model based on the first response only, which results in a much less efficient estimate of mean WTP (and much less conservative too) than any of the double-bounded treatments but can be considered free of any of the biases due to question effects.[17]

Our most flexible model, described by Equation 7.13, which allows for shift effects and heterogeneous anchoring, as well as for heteroscedastic errors, was estimated through maximum likelihood by considering the following four prob-

---

[17] Whether one would rather discard the information provided by the second response (at least in cases where significant question effects have been detected), it may be still desirable to use the double-bounded format, after accounting for the tradeoff between the likely downward bias generated by the question effects and the tighter confidence intervals afforded by the use of additional information from the second response (Alberini, 1995a).

abilities in Equations 7.15-7.18 behind each possible response obtained.[18]

$$Pr(nn) = \Phi\left(\frac{\frac{bid_{2i}-\gamma_i bid_{1i}-\delta}{1-\gamma_i} - x_i\beta}{\sigma_i}\right) \tag{7.15}$$

$$Pr(ny) = \Phi\left(\frac{bid_{1i}-x_i\beta}{\sigma_i}\right) - \Phi\left(\frac{\frac{bid_{2i}-\gamma_i bid_{1i}-\delta}{1-\gamma_i} - x_i\beta}{\sigma_i}\right) \tag{7.16}$$

$$Pr(yn) = \Phi\left(\frac{\frac{bid_{2i}-\gamma_i bid_{1i}-\delta}{1-\gamma_i} - x_i\beta}{\sigma_i}\right) - \Phi\left(\frac{bid_{1i}-x_i\beta}{\sigma_i}\right) \tag{7.17}$$

$$Pr(yy) = 1 - \Phi\left(\frac{\frac{bid_{2i}-\gamma_i bid_{1i}-\delta}{1-\gamma_i} - x_i\beta}{\sigma_i}\right) \tag{7.18}$$

The probabilities for the conventional double-bounded, anchoring model can be obtained by imposing restrictions in Equations 7.15-7.18. Restricting $\sigma = \sigma_i$ for all respondents leads to the homoscedastic counterparts of all the models while restricting γ to take the value of zero leads to the shift model. Imposing $\delta = 0$ yields the anchoring model and imposing $\delta = 0 = \gamma$ results in the conventional interval model.

---

[18] After the elimination of "don't know" responses.

Part V

Further issues in the valuation of risk reductions

Chapter 8

Risk perception

Risk can be defined in a number of ways. However, in the most simplistic of terms it can be described as the likelihood that an individual will experience the effect of danger (Short Jr., 1984). Risk perception involves the evaluation of the probability and consequences of an unwanted outcome occurring. This evaluation of risks is influenced by both individual and social characteristics (Sjöberg et al., 2004).[1]

The estimated value of a statistical life, in particular as obtained through stated preference methods, depends on how individuals perceive mortality risks. It depends, in particular, on the individual perception of baseline risks and of the probability changes valued. If, for example, risks are perceived to be higher than they actually are, estimates of the value of risk reductions are expected to

---

[1] For a depiction of the fascinating history of the science of risk management, see Bernstein (1996).

be higher than if individuals were better informed (Gayer et al., 2000; Bleichrodt and Eeckhoudt, 2006).

Andersson and Treich (2011) point out that a strong and diverse corpus of empirical evidence suggests that individuals are generally quite rational in their decision-making involving risks in the marketplace (Viscusi and Aldy, 2003; Blomquist, 2004) but some other results imply that the estimated "risk-dollar" tradeoffs may not be accurate (Viscusi and Magat, 1987). Stated preference studies also tend to show evidence of ordinal but not cardinal risk comprehension (Hammitt and Graham, 1999), so individuals can be seen as responding correctly to risks, while "their ability to perceive risk in a cardinally correct way is questioned" (Blomquist, 2004, p. 99).

The next sections briefly describe the different approaches used to explain how individuals approach, perceive, and judge risk levels in general, before focusing on road traffic risks. One of the key issues of valuation studies that deal with the effects of risk reductions on economic welfare involves the fact that most individuals have quite a lot of difficulties understanding the meaning of risk measures and understanding the implications of changes in risk. This is particularly problematic in the case of remote, or small in general, and unfamiliar risks. There seem to be serious problems of risk perception both among non-experts and, in some cases, also experts. In the next sections we consider how the issue of risk perception has been dealt with in the specialized literature and, more particularly, how it affects the valuation of risk reductions in road traffic.

8.1 Heuristics and risk perception

Initially, risk perception was predominantly studied through the scientific application of judgmental heuristics. This theory argued that people did not make valid intuitive judgments of probabilities as defined and computable by probability calculus (Tversky and Kahneman, 1974) but are instead unduly affected by irrelevant factors and by the availability of evidence (Tversky and Kahneman, 1973). Overall, availability heuristics' research on risk perception resulted in counterintuitive findings and, as a result, it is considered to be a fairly unimportant approach used to study risk perception today (Sjöberg, 2000).

Commonly studied using the availability heuristic is the influence of the media on risk perception. Public attitudes towards risks (whether they accept, reject, modify, tolerate or eliminate a risk) are expected to be influenced by the media's presentation of these risks (Boholm, 1998). If there is an over-representation of a risk in the media, this may lead the public to overestimate this risk. However, while looking at their results, Sjöberg et al. (1996) found that there was no evidence which implied that the media influences the public's perception of risks, as far as quantitative correlations between perceived levels of risk and coverage in the media goes (Boholm, 1998). Conversely, researchers like Kasperson et al. (1988) found that coverage of risk in the media can lead to an 'amplification' of fears about risks. Clearly the role of the media in risk perception is very much under debate (Sjöberg, 2000).

One interesting phenomenon that the heuristics approach brought to the study of risk perception is "risk denial". This is characterized by an unrealistic

141

type of optimism leading people to believe that they are less at risk than others (Sjöberg, 2000). For example, risk denial is found in a study on risk perception and driving by Matthews and Moran (1986). In this study young and older men were asked to rate their confidence in their ability to maintain control of their vehicle, their confidence in their ability to make safe vehicle-handling decisions and their confidence in their driving reflexes necessary to avoid accidents. Both younger and older male drivers rated their own individual driving abilities as higher than those of their peers and rated their likelihood of experiencing an accident as lower than their peers' (Matthews and Moran, 1986). This phenomenon is not only observed in driving risk scenarios but in most risk contexts and it is not only a condition found in young and old men but tends to be a universal trait (Sjöberg, 2000). Despite the strengths found in the heuristics model, nowadays the implementation of this approach is fairly uncommon (as mentioned previously) among academics studying risk perception (Sjöberg, 2000). Risk perception is now most commonly studied using cultural theory or the psychometric paradigm.

8.2   Cultural theory and risk perception

Cultural theory argues that the perception of risks is influenced by "worldviews or ideologies entailing deeply held values and beliefs defending different patterns of social relations" (Wildavsky and Dake, 1990, p. 43). The term "social relations" refers to patterns of interpersonal relationships found within a given

culture. These patterns of interpersonal relationships exist between the four different types of people found within any given culture. These four different types of individuals are concerned with different types of risks and also look to different solutions when dealing with these risks (Wildavsky and Dake, 1990). *Egalitarians* are concerned with technological and the environmental risks; *individualists* with war and other threats to the market; *hierarchists* with law and order; and *fatalists* are concerned with none of the above (Sjöberg, 2000). When attempting to reduce risks, *hierarchists* turn to expert committees and universal safety standards for solutions. *Egalitarians* favour decision making processes which encourage public participation. *Individualists* promote economic factors, and in particular CBA to ensure rational decision-making. Finally, *fatalists* believe that decisions are beyond their control and will feel obliged to accept whatever is imposed upon them (Marris et al., 1996).

When attempting to evaluate the risk perception associated with cultural bias and various socio-demographic factors, a statistically significant relationship was found to exist between the two (Marris et al., 1996). They discovered that fatalists tended to be men who have less formal education. Hierarchists tended to be older, have less formal education, and lower incomes. Individualists tended to be older and have less formal education. Finally, egalitarians tended to be women and have higher education. Despite these findings, however, Marris et al. (1996) stated that, unlike the relationship found between cultural biases and socio-demographic factors, "[n]one of the correlations between cultural biases and risk perceptions were very high" (Marris et al., 1996, p. 22)

and that depending on how risk was being defined, either in terms of riskiness, fatalities, environmental harm, injuries or unacceptability the study came back with different correlations between risk perceptions and cultural biases.

Upon evaluating their findings, Marris et al. (1996) concluded that, based on their data, the greatest number of and most significant correlations were observed when risk was defined in terms of *Riskiness, Environmental Harm* or *Unacceptability*, as opposed to when it was defined in terms of *Fatalities* or *Injuries*. They found that an egalitarian worldview was correlated with high risk perceptions for environmental threats of a potentially catastrophic nature, and also for risks perceived as 'unnatural'. The hierarchical worldview was, as predicted, associated with high scores for social threats such as mugging and terrorism. The individualist worldview encompassed a low concern for environmental issues, as well as a low concern for risks which would be perceived (by individualists) as 'personal risks', alcoholic drinks, car driving, and food colourings, to name but a few (Marris et al., 1996). After testing to assess the relationship between cultural biases and risk perception, Marris et al. (1996) concluded that none of the correlations obtained were very high: in particular only 11%, at most, of the variation in risk perception was linked to cultural biases (Marris et al., 1996). However, they did state that the theory did shed some light on some of the factors that shape risk perception (Marris et al., 1996).

8.3 The psychometric paradigm and risk perception

According to Slovic et al. (1985), the psychometric paradigm "assumes risk to be subjectively defined by individuals who may be influenced by a wide array of psychological, social, institutional, and cultural factors" (Sjöberg et al., 2004, p. 10). The paradigm suggests that the relationship between these factors and risk perception can be quantified through the use of appropriate survey instruments. When this paradigm was first being developed researchers found that "experts" and "laypeople" defined "risk" in different ways. "Experts" typically based their definition of "risk" on the expected number of fatalities; "laypeople", on the other hand, had a definition which was comprised of more qualitative characteristics. Slovic et al. (1985) divided these qualitative characteristics into three broad categories: dread; unknown; and exposure. The public perceived the highest degree of risk to be associated with issues which were rated high on both the 'dread' and the 'unknown' categories (e.g. nuclear power, lasers, herbicides), and that the public perceived low-levels of risk in situations which they felt were not subject to the 'dread' or the 'unknown' categories (e.g. motor vehicles, alcoholic beverages, downhill skiing). This was the case regardless of the expected number of fatalities associated with 'risky' situations. The findings in Slovic et al. (1985) were ground-breaking at the time, since they challenged the assumption made by 'experts' and policy makers that non-experts were irrational and/or ignorant when they displayed high levels of concern for issues

such as nuclear power, while at the same time ignoring more common issues such as road accidents (issues such as road accidents were rated more high by 'experts') (Marris et al., 1996).

Typically, studies done using the psychometric paradigm have attempted to investigate variations in the risk perception of individuals based on standard socio-demographic variables, such as gender, age, occupation, nationality, etc. (Marris et al., 1996). However, many researchers have found that "very few of these variables were found to correlate consistently with risk perception. Furthermore, even when differences were identified, this approach provided no understanding about why different people perceive risks differently" (Marris et al., 1996, p. 2). Through their research, Marris et al. (1996) found some success when attempting to evaluate the relationship between thirteen risk issues: sunbathing, food colourings, genetic engineering, nuclear power, mugging, home accidents, ozone depletion, car driving, microwave ovens, AIDS, war, terrorism and alcoholic drinks; and nine characteristics, some of which were in line with those proposed by Slovic (1987). These nine characteristics were: involuntariness, delayed effects, severity, dread, catastrophic potential, harm to future generations, lack of knowledge to those exposed, lack of knowledge to scientists, unfairness. Marris et al.(1996) found that "eight out of the nine risk characteristics were closely related to risk perceptions. The exception was 'lack of knowledge to science', which was not correlated" (Marris et al., 1996). Specifically, they found that 'dread' and 'harm to future generations' appeared to be important concepts in the framing risk perceptions (Marris et al., 1996).

Using the psychometric paradigm, Marris et al. (1996) also evaluated risk perceptions of the thirteen risk issues based on socio-demographic factors: sex, age, level of education and household income. When analysing the data at the level of individual risk issues, Marris et al. (1996) found a significant, however weak, relationship between the socio-demographic attributes of respondents and the risk ratings they gave to each of the thirteen risk issues. When analysing risk perceptions in a general sense, Marris et al. (1996) had more luck in finding a significant relationship. They found that "women tended to rate risk higher than men [...], and analysis of variance indicated that this difference was statistically significant at the 95% confidence level" (Marris et al., 1996, p. 30). Although some researchers reported similar findings, others did not. Sivak et al. (1989), when evaluating cross-cultural differences in driver risk perception, found that there were no significant differences between the two sexes when participants were required to rate how likely they would be to get in an accident if put in various driving situations. In reference to age, Marris et al.'s (1996) findings suggested that, "older respondents tended to rate risk lower than younger people, but only when risk perception was defined as Fatalities, Injuries or Environmental Harm" (Marris et al., 1996, p. 30). A study done on age and driver risk perception by Matthews and Moran (1986) found a relationship to exist between the two variables. With younger drivers seeing themselves as immune from the effects of higher levels of risk, however they stated that this is more a problem for young males than young females. In terms of education level, Marris et al. (1996) found that those with a university degree tended to rate

the riskiness scale lower across all risk issues. Turning to income, respondents with higher household incomes typically rated risks lower than respondents with lower household incomes (Marris et al., 1996). Upon summarizing their findings on general risk perceptions and socio-demographic variables (especially sex, age and income), Marris et al. (1996) concluded that despite their being a statistically significant relationship, correlations were not very high, and that at most, only 10% of the variation could be explained by any one variable (Marris et al., 1996).

8.4 Limitations of the cultural theory and psychometric paradigms

Marris et al. (1996) concluded that the psychometric paradigm is more successful at explaining risk perceptions than cultural theory (Marris et al., 1996). However, they stated that despite the fact that the psychometric paradigm's methodology generates robust quantitative results, these do not provide much insight into the reasons why particular risk characteristics are closely correlated with risk perceptions. Marris et al. (1996) held a similar view on socio-demographic variables and their relationship to risk perception, stating that the relationships that were found to exist do not shed much light on why they exist. On the other hand, they found that cultural theory does provide explanations for risk perceptions by indicating how they fit coherently into worldviews held by respondents (these worldviews being either egalitarian, individualist, hier-

148

archist, or fatalist) (Marris et al., 1996). Overall, Marris et al. (1996) believe that both approaches have contributed to understanding how risk perception is shaped (Marris et al., 1996). They also believe that both approaches bring something unique to the table when studying risk perception, as a result an implementation of the two theories may produce the best insight (Marris et al., 1996).

Although Marris et al. (1996) highlight both the strengths and weaknesses of cultural theory and the psychometric paradigm, other researchers have a stronger opinion of the limitations of these theories. According to Sjöberg (2000), measuring risk perception is very problematic and neither of the two theories is very successful at doing so. Sjöberg (2000) found that, "in general, neither the cultural theory nor the psychometric model explain much of the variances found in risk perception. The latter model accounts, in its original form, for only some 20% of the variance of risk perception. [. . .] Cultural Theory is even less successful" (Sjöberg, 2000, pg. 8). Sjöberg (2000) states that:

> some of the models suggested for risk perception have failed to explain more than a rather small fraction of it. Some investigators have apparently been satisfied with statistical significance as a criterion of validity, but that is a counterproductive strategy. Others have presented seemingly persuasive results, but they have been based on averages and therefore quite misleading as to the explanatory power of the models. (Sjöberg, 2000, p. 9)

Despite the obvious issues surrounding the study of risk perception, it is

crucial for policy makers to attempt to understand risk perception to the best of their ability if they are to address issues found within societies.

8.5   Risk perception and road traffic risks

The perception of traffic risks is one of the most studied aspects of the perception of risks. And a key result of the risk perception literature is that, as Lichtenstein et al. (1978) showed, individuals tend to overassess small fatality risks and underassess large fatality risks, a pattern today regarded as an established fact (Benjamin and Dougan, 1997; Viscusi et al., 1997; Hakes and Viscusi, 2004; Armantier, 2006).[2]

We mentioned in Section 8.1 some of the results found in terms of risk denial (Matthews and Moran, 1986). Indeed road traffic risks tend to be underestimated by most people, matching the notion of the common type of risk mis-perception whereby individuals tend to overestimate death risk due to low probability events and underestimate death risk due to high-probability events .

Finn and Bragg (1986), for example, found that young drivers perceived the risk of an accident as significantly lower than did older drivers. When both young and older drivers were asked to estimate the risk of an accident to three groups: all drivers, young drivers, older drivers and themselves, both groups considered young male drivers to face a greater risk of an accident than

---

[2] Although Benjamin et al. (2001) raised some doubts over the results in Benjamin and Dougan (1997) , Armantier (2006) confirmed the main conclusions in the former study.

all drivers, who in turn had a greater risk of an accident than older drivers. However, young drivers saw their own individual risks as significantly lower than those of other young male drivers, while older drivers saw their risk as comparable to those of their peers.

Since a MVC is a low probability event we might expect individuals to overestimate their own death risk. However, familiarity with a risk and degree of perceived control (Sjöberg et al., 2004) tend to reduce perception of a risk. Since driving on the highway is both voluntary and familiar as well as something individuals tend to feel a lot of control over (driving), this should reduce the risk perception. Andersson (2011) found that females tend to over-assess their own road traffic risk perceptions and males tend to under-assess their own traffic risk. This may explain why females tend to be willing to pay more than males for traffic risk reductions (Andersson and Lundborg, 2007). Overall, it is expected that most people will understate their risk, due to the high degree of familiarity and control, and that younger males will understate risk more than others.

Chapter 9

Cognitive skills effects in the valuation of risk reductions

When estimating the benefits of risk reductions using stated preference methods, economists try to identify the marginal monetary tradeoffs that individuals would be willing to achieve very small changes in their risk of death or injury. In particular, CVM studies assume that respondents' preferences can be identified if the hypothetical market scenario included in the survey instrument is plausible, meaningful, and understandable for the respondents (Carson et al., 2001). When it comes to studies that value risk reductions, however, the empirical evidence overwhelmingly suggests that people are very imprecise when

stating their preferences about their own safety (Dubourg et al., 1997; Hammitt and Graham, 1999; Andersson and Treich, 2011), as we have already pointed out elsewhere in this report. This makes the use of the CVM for the elicitation the benefits from small changes of probability is particulary problematic (Carson et al., 2001), likely because of the general difficulty in understanding small (changes in) probabilities we have been considering in Chapter 8.

For this reason, a constant concern of those trying to examine preferences for risk reductions has been trying to identify the source of the problem and to ameliorate it. In particular, much work has been devoted to try to identify the roots of the problem of insensitivity to scale found in many CVM studies (Hammitt and Graham, 1999), which may be linked to the individuals' lack of understanding and perception of the remote probabilities involved in the usual risk reduction scenarios. For example, as we described in Section 4.2.9, visual aids and by training respondents in trading wealth for safety has been proven as somewhat successful in achieving scale sensitivity in line with the theoretical predictions (Corso et al., 2001b; Alberini et al., 2004).

A different line of approach has been followed, with researchers trying to identify to which extent indicators of cognitive skills about dealing with probabilities, fractions, and proportions correlate with the ability of respondents to respond to stated-preference questions about WTP for risk reductions in a way that reflects theoretical expectations. Information about who is expected to be better at handling the payment questions in the CV survey can be used to eliminate some responses from the analysis, to give them less weight, or at least

to show that those with the strongest skills tend to provide the most consistent and reliable responses, while those who have the most trouble comprehending small changes in risk rates tend to be behind most of the problem of scope insensitivity.

For example, Hammitt and Graham (1999) asked the following simple probability question to a subsample of their survey respondents:

> Which is a larger chance, 5 in 100,000 or 1 in 10,000? By this I mean which has the greater probability of occurring?

After that they analysed the sensitivity of WTP to risk increments separately for the 61% of the subsample of respondents who answered the question correctly and the 32% who did not.

Krupnick et al. (2002) used simple comprehension questions to help identify respondents who seemed to have trouble understanding the quantitative risk information, finding that a small proportion of their respondents were confused about it or unwilling to put in the effort required, so they were excluded from the analysis (Chestnut and De Civita, 2009).

Andersson and Svensson (2008) tested the cognitive abilities of two hundred Swedish students before they took part in a CVM-study asking them about their WTP to reduce bus-mortality risk. In this context, they found a significant relationship between cognitive skills and sensitivity to the scope of the risk reduction policy. They showed that more cognitively skilled respondents gave answers less affected by scope bias. Additionally, they identified that some parts of the cognitive test based on questions that demanded skills in handling

probabilities, were more strongly linked with scope bias.

Our study included the use within the survey of four questions aimed at trying to proxy the skills of respondents to handle questions about fractions and probabilities, described in Section 10.7. Appendix D shows the results of some preliminary attempts to model the value of the resulting cognitive skill index as a function of respondent characteristics. Additionally, Appendix B shows the results of trying to model the stated levels of self-perceived own risk (both morbidity and mortality risk) as a function of individual characteristics.

Part VI

Data description

Chapter 10

The survey

Our analyses were conducted on a sample of data collected through a phone survey of Newfoundlanders. The survey was administered by a professional market research company (OpinionSearch Inc.) via telephone using random-digit dialing in the island part of the province of Newfoundland and Labrador in the spring of 2013.[1] Residents of Labrador were not approached, since the issue of moose-vehicle collisions is less pressing in that part of the province. After several rounds of refinement, the survey instrument was initially tested by the researchers and a small group of pre-testers. After some further modifications and questions added to the survey, a field pre-test of 150 real respondents was conducted. Due to the lower than average proportion of respondents in the 19-30 age range after the pre-test, the survey was modified to ask to speak

---

[1] A virtually negligible number of calls were made to numbers with Labrador postal codes. There are so few that we kept them in the sample.

with the youngest adult in the household. The bid vector was also adjusted to include further higher bids, since it was found that a relatively high proportion of respondents in the pre-test had accepted the highest ones in the original bid vector.[2] Since no other substantial adjustments were made, the observations from the pre-test sample were appended to those from the final survey and included in the final sample.

A total of 7,599 households were reached (after up to five recall attempts before replacement). Out of the 1,449 respondents who agreed to start answering the survey, those under 19 years of age and those who lived in Newfoundland for less than six months were thanked and dismissed as non-eligible after a couple of brief screening questions or because they would have overfilled the quota aimed for the share of residents in the Avalon Peninsula versus outside the rest of Newfoundland. Several additional respondents chose to terminate the interview after these preliminary questions, resulting in a total number of respondents who completed the questionnaire of n = 150 in the pretest and n = 1,207 in the final survey, adding up to a total of N = 1,357. The response rate was around 19% in both cases (just slightly higher in the final sample, 18.91%, but not statistically different from that in the pretest, 18.66%).

Survey respondents were asked a series of questions about their driving habits, self-rated health status, experiences with moose on the road, and their risk perceptions of a moose-vehicle collision. Additionally, the usual questions

---

[2] To a lesser extent, this problems persisted even after we included additional high bids, so a third round of bid refinement will be applied in an extension of this fieldwork planned for 2014.

on socio-demographic factors such as age, education, gender, and income were asked. In addition to these typical questions asked in a CV survey, this survey included four questions assessing skills in interpreting decimals and fractions and numerical computation. The intent of these questions was to assess how mathematical and computational skills affect various aspects of the WTP distribution. The central questions in our study were, however, the payment questions themselves.

As further described below, we used different versions of the payment scenario. In the versions that asked respondents to consider paying for the provision of a public good, respondents were asked about their annual willingness to pay for a proposed five-year policy involving highway fencing that would reduce the suggested province-wide average baseline risk (which was randomized across respondents) of a MVC in Newfoundland in a certain proportion (which was also randomized across respondents). In the versions involving a private good, the question was about an annual rental payment for a safety device that respondents could install themselves in their cars to reduce their own individual risk of suffering a MVC. In both cases, a double-bound dichotomous choice elicitation format (described in detail in Chapter 7) was used and debriefing follow-up questions were asked in order to obtain a numerical certainty scale (as described in Section 6.2) and in order to be able to assess whether a "no-no" response was the results of a protest response (as explained in Section 11.1). Details of the information obtained and the preliminary manipulation of the data follow.

10.1 Survey versions

Respondents were asked to consider different types of hypothetical MVC risk mitigation strategies. In particular, a subsample of the respondents received a question based on the provision of a private good (a safety device that could be rented yearly and installed in one's car) that could reduce both the mortality and morbidity risks in the case of suffering a MVC (Version A). This device, although it was not specified at all, could be understood as equivalent to an airbag or a safety belt.[3] This is because it would not reduce the risk of a collision but rather the risk of getting injured or dying should a collision occur. Another subsample received a question based on the provision of the same private good that would, instead, reduce only the mortality risk (but not the injury risk) associated with a MVC (Version B). Some other respondents received a question based on the provision of a public good consisting specifically of the erection of fences along highways that would reduce both the mortality and morbidity risk of a MVC (Version C) not only for the respondent but for the general population of the Province. A fourth subsample received a question based on the provision of that same public good, although in this case the type of risk reduction strategy was not specified (*fences* were not mentioned) and the risk reduction would only affect the mortality risk (and not the injury risk) associated with a MVC (Version D). Finally, a fifth subsample was based on questionnaires that combined both types of mortality only risk reduction policies (Version E,

---

[3] A similar approach was used in many previous similar studies (Jones-Lee et al., 1985; Jones-Lee et al., 1995; Johannesson et al., 1996; Hammitt and Graham, 1999; Persson et al., 2001; Hultkrantz et al., 2006; Andersson, 2007; Andersson and Lindberg, 2009)

combining Versions B and D). Additionally, about half of the respondents within Version E were asked the question about the public good first (again with no specific mention of fences but just a reference to a generic strategy to reduce the risk of collisions) and the rest were instead asked first the question about the private good. The distribution of respondents in the main sample according to version is shown in Table 10.1. Only respondents who stated at the beginning of the survey that they regularly drove a vehicle were allocated to the versions that included the private good.

Note that Versions B, D, and E included a valuation scenario based on a policy explicitly involving only a reduction in the mortality risk from MVC, while Versions A and C asked respondents to value a more *comprehensive* good that included reducing the risk of injury from a MVC in exactly the same proportion as the risk of death from a MVC. Therefore, the versions can differ in terms of the nature of the good value (*publicgood* $=0,1$) and whether this good includes reductions in only mortality or also injury risks (*comprehensive* $= 0,1$). Version E combines the public and private good scenarios in a single questionnaire but deals only with reductions in the risk of death.[4] Table 10.1 summarizes the different treatments.

Our exercise takes into account that there are losses associated with a collision that would "be linked to suffering, loss of quality of life, and the pain inflicted on friends and relatives or other individuals in the society" (Dionne

---

[4]We are planning an extension of the survey effort that will include collecting data based on the remaining combination of survey types, namely a Version F that combines public and private good scenarios and a *comprehensive* policy of risk reduction.

Table 10.1: Frequency distribution of respondents by survey version.

	n	public	comprehensive
Version A	271	NO	YES
Version B	271	NO	NO
Version C	275	YES	YES
Version D	270	YES	NO
Version E	270	YES+NO	NO
Total	1357		

and Lanoie, 2004, p. 251) as well as losses imposed by the mere risk of a collision (having to drive more slowly and attentively, avoiding driving at certain times, worrying while driving, etc.). There is clearly no market for these welfare losses. Note, however, that we have abstracted from any consideration of private insurance costs drivers face in the existing auto insurance market to cover the strictly financial and material losses (vehicle damage, loss of income, hospital bills, etc.) associated with a collision. One could expect that a reduced risk of collisions would result in a reduction in insurance premia in the long run.

Our elicitation of WTP for a risk reduction assumes that either those financial costs are fully covered by the drivers' insurance policies or that these fully take them into account when expressing their WTP for a given reduction in risk. In any case, we should expect that the WTP for reducing the risk of a collision will be higher than the WTP for a reduction in the risk of dying should this collision occur. This will then be one of the reasons why we expect the WTP for a given risk reduction level under Versions C and D and under the public component of Version E to exceed the WTP obtained from the other versions.

10.2 Socioeconomic characteristics of Newfoundlanders in the sample

In this section, the five different subsamples are described in terms of a series of socio-demographic characteristics. The first variable we consider, *children-number,* represents the number of members in the household who are under 18, that is, children and teenagers. We might expect that those households with a higher number of young dependants to support would be more willing to pay for the provision of a good that reduced the possible risks of a MVC. As shown in Table 10.2, the number of dependent persons per household is quite small across the different versions, with the average number of children and teenagers in the household well below one. In fact about 65% of our respondents lived in households with no members under 18.

Table 10.2: Average respondent characteristics, by version.

	childrennumber	age	male	Avalon
Version A	0.646	2.595	0.4797	0.5461
Version B	0.641	2.651	0.4982	0.5387
Version C	0.460	2.818	0.4836	0.5164
Version D	0.560	2.833	0.4815	0.5185
Version E	0.617	2.659	0.5074	05370
Total	0.584	2.711	0.4901	0.5313

The variable *age* was, when possible, recorded as a continuous variable. However, since some respondents were not comfortable providing their exact age, we asked them to just place themselves in one of four age intervals instead. Therefore, we also recoded the information on age into an ordered categorical

variable, with respondents belonging to one of the following intervals (in years): [18-29];[30-49]; [50-64];[65- $+\infty$]. Younger respondents ended up being under-sampled relative to their population proportion,[5] because they rarely own land-lines. Therefore, we considered the information on age intervals when designing the sampling weights applied to try and make the sample more representative of the population (Section 11.4).

Furthermore, we used chained-equation imputation techniques to impute the missing values of age in a continuous format, using both the information about the age range provided by the respondent, when available, and the values of other variables. Details about the imputation of missing values for this and other variables are included in Section 11.6.

Table 10.3: Gender distribution by version.

| | Version | | | | | |
	A	B	C	D	E	Total
female	52	50	52	52	49	51
male	48	50	48	48	51	49
Total	100	100	100	100	100	100

Almost half of all the respondents were male, with females only slightly outnumbering them. It is noteworthy that, as shown in Table 10.3, the gender distribution is about the same across versions, suggesting that both women and men are equally likely to drive regularly in the province, since being a regular driver was the only factor that systematically allocated respondents among survey versions.

---

[5]On average, only 6.21% of respondents in the sample fell in the first age category (18-29).

As shown by the mean values of the variable *Avalon*, slightly over half of the respondents in the sample reside within the more urban, more densely populated Avalon Peninsula, while the rest live in the more rural and sparsely populated part of the island of Newfoundland.

Household income before taxes is an ordered categorical variable, with households belonging to one of the following intervals (in Canadian dollars per year): [0-30,000]; [30,000 - 50,000]; [50,000- 70,000]; [70,000 - 90,000]; [90,000 - 110,000]; [110,000 - 130,000]; [130,000 and 150,000]; [150,000 - $+\infty$]. The graphs in Figure 10.1 represent histograms of the income distribution for each version. It can be seen that the distribution differs, as expected, between those versions that value the private good (since they include only those who regularly drive their, presumably in most cases, own vehicle) and those versions that also include non-drivers.

As usual, a great proportion of respondents (slightly over 20% in our case) did not volunteer their income. We, as explained in Section 11.6, imputed the missing values of the income variable in order to be able to use the rest of the information provided by the respondents who did not volunteer their income bracket.

Education is measured as a multinomial categorical variable that reflects different levels and types of education attainment by the respondent (1: Some grade school/high school, 2: High school graduate, 3: Some tech/ vocational/ trade school, 4; Graduate of a tech/ vocational/ trade school, 5: Professional undergraduate degree (nurse, teacher, community college), 6: Some university,
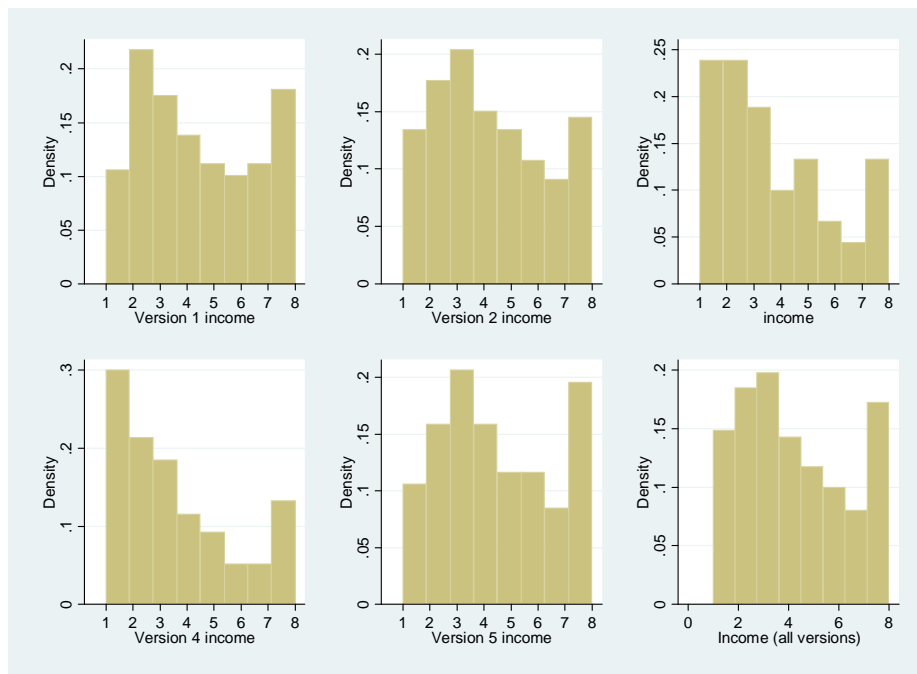
Figure 10.1: Income distribution, by version and in the whole sample.

7: University graduate, and 8: Masters/ doctorate/ professional degree). Although, in strict sense, this variable is nominal, there is a definite ordering behind most of its categories, with the exception perhaps of Categories 4 to 7. Table 10.4 shows the distribution of education levels by version. Over a third of the respondents finished a university degree, which suggest that we might have oversampled relatively educated households. Therefore, we also considered the variable education when designing our sampling weights (Section 11.4).

Since there is only an approximate ordering for the categories of the education variable, we built the binary indicator *college*, which identifies those individuals in categories 5, 7, and 8, which correspond to having finished a

Table 10.4: Distribution of highest levels of education, by version (%).

What is your highest level of education?	Version					
	A	B	C	D	E	Total
Some grade school/high school	7	10	18	18	10	13
High school graduate	13	10	18	17	10	14
Some tech, vocational, or trade school	3	4	4	4	2	3
Graduate of a tech, vocational, or trade school	12	14	10	11	18	13
Professional undergraduate degree	14	16	11	10	16	14
Some university	9	8	10	6	8	8
University graduate	29	29	16	23	25	24
Masters, doctorate, or professional degree	12	9	13	10	11	11
Total	100	100	100	100	100	100

postsecondary education degree or some kind. We found that 49% of our respondents fall in this category.

Only fewer than 1.5% of respondents refused to volunteer their highest level of education. We imputed those missing values (Section 11.6).

Table 10.5: College-educated respondents, by version (%).

	A	B	C	D	E	Total
No	44	46	60	57	48	51
Yes	56	54	40	43	52	49
Total	100	100	100	100	100	100

Although their phone area codes gave us an idea about the location of their current residence, we were interested in learning about the our respondents' place of birth too, so we asked them whether they were originally from Newfoundland or had moved to the province either from another Canadian province or from abroad. Table 10.6 shows that most of the respondents are from Newfoundland, with only a few coming from another province and even fewer from abroad.

Table 10.6: Responses to *What best describes you?*, by version (%).

| What best describes you? | Version | | | | | |
---	A	B	C	D	E	Total
I was born in Newfoundland and have lived here for the last 6 months	89	89	91	92	89	90
I moved to Newfoundland from another Canadian province and have lived here for the last 6 months	7	9	7	5	9	8
I moved to Newfoundland from another country and have lived here for the last 6 months	3	2	1	3	2	2
Total	100	100	100	100	100	100

We also asked how long respondents (even those originally from the province but who moved elsewhere for a period of time) had been residing in Newfoundland. We did not interview anyone who had not lived in the province for at least six months. Table 10.7 summarizes this information by version. In general, it is easy to see that most respondents have been in the province for a long period of time, all their lives in a majority of cases. Around 90% of respondents spent at least the last 10 years in Newfoundland.

Table 10.7: Responses to *How long have you now lived in Newfoundland?*, by version (%).

| Q2. How long have you now lived in Newfoundland? | Version | | | | | |
---	A	B	C	D	E	Total
I moved to Newfoundland less than 5 years ago	3	3	2	0	3	2
I moved to Newfoundland between 5 and 10 years ago	1	2	4	1	3	3
I moved to Newfoundland more than 10 years ago	13	13	10	9	11	11
I have lived here all my life	71	68	70	77	69	71
I used to live elsewhere and moved back here less than 5 years ago	3	3	1	2	3	2
I used to live elsewhere and moved back here between 5 and 10 years ago	3	4	2	4	2	3
I used to live elsewhere and moved back here more than 10 years ago	7	7	10	5	8	8
Total	100	100	100	100	100	100

10.3   Descriptive statistics on risk profiles

Next, we present some descriptive statistics on respondents' characteristics and risk factors associated with MVCs.

Most of the respondents (over 90%) in our sample regularly drive a vehicle, as shown in Table 10.8. Since different types of vehicles could be proxies of attitudes towards road traffic risks and indicators of measures of self-protection, we asked respondents what type they drove most often, finding out that the most common type of vehicle is a SUV-pickup truck,[6] followed by a small-midsize car (Table 10.9). A newer vehicle would also tend to be safer, and we found that, in general, the vehicles owned by the respondents in our sample are relatively new, being 2009-2013 the modal interval (Table 10.10).

Table 10.8: Distribution of frequent drivers, by version.

| | Version | | | | | |
	A	B	C	D	E	Total
Yes	271	271	217	210	270	1,239
No	0	0	57	60	0	117
Total	271	271	274	270	270	1,356

We also asked those respondents who regularly drove a vehicle to tell us about the approximate number of Km they drove per year. We thought it would be useful to have information on who used the roads of the province more, in case that could also explain their WTP for a reduction in the risk of collision with a moose. Although the figures in Table 10.11 show reasonable averages

---

[6]For this reason, we also constructed later a dummy variable SUV that takes the value one if the respondent drives most frequently a SUV and zero otherwise.

Table 10.9: Distribution of type of vehicle, by version (%).

| | Version | | | | | |
	A	B	C	D	E	Total
SUV-pickup truck	44	46	42	38	46	43
Small-midsize car	35	34	35	40	30	35
Full size car	10	13	13	10	13	12
Minivan	11	8	9	11	10	10
Other	0	0	0	1	0	0
Total	100	100	100	100	100	100

Table 10.10: Distribution of vehicle year, by version (%).

| | Version | | | | | |
	A	B	C	D	E	Total
2009-2013	56	60	57	53	57	56
2005-2008	30	28	30	31	26	29
2000-2004	12	10	11	12	13	11
1995-1999	1	1	0	1	1	1
Older than 1995	0	0	0	0	1	1
Don't Know/Refused	1	1	2	3	2	2
Total	100	100	100	100	100	100

for all survey versions for this variable, closer inspection would reveal that several respondents stated an unreasonably high mileage, which we attributed to professional drivers.

Table 10.11: Average number of Km. driven annually, by version

	Version
Version	mean
A	22801.11
B	25277.45
C	25346.26
D	24790.57
E	24205.89
Total	24422.73

We asked respondents to let us know whether they commuted more than 30 Km. for work (Table 10.12). This is because, since most people drive to work in the early morning, and because there is usually relatively little flexibility in the time of day one can do the commute to work, a long commute could be associated with an increased risk of hitting a moose.

Table 10.12: Responses to *Do you drive more than 30 Km to work?* (*drives30towork*), by version (%).

	Version					
	A	B	C	D	E	Total
YES	16	17	14	17	23	17
NO	84	83	84	82	76	82
Don't Know/Refused	0	0	2	1	1	1
Total	100	100	100	100	100	100

The majority of MVC occur between dusk and dawn, as this is the time when driver visibility is reduced by darkness, and when moose are more active. In our sample only a few respondents stated that their job involved frequently

driving at night, as shown by Table 10.13.

Table 10.13: Responses to *Does your job involve frequently driving between 12 midnight and 6 am?* (Variable *job12to6am*), by version (%).

| | Version | | | | | |
	A	B	C	D	E	Total
YES	9	14	13	14	13	12
NO	91	86	87	86	87	88
Total	100	100	100	100	100	100

Personal experience, especially recent events, strongly influences the perception of risk. Therefore, we wanted to include in our analysis some indicator of the degree of experience respondents had with moose on the roads of Newfoundland. As shown in Table 10.14, 85% of the respondents had seen a moose crossing the highway in the previous 3 years. A lower percentage of the sample, but still relevant, about 55%, stated that they had hit a moose or had a near miss (defined as having had to swerve/brake suddenly to avoid the moose) ever (Table 10.15). Most of our respondents also know someone who has hit a moose while driving (Table 10.16).

Table 10.14: Respondents who had seen a moose crossing the highway in Newfoundland in the last 3 years (variable *seenmoosecross*), by version (%).

| | Version | | | | | |
	A	B	C	D	E	Total
YES	91	86	81	82	88	85
NO	9	14	19	18	12	15
Total	100	100	100	100	100	100

Apart from information on these objective stated risk factors, we also wanted to know about the subjective perceptions of risk help by our respondents. This

Table 10.15: Respondents who ever hit a moose (variable *hitmoose*), by version (%).

| | Version | | | | | |
	A	B	C	D	E	Total
YES	61	55	52	49	56	55
NO	39	44	48	51	44	45
Total	100	100	100	100	100	100

Table 10.16: Respondents who know of anyone who ever hit a moose (variable *knowelse*), by version (%).

| | Version | | | | | |
	A	B	C	D	E	Total
YES	78	80	71	74	76	76
NO	22	20	29	26	24	24
Total	100	100	100	100	100	100

is because we assumed that their perception of the baseline risk would matter more than the objective risk when considering their WTP for risk reductions. We expected that they would have a great deal of trouble coming up with a measure of risk by themselves, let alone one that could be reasonably considered comparable across respondents. For this reason, we chose to present them with some objective estimates of 10-year death risks and injury risks associated with general driving and with MVCs in particular in the province. We also provided an explanation of what those measures of risk implied in intuitive terms, using the verbal community analogy[7] approach (Carlsson et al., 2004; Hultkrantz et al., 2006; Andersson and Svensson, 2008; Andersson and Svensson, 2013) with reference to the biggest city in the province, its capital St. John's:[8]

---

[7]See Hammit and Graham (1999) for a comparison of alternative probability analogies in this type of valuation study.

[8]See full text of the payment scenario in the Appendix.

Table 10.17: Summary of risk related variables (%).

Version	drives30towork	job12to6am	seenmoosecross	hitmoose	knowelse
A	16	9	91	61	78
B	17	14	86	55	80
C	14	13	81	52	70
D	17	14	82	49	73
E	23	13	88	56	76
Total	17	12	85	55	75

In parts of Newfoundland drivers often hit moose. Which, apart from the death of the moose themselves, results in car damages, injuries, and, even in some occasions, human deaths. The 10-year average traffic mortality risk in Newfoundland is <RR> in 100,000. So in 10 years in a city like St. John's (with its 200,000 people) one would expect about <RRX2> people to die in car accidents. Similarly, the 10-year average risk of dying from hitting a moose in Newfoundland is <RM> in 100,000. Of course, this risk varies from person to person depending on: where one lives, how much one drives, the type of vehicle one drives, whether one drives at night or not, the type of roads used, and how carefully one drives...

The following questions varied slightly depending on whether the respondent had self-identified as a frequent driver or not during one of the earlier screening questions. Drivers were asked:

Now, considering all this, how high do you think is your own risk of dying in a moose-vehicle collision in the next 10 years? That is, how many times in 100,000 you think your own risk is if the average in

Newfoundland is <RM > in 100,000?

while non-drivers were asked instead:

Now, considering all this, how high do you think is your own risk of

dying **in a car accident involving a moose** in the next 10 years?

That is, how many times in 100,000 you think your own risk is if the

average in Newfoundland is <RM > in 100,000?

Table 10.18: Frequency distribution of "Objective" mortality rates from MVCs (RM) by version (%).

	Version A	Version B	Version C	Version D	Version E	Total
4 in 100,000	16.73	19.48	16.98	17.83	20.53	18.31
6 in 100,000	19.70	22.85	16.60	24.81	17.87	20.35
8 in 100,000	23.79	22.47	22.26	16.67	19.77	21.03
10 in 100,000	15.99	14.61	18.49	20.54	19.01	17.70
12 in 100,000	23.79	20.60	25.66	20.16	22.81	22.62
Total	100.00	100.00	100.00	100.00	100.00	100.00

Therefore, after describing to respondents the objective 10-year death and injury risk rates (per 100,000) associated with both road accidents (RR) and accidents involving MVC in Newfoundland (RM and RI), we asked them to estimate their own risk rates. However, when describing the road traffic risks in the province as part of the payment scenario, we, while pretending to be quoting true official statistics, actually randomized across respondents for RR four plausible values around the true estimate of about 85-100 in 100,000 (for 10 years)[9] and five plausible values for RM close to the true value of around 6

---

[9] We considered the figures published in StatsCanada (2009) and the equivalent tables for the years 2006 and 2008.

per 100,000 in 10 years (Guo, 2011). Having several values purported to be the true baseline risk would allow us to test for anchoring effects on the respondents estimation of their own risk and also to use baseline risk (or a transformation thereof) as an independent variable in the WTP models. The distribution of these values is shown in Table 10.18.

The objective injury (morbidity) rates (RI) where simply calculated as 30 times the respective death rates. The variable RM includes these objective values of the death rate due to MVC in Newfoundland and variable RI includes its objective injury rate counterpart. We used these values as baseline risks for the public good questions, since a general reduction in the risks of accidents across the province because of the implementation of the fence building program would reduce the risk on average for all the residents in the province. However, we used the perceived own risks (from Questions $Q12$ and $Q13$, whose full text is available in Appendix E) as the baseline risk rates for the private good (labelled *baseline* and *baselineI*), since the in-vehicle safety device was described as reducing only the personal risks for the driver. When respondents failed to provide an estimate of their own mortality risk,[10] knowing that we would need that information for the construction of the baseline and associated measures of scope of risk reduction, we ensured that they were assigned one of the public good versions (Version C or Version D). Therefore, no missing values are found in Versions A, B, and E for Q12 or Q13.

The public good versions do have missing values for the respondent's own

---

[10] Or provided a perceived risk of less than 0.5 in 100,000, usually exactly zero, since we deemed those respondents as unable to fully grasp the concept of the risk measure provided.

176

estimate of the death risk (about 23% of cases) or injury risk (slightly over 21% of cases) rates. Variables *baseline* (which refers to mortality risks) and *baselineI* (which refers to injury risks) are then based on the randomised values of RM for the cases where *publicgood* $= 1$ and the cases where $Q12$ is missing and the randomised values of RI (which equals 30 times RM) for the cases where *publicgood* $= 1$ and the cases where $Q13$ is missing.

It should be noted that we found several very implausibly high values[11] given for $Q12$ ad $Q13$, which, as explained in Chapter 11, led us to eliminate some observations from the analysis, namely those 75 cases whose $Q12$ exceeded 100 in 100,000 (which would imply that about 50 people would be killed on average my MVCs every year in the province) and/or those 72 cases for which, equivalently, $Q13$ exceeded 3000 in 100,000 (implying that every year moose would injure 1,500 people in the province).[12] In a total of 42 cases, $Q12$ and $Q13$ both took extreme values for the the same respondent, so overall only 105 cases (instead of 147) were eliminated because of their having outlier values for perceived baseline risks. Tables 10.19 and 10.20 in the next section, show the effect of the wildly exaggerated estimates we eliminated from the original sample of the mean values of the relevant variables.

---

[11] This made, for example, the mean of $Q12$ exceed 2000 per 100,000 while the median was, much more reasonably, equal to only 5.

[12] An reasonable estimate would be about 100 for general injuries and around 10-15 for major injuries.

10.4 Measures of scope and scale

As noted in Footnote 14 in Section 4.2.9, *scale* and *scope* are used interchangeably to refer to the size of the risk reduction. We will only be carefully specific when using *scope* or *scale* in a sense different from the size of risk reduction. Otherwise, we will use them interchangeably.

Table 10.19: Measures of self-perceived mortality and morbidity baseline risks, proportions of risk reduction in the policy scenario, and resulting absolute magnitudes of mortality and morbidity risk reductions, by version. Full sample.

Version		Q12	Q13	MULTI	diffM	diffI
A	mean	2404.60	1879.02	3.03	1490.86	1164.47
	median	5.00	50.00	3.00	3.00	33.33
	st. dev.	12224.42	10249.68	0.82	7337.91	5954.81
B	mean	2413.70	1638.31	3.03	2290.26	0.00
	median	4.00	50.00	3.00	2.50	0.00
	st. dev.	13191.07	9521.32	0.80	12737.10	0.00
C	mean	3222.08	3834.61	2.91	5.28	158.30
	median	5.00	40.00	3.00	5.00	150.00
	st. dev.	14683.49	15908.14	0.82	2.06	61.73
D	mean	1581.79	4432.55	3.02	5.17	0.00
	median	4.00	22.50	3.00	5.00	0.00
	st. dev.	10850.10	19088.48	0.79	2.05	0.00
E	mean	989.84	2360.26	3.07	33.15	0.00
	median	4.00	50.00	3.00	4.00	0.00
	st. dev.	8553.85	13345.42	0.82	412.71	0.00
Total	mean	2034.89	2582.80	3.01	774.58	268.68
	median	4.00	50.00	3.00	4.00	0.00
	st. dev.	11807.22	13321.66	0.81	6673.68	2720.96

We used variables *Q12* and *Q13* to build the measures of scope of the risk reduction policy by suggesting in the payment scenarios that either the safety device or the fence installation policy would reduce the baseline risk in a certain proportion *MULTI* (taking the values 2, 3, and 4, randomized across respondents). Therefore, depending on the size of the baseline risk and the value of

Table 10.20: Measures of self-perceived mortality and morbidity baseline risks, proportion of risk reduction in the policy scenario, and resulting absolute magnitudes of mortality and morbidity risks, by version. Trimmed sample.

Version		Q12	Q13	MULTI	diffM	diffI
A	mean	6.73	101.93	3.03	4.31	65.89
	median	4.00	38.00	3.00	2.59	23.25
	st. dev.	9.97	147.24	0.83	6.42	96.46
B	mean	5.83	93.33	3.02	3.78	0.00
	median	3.00	50.00	3.00	2.00	0.00
	st. dev.	9.71	139.47	0.78	6.59	0.00
C	mean	7.20	72.80	2.88	5.36	160.80
	median	4.00	20.00	3.00	5.00	150.00
	st. dev.	10.70	104.00	0.84	2.13	63.80
D	mean	8.16	89.94	3.02	5.20	0.00
	median	4.00	20.00	3.00	5.17	0.00
	st. dev.	16.32	180.95	0.75	1.92	0.00
E	mean	4.86	98.20	3.09	4.14	0.00
	median	3.00	50.00	3.00	3.75	0.00
	st. dev.	5.37	137.96	0.82	3.18	0.00
Total	mean	6.20	94.75	3.03	4.32	34.32
	median	4.00	35.00	3.00	3.00	0.00
	st. dev.	9.93	143.09	0.81	5.15	74.82

MULTI, each respondent ended up with a different value for the *absolute* reduction in risk assigned to their payment scenario. Some respondents were asked to value a *comprehensive* policy scenario promising both an absolute reduction in the mortality risk, measured by the variable *diffM*, and an absolute reduction in morbidity risks, measured by *diffI*. It must be stressed, though, that the scenario description only mentioned the baseline risk and the final lower risk level. Variables *diffM* and *diffI* were constructed from the data as the difference between these two measures.

Since, as explained above, and more in detail in Chapter 11, we eliminated observations with implausibly high values of *Q12* and *Q13*, we show in Tables 10.19 and 10.20 the summary descriptives of *diffM*, *diffI*, *MULTI*, *Q12* and *Q13* both before and after the elimination of these problematic cases. It can be seen in these tables that, in the versions including only public good valuations, there are no issues related to extremely high levels of risk, because self reported risk levels were not applied to those versions. Note also that the medians are always reasonable as well.

10.5 Willingness to pay and different proposed policies

We used a double-bounded dichotomous choice format in our willingness to pay question. Respondents initially faced a randomly selected bid ranging from $15 to $150. If the first payment question was answered positively, the bid amount

was doubled for the follow-up, while it was halved if the first response was negative.

The full text of the with a budget reminder and the payment question read as follows in the case of the public good versions. For Version C:

> We now would like to know if you would support a program aimed at reducing the risk of moose-vehicle collisions in Newfoundland by installing and maintaining fences along the highways, together with under and over passes for the moose to cross the road safely. There are many good reasons why one is willing to pay or not. Before answering the question – we would like to remind you that there may be other causes to support, including programs aimed at promoting health and safety in other ways, and we would also like to remind you that supporting the program would mean having less money for other personal expenditures such as rent, food, gas, and so on.
>
> Imagine a [randomized *federal/provincial*] fencing program that would run for five years. If carried out the program would reduce the mortality risk from a MVC (Moose-Vehicle Collision) for the general population of the province from RM to $(RM/MULTI)$[13] in 100,000. Note that the risk of injury to you, your passengers or other drivers would be now also proportionally less as well as the effects of injuries to the moose. In particular, the 10-year injury

---

[13] With RM taking a random value from the vector [4,6,8,10,12] and *MULTI* taking the value of either 2, 3, or 4, as explained above.

risk from moose-vehicle accidents would go down from from *RI* to (*RI/MULTI)* in 100,000. The funds for this would come from an increase in [randomized: *annual federal taxes/annual driver license fees*]. The extra [*taxes/fee*, depending on previous sentence] your household would have to pay would be $*firstbid* (randomly assigned from the vector [15, 30, 45, 60, 75, 100, 150] per year for the five years of the program if it went ahead.

If in trying to decide about the program the [as above: *federal/provincial*] government conducted a referendum. [so if a majority of Newfoundlanders were in support of the policy, this would go ahead; otherwise there would be no policy and no increased tax/license fee][14] Would you vote yes in this referendum?

For Versions D, and E, since they involved only mortality risk reductions (and not morbidity risk reductions), we thought it would be more plausible for the respondents not to hear any mention of the specific type of risk reduction strategy, so the payment scenario read simply:

We now would like to know if you would support a public policy aimed at reducing the general mortality risk of car drivers by reducing the risk of moose-vehicle collisions in Newfoundland. There are many good reasons why one is willing to pay or not. Before answering the question, we would like to remind you that there may

---

[14]This sentence in square brackets is a randomized provision constraint mechanism applied to 50% of respondents in Versions C, D, and E. This resulted in variable *referendumreminder*.

be other causes to support, including programs aimed at promoting health and safety in other ways, and we would also like to remind you that supporting the program would mean having less money for other expenditures such as rent, food, gas, and so on.

Imagine a [randomized *federal/provincial*] program that would be applied during five years. If carried out, the program would reduce the mortality risk from a MVC (Moose-Vehicle Collision) for the general population of the province from *RM* to $(RM/MULTI)^{15}$ in 100,000. Although this is a bit unrealistic, please assume that the risk of injury to you or other drivers, or to any passengers, would remain the same and that the risk of injuries to the moose would also remain unchanged with the program. In other words, the policy would only reduce your mortality risk, nothing else. The funds for this would come from an increase in [randomized: *annual federal taxes/annual driver license fees*]. The extra [*taxes/fee*, depending on previous sentence] your household would have to pay would be $*firstbid* (randomly assigned from the vector [15, 30, 45, 60, 75, 100, 150] per year for the five years if the program went ahead.

If in trying to decide about that program the [as above: *federal/provincial*] government conducted a referendum. [so if a majority of Newfoundlanders were in support of the policy, this would go ahead; otherwise there would be no policy and no increased

---

[15] With RM taking a random value from the vector [4,6,8,10,12] and *MULTI* taking the value of either 2, 3, or 4, as explained above.

tax/license fee][16] Would you be willing to pay \$[*firstbid*] for this

program?"

The differences between Version D and the public good question in version E versus Version C would be only that version C also involved a more comprehensive policy, since it included a reduction in the corresponding injury risk for the residents of the province and that Versions D and E made no mention of the specific type of public program aimed at reducing the risk of dying from a collision with a moose. The full text of all the versions is included in the Appendix E.

The full text of the payment question read as follows in the case of the private good versions (Versions A, B, and one of the questions in Version E):

We now have some questions about your willingness to pay for increased traffic safety. Imagine that you are offered a new safety device that is not inconvenient, ugly, or complicated to use. In fact, you would not notice it. It reduces only your own mortality risk from its current level of *Q12*[17] down to *Q12/MULTI*[18] in 100,000 should you be involved in a moose vehicle collision. It is only you as a driver who can personally benefit from it by reducing the risk of dying but only from hitting a moose: it does not help reduce the risk of you dying in other types of car accidents or your risk of

---

[16] This sentence in square brackets is a randomized provision constraint mechanism applied to 50% of respondents in Versions C, D, and E. This resulted in variable *referendumreminder*.

[17] Or RM if Q12 was missing.

[18] With RM taking a random value from the vector [4,6,8,10,12] and *MULTI* taking the value of either 2, 3, or 4, as explained above.

getting injured; it does not protect your passengers, or other drivers, or the moose, or the vehicle and you could not lend it to anyone, even in your household. Assume that this personal device can be used in any of the vehicles you drive. Assume that its effect lasts only one year, so, after that, you must make another payment if you want to continue the risk reduction. Remember that there would be other ways to improve your safety and that paying for this device would mean having less money for other personal expenditures such as rent, food, gas, and so on.[19] Would you be willing to pay $*firstbid* (randomly assigned from the vector [15, 30, 45, 60, 75, 100] per year for this device?

The difference between Versions A and B and the question concerning a private good within Version E would be only that version B also involved a more comprehensive policy, since it included a reduction in the corresponding injury risk faced by the driver. The full text of all the versions is included in the Appendix E. Version E included both the public good policy and the private good question in randomised orders.[20]

In the following tables, the frequency distribution of the bids actually used across versions is presented. Cross-tabulations by version of response frequency and bid values can be seen in Table 10.22. However, the notion of using a double-bounded dichotomous choice approach makes the information in Table

---

[19] The interviewer, if asked, stressed that the level of risk would revert to *Q12* (or *RM* is *Q12* was missing) in 100,000 after discontinuing the use of the device.

[20] Those who received the private good question first were assigned a value of one for variable *privatefirst*.

Table 10.21: Frequency distribution of variable *COST* by version (%).

| | Version | | | | | |
	A	B	C	D	E	Total
$15	17.47	12.36	11.70	10.85	9.13	12.33
$30	10.04	14.61	12.45	8.53	13.69	11.88
$45	11.90	11.99	10.19	10.85	14.07	11.80
$60	16.36	13.11	12.45	14.73	16.35	14.60
$75	11.90	16.10	14.34	12.02	11.79	13.24
$100	12.64	12.73	13.21	15.50	14.83	13.77
$120	10.04	8.24	12.83	15.50	11.03	11.50
$150	9.67	10.86	12.83	12.02	9.13	10.89
Total	100.00	100.00	100.00	100.00	100.00	100.00

10.23 likely more informative. We can see that, as expected, the combination NO-NO was more frequently observed at higher levels of the initial bid proposed, while the pattern for YES-YES is more or less the inverse. These patterns are not fully monotonic but the general tendency to reject the bid when its size is larger is rather clear in these type of responses. Also, as expected, almost no patterns can be found with respect to the bid value in the YES-NO and NO-YES responses. However we can comment on one additional aspect of Table 10.23. As explained in Section 7.3, with proper bid design and under no undue influences of the first response or the first bid on the second response, the DBDC question format, which estimates the location of the respondents' WTP by attempting to bracket it with the two bids, should result in an approximate equal distribution of responses across the four combinations shown in Table 10.23. However, it can be seen that the combinations YES-YES and NO-NO are much more prevalent than the combinations NO-YES and YES-NO. As pointed out by Cameron and Quiggin (1994), this tendency of the respondents to provide consistent responses

is an indicator of *anchoring*, an issue that we described in Section 7.

Table 10.22: Willingness to pay bid ($ per year)(%) in response to initial payment question.

Version	Willing to pay (%)	Bid value ($/year)								Total
		15	30	45	60	75	100	120	150	
A	Yes	68	63	53	52	48	37	41	35	51
	No	32	37	47	45	48	60	59	65	48
	Don't Know/Refused	0	0	0	2	3	3	0	0	1
B	Yes	56	49	41	39	39	40	32	52	44
	No	41	49	59	58	59	57	68	48	55
	Don't Know/Refused	3	3	0	3	2	3	0	0	2
C	Yes	68	68	38	54	51	47	56	40	53
	No	29	29	52	37	36	39	38	54	39
	Don't Know/Refused	3	3	10	9	13	13	6	6	8
D	Yes	32	54	55	42	50	55	47	44	47
	No	57	38	35	55	35	45	47	47	45
	Don't Know/Refused	11	8	10	3	15	0	7	9	7
E1 private good	Yes	47	62	44	44	29	50	29	27	43
	No	53	38	56	56	71	44	71	73	56
	Don't Know/Refused	0	0	0	0	0	6	0	0	1
E1 public good	Yes	67	62	56	44	50	33	36	27	48
	No	27	33	33	52	50	61	57	64	46
	Don't Know/Refused	7	5	11	4	0	6	7	9	6
E2 public good	Yes	70	40	47	41	41	43	44	7	41
	No	20	60	53	47	53	43	50	73	51
	Don't Know/Refused	10	0	0	12	6	13	6	20	8
E2 private good	Yes	70	47	26	29	12	52	19	27	34
	No	30	53	74	65	88	48	81	73	65
	Don't Know/Refused	0	0	0	6	0	0	0	0	1

In order to allow the possibility of using a numerical certainty scale for the purpose of investigating its potential for reducing hypothetical bias, for improving estimation efficiency, for alleviating inconsistencies between responses to double-bound dichotomous choice questions, and for ameliorating issues of insensitivity to scope in exercises dealing with the valuation of risk reductions, respondents were asked to rank on a scale of one to ten how sure they were

Table 10.23: Frequency distribution of response patterns by initial bid (CAD $), in %.

	No-No	No-Yes	Yes-No	Yes-Yes	Total
$15	34.78	5.59	12.42	47.20	100.00
$30	31.37	9.80	19.61	39.22	100.00
$37.5	0.00	0.00	0.00	100.00	100.00
$45	44.14	5.52	22.07	28.28	100.00
$50	0.00	0.00	0.00	100.00	100.00
$60	41.58	10.53	23.16	24.74	100.00
$75	44.85	6.06	18.79	30.30	100.00
$100	44.77	8.14	16.86	30.23	100.00
$120	45.77	9.15	19.72	25.35	100.00
$150	50.75	8.21	20.15	20.90	100.00
$300	100.00	0.00	0.00	0.00	100.00
Total	42.06	7.91	19.05	30.99	100.00

of their responses to the WTP questions. The distribution of values of this numerical certainty scale is shown in Table 10.24.

10.6 Additional variables about the payment scenario characteristics

In the case of the payment scenario based on a public good we split the sample 50-50 using the a control for the agency in charge of implementing the policy. In the end, our usable observations reflected that pattern too. The variable *federal* identifies with the value 1 the observations for which the agency in charge of the policy was supposed to be the Canadian Federal Government, while a zero indicates that the provincial government was mentioned. In the first case, the payment vehicle used was an increase in annual income taxes and in the second

Table 10.24: Frequency distribution of numerical certainty scale (how sure respondents are of their responses to the WTP questions), by version.

	Version					
	A	B	C	D	E	Total
1 - Not very sure	3	6	5	10	6	6
2	1	2	3	1	2	2
3	4	2	4	2	2	3
4	4	3	1	0	3	2
5	15	15	9	16	10	13
6	3	5	5	5	5	5
7	8	7	5	6	9	7
8	15	16	11	13	13	14
9	6	7	5	3	7	6
10 - Very sure	40	37	45	41	42	41
Don't Know/Refused	1	1	5	3	1	2
Total	100	100	100	100	100	100

case a surcharge on the annual fees paid for a drivers' licence.

The two values of the variable *referendumreminder,* which indicates that a reminder of a provision rule whereby the risk reduction policy would only be effected if a majority of respondents to a referendum voted in favour of the policy, was presented to the respondent, were also distributed equally among those respondents who receive the public policy version of the questionnaire.

10.7   Additional variables about respondent characteristics

Since there are several theoretical reasons why the WTP for a reduction in the risk of dying could be affected by the health status of the respondent, we wanted to have some indication of how healthy respondents perceived themselves

to be. The self-perceived *health* status of the respondent is based on a bounded variable, i.e., respondents were asked to rank their health on a 100-point scale, using the question:

> How healthy would you say you are in general? Consider a 1 to 100
> scale where 1 means very sick and 100 is perfectly healthy. How
> much do you think you would currently score?

Figure 10.2 shows the distribution of this variable in our sample.
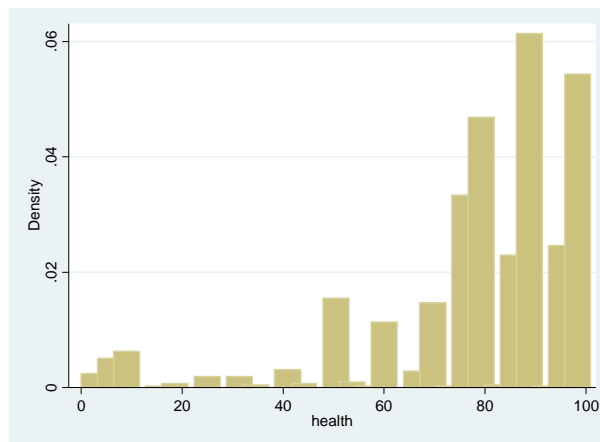
Figure 10.2: Distribution of values of self-reported overall level of health.

We also wanted to know about any feelings the respondents might have towards alternative solutions to the MVC problem, particularly those associated with culling them or increasing the number of hunting licenses. Therefore, we asked them whether they hunted moose in the past five years (which resulted in the binary variable *huntedmoose*) and also whether they consumed moose meat in the past year (variable *atemoose*).

In order to obtain a further proxy for the respondents' attitudes towards risk, we asked them:

Have you smoked more than 10 cigarettes in the past 10 days?

In order to test whether it had any influence in their attitudes towards risk and in their ability to handle questions about small risks we also asked about the respondents month of birth.

Finally, the variable *mathscore* was based on a series of four questions that were asked at the end of the survey. The questions assessed math skills such as numerical computation and understanding small fractions and decimals, both of which are skills used when assessing various mortality risk scenarios. The value of *mathscore* was simply the sum of correctly answered questions by the respondent.[21]

Table 10.25 includes summary statistics by version of these additional covariates.

---

[21]Krupnick et al. (2002) identified, using several simple comprehension questions, a small fraction of respondents who seemed unable to understand the quantitative risk information and excluded their responses from the analysis.

Table 10.25: Summary descriptives for additional variables, by version.

Version	Statistic	health	smoker	monthofbirth	atemoose	huntedmoose	mathscore
A	mean	79.19	0.16	6.49	0.67	0.32	1.62
	median	85.00	0.00	6.00	1.00	0.00	2.00
	std. dev.	20.08	0.36	3.36	0.47	0.47	0.95
	minimum	1.00	0.00	1.00	0.00	0.00	0.00
	maximum	100.00	1.00	12.00	1.00	1.00	4.00
B	mean	80.23	0.16	6.50	0.66	0.28	1.59
	median	85.00	0.00	7.00	1.00	0.00	2.00
	std. dev.	19.72	0.37	3.22	0.47	0.45	0.98
	minimum	7.00	0.00	1.00	0.00	0.00	0.00
	maximum	100.00	1.00	12.00	1.00	1.00	4.00
C	mean	78.25	0.21	6.42	0.67	0.32	1.57
	median	80.00	0.00	6.00	1.00	0.00	2.00
	std. dev.	23.01	0.41	3.21	0.47	0.47	0.85
	minimum	1.00	0.00	2.00	0.00	0.00	0.00
	maximum	100.00	1.00	12.00	1.00	1.00	3.00
D	mean	70.49	0.19	5.79	0.56	0.19	1.62
	median	80.00	0.00	5.00	1.00	0.00	2.00
	std. dev.	28.90	0.39	3.59	0.50	0.39	0.86
	minimum	3.00	0.00	1.00	0.00	0.00	0.00
	maximum	100.00	1.00	12.00	1.00	1.00	4.00
E	mean	80.13	0.13	6.82	0.71	0.31	1.73
	median	85.00	0.00	7.00	1.00	0.00	2.00
	std. dev.	19.24	0.34	3.33	0.46	0.47	0.92
	minimum	1.00	0.00	1.00	0.00	0.00	0.00
	maximum	100.00	1.00	12.00	1.00	1.00	4.00
Total	mean	78.72	0.16	6.50	0.67	0.29	1.64
	median	85.00	0.00	6.00	1.00	0.00	2.00
	std. dev.	21.27	0.37	3.33	0.47	0.46	0.93
	minimum	1.00	0.00	1.00	0.00	0.00	0.00
	maximum	100.00	1.00	12.00	1.00	1.00	4.00

Chapter 11

Data screening and

manipulations

11.1 Treatment of 'don't know' responses

As stated before, we used the double-bounded dichotomous choice method (DBDC), so most respondents were asked two WTP questions that used bid values taken from three vectors given by the variables *COST*, *COSTH*, and *COST2*, as shown in the full text of the questionnaire (available in Appendix E). In order to foster consistency with the notation in most of the literature, let us label *COST* as B, *COSTH*, which just takes one half of the values in *COST*, Bl, and *COST2*, which contains the doubled values of *COST*, Bh.

First, each respondent was randomly assigned an initial bid B (with values
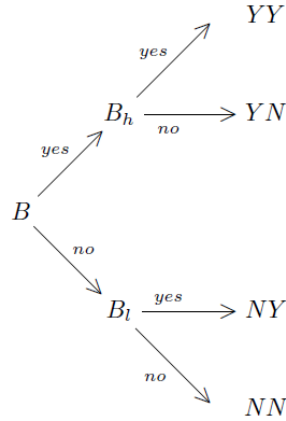
Figure 11.1: Diagram for YES and NO responses to initial question.

ranging from \$15 to \$150). If the answer was positive for that bid, the respondent was then assigned a doubled bid Bh (with values ranging from \$30 to \$300) for the follow-up question [and the response to a lower, halved, bid Bl is implicitly assumed to be also positive]. If the answer was negative for the initial bid B, the respondent is thus determined to follow the descending sequence and the response to the higher bid Bh is implicitly assumed to be negative. The follow-up question used in this second case on the halved bid Bl (with values ranging from \$7.5 to \$75). In sum, most respondents face two bids, B and either Bh or Bl, depending on their response to B. However, the underlying approach involved assigning three bids to each respondent, even if in most cases one of them was never presented to them.

Unless the response to the follow-up WTP question was a "don't know"/no-response (DK), the values of a polytomous variable can be worked out from the

two responses that trace the type of willingness to pay sequence: NN, NY, YN and NN. It is also clear that most of the information about the respondent's actual WTP is obtained from the observations where it is doubly bracketed (those with sequence NY and YN). Figure 11.1 illustrates diagrammatically the elicitation process.

A third possibility, though, is that the respondent replied "don't know" to the initial question based on bid B or failed to provide a response. A total of 71 cases fall in this category. In this case, the respondent was then asked both the higher and the lower bid in a random order which, due to a flaw in the questionnaire programming design,[1] we can no longer determine. However, if both Bh and Bl obtained positive responses, we can infer that the respondent was asked first about Bl and then about Bh (since the reverse order would have made the second question redundant) and the logical order would have been the opposite if the responses are both negative. However, it is impossible to find out the order in which the bids (Bh and Bl) were presented when the answers are positive and negative or *vice versa*. Therefore, we omitted those (35) observations from the analysis, since knowing the order in which responses were given is essential for most analyses involving DBDC data. Note also that in the few cases where the first response was a DK, the "firstbid" could in principle have been higher than $150 (one case) and lower than $15 (no actual cases, though). Similarly, the most infrequent cases of firstbid = $22.5 (one case) and firstbid = $37.5 (one case) are the result of this treatment of the

---

[1]This issue has been resolved for the second wave of the data collection process.

DKs. Additionally, we eliminated the cases with remaining DKs as responses to questions other than the very first one, since they would not yield enough information for us to use DBDC models.[2] Abstracting from these DK responses to questions other than the first one, Figure 11.2 illustrates diagrammatically the elicitation process when the first response was a DK.
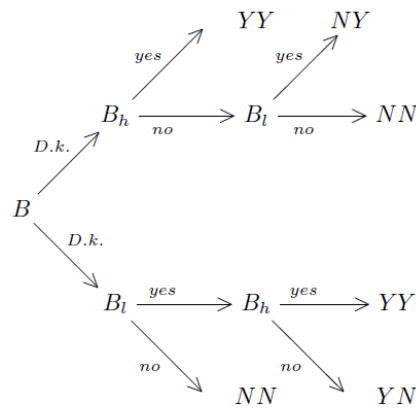
Figure 11.2: Diagram for "Don't know" responses to initial questions.

Additionally, and as noted in Section 10.3, respondents who did not provide an estimate of own mortality risk,[3] because we would need that information for the construction of the baseline risk measure and associated measures of scope of risk reduction, were assigned one of the public good versions of the questionnaire (Version C or Version D). Therefore, there were no missing values in Versions A, B and E for Q12 or Q13. The public good versions of the survey do have

---

[2] 95 observations originally had a DK as the response to the initial question and 44 in response to the follow-up question.

[3] Or provided a perceived risk of less than 0.5 in 100,000, usually exactly zero, since we deemed those respondents as unable to fully grasp the concept of the risk measure provided.
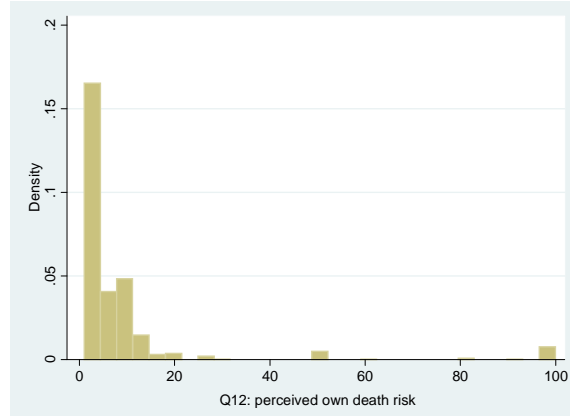
Figure 11.3: Perceived own death risk rate due to MVCs (trimmed at 100 in 100,000).

missing values for the respondent's own estimate of the death risk (about 23% of cases) or injury risk (slightly over 21% of cases) rates.

11.2 Baseline risk outliers

Given the hypothetical nature of the survey, we must be mindful that not all responses are going to convey meaningful information and that some respondents will have responded to the questionnaire having explored more information, more carefully, more reliably, and/or more thoughtfully. We have a few ways to at least try to ameliorate this problem *ex post* based on the answers to the questionnaire. In particular, we worry about a pervasive problem in the literature that deals with the valuation of risk reductions, namely that typical respondents have quite a lot of difficulty understanding differences between small risk levels, as explained in further detail in Chapter 8.
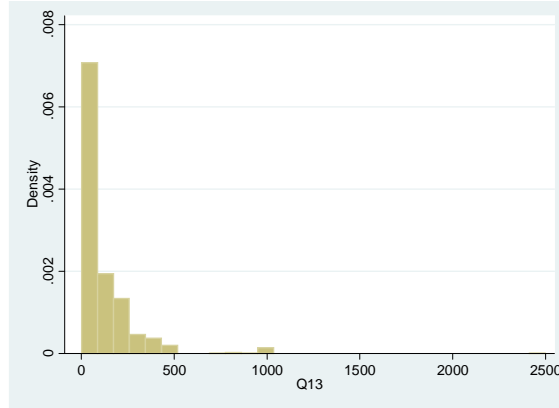
Figure 11.4: Perceived own injury risk rates due to moose-vehicle collisions (trimmed at 3,000 in 100,000).

We chose to discard from the analysis the information from those observations for which the stated mortality risk perception was greater than 100 out of 100,000 (with a few respondents stating levels in the tens of thousands) or for which the stated level of own injury risk was higher than 3000. The relation between these two limit points was equal to the 30 to 1 ratio we used when describing generic averages of risk of mortality and morbidity in the province. We thought that using information from these respondents (about 6% of the whole sample) who stated values orders of magnitude larger than the average risk and so much higher than the proposed levels of general risk would result in implausible levels of scale of risk reduction and that these respondents seemed unable to fully interpret the valuation questions and to place a meaningful value upon small changes in risk reductions.

The mean and median values of the self-perceived death risk (variable *Q12*)

after the trimming were 6.09 in 100,000 and 4 in 100,000, respectively, and the distibution of these values can be seen on the histogram in Figure 11.3. Their injury risk rate counterparts are 105.18 in 100,000 (mean) and 50 in 100,000 (median). More detailed summary statistics by version can also be found in Tables 10.19 and 10.20 in Section 10.3.

11.3   Protest responses

As explained in Section 6.4, the literature tends to agree that, although there is no established theoretical basis for excluding protest responses, most often the elimination of these responses from the sample is the strategy of choice (Morrison et al., 2000; Boardman et al., 2001), together with the reporting welfare estimates obtained with and without the protest responses. Eliminating protest responses will increase the welfare estimate, while keeping the protests will yield more conservative estimate.

In our case, responses reflecting the pattern NN[4] were further screened for protest responses in our survey. We included a debriefing question asking why these respondents would not be willing to pay any of the bids presented to them. We allowed each respondent to provide as many reasons as they wanted and we had the interviewer classify each of these into a series of 15 predetermined categories. In addition, the interviewer recorded verbatim any "other" reason that did not immediately fall into any of these categories. After that, we manu-

---

[4]Or NN NN in the case of version E, since those respondents were asked at least four questions, or DK-NN in any of the versions.

ally classified ourselves some of these other reasons into the categories *protest* or *protestclear*. As the labels of these variables suggest, the former includes cases where there might be some doubt as to whether the response is a protest and not genuine negative to pay, while the former includes only the cases most researchers would comfortable classify as protests.

Among the responses to the debriefing question (whose full text is available in Appendix E), the following are the types of responses that were identified as definitively protest responses in our dataset (identified by the binary variable *protestclear*):

- "I don't believe the money would be spent on that" (01);

- "I would not trust the government to do the job properly" (02)

- "It should not be financed through taxes/not everyone should have to pay their share to protect drivers" (05)

- "I should not have to pay individually: the province/government should pay for that without raising taxes" (07)

- "I do not believe that the program would be effective"(11)

- "The drivers should pay for that themselves" (12)

- "The drivers' insurance should pay for that" (13)

- "The government should fund the program with existing revenues, and not ask for additional taxes"   (14)

- "Brush should be trimmed from roadsides to enable visibility" (22)

- "Should not have to pay for the poor habits of other drivers" (32)

while, in addition, the following reasons made us classify the case as a mere

protest:

- "Moose population should be decreased/culled" (26)

- "MVC prevention should focus on driver safety/awareness" (27)

- "Need more information/proof/evidence of effectiveness" (28)

- "The problem exists because moose are not native to the area" (31)

Table 11.1: Protest responses, by version (%).

	Version A	Version B	Version C	Version D	Version E	Full sample
genuine no	84	84	80	73	78	80
protest no	16	16	20	27	22	20
Total	100	100	100	100	100	100

Table 11.2: "Clear" protest responses, by version (%).

	Version A	Version B	Version C	Version D	Version E	Full sample
genuine no	90	88	82	78	83	84
protest no	10	12	18	22	17	16
Total	100	100	100	100	100	100

11.4   Sampling weights

Our sampling strategy overestimated slightly respondents from outside the Avalon

Peninsula, since we wanted to make sure we had enough representation from the

Table 11.3: Frequency distribution of reasons to reject both bids proposed within the payment question.

variable	N
I don't believe the money would be spent on that	17
I would not trust the government to do the job properly	7
Too expensive/I cannot afford that	55
I already pay too much tax	37
It should not be financed through taxes/not everyone should have to pay their share to protect drivers	212
I already contribute to other environmental programs/causes	1
I should not have to pay individually: the province/government should pay for that without raising taxes	222
I do not care about MVC	13
I do not drive	9
Drivers should just slow down	96
I do not believe that the program would be effective	88
The drivers should pay for that themselves	9
The drivers' insurance should pay for that	5
The government should fund the program with existing revenues, and not ask for additional taxes	235
The government has other higher priorities for spending taxpayer's money	13
Not necessary/MVC risk is low/waste of money (unspecified)	43
I am a careful driver/am not worried about hitting moose	26
Brush should be trimmed from roadsides to enable visibility	6
Device protects only drivers/not the car/other passengers	46
I don't/rarely drive at night/these accidents occur at night	14
I don't/rarely drive on the highway	13
Moose population should be decreased/culled	8
MVC prevention should focus on driver safety/awareness	34
Need more information/proof/evidence of effectiveness	33
There are too few/no moose in my area to worry about it	13
I do not drive enough in moose-populated areas	6
The problem exists because moose are not native to the area	6
Should not have to pay for the poor habits of other drivers	4
Other	42

least populated parts of insular Newfoundland. Additionally, we compared the frequency distributions of our sample with official statstics for the insular part of Newfoundland and Labrador and noticed slight discrepancies in terms of several socio-demographic components. Therefore, we constructed sampling weights to account for the overrepresentation of some individuals and used those weights in our regressions. We thus report standard errors based on these sampling weights.

11.5   Uncertain responses

We collected information about how sure respondents were of the answers they provided to the payment questions. The distribution of values of this numerical certainty scale is shown in Table 10.24. We used this information to construct variable *howsure,* which takes values from 1 to 10. In principle, this information could be used to try and reduce potential hypothetical bias, to try and improved estimation efficiency, as explained in detail in Section 6.2, and perhaps also to alleviate issues of inconsistency between responses to DBDC payment questions (Flachaire and Hollard, 2007a; Jeanty et al., 2007; Donfouet et al., 2011), as well as to ameliorate problems of insensitivity to scope.[5] However, and although we are planning to explore the possibilities open by our having this numerical certainty scale in future work, we have not, in the preliminary analysis we present in this report, made use of this variable.

---

[5] See, for example, Hammit and Graham (1999) and Alberini et al. (2004).

11.6 Imputation of missing values

We used multivariate imputation techniques to impute missing values for the variables measuring income (18% of cases), education (9% of cases), age (1% of cases), and health (1% of cases), rather than fully discarding the incomplete observations. In order to impute the missing values for these variables we followed the approach developed by Royston (2004, 2005a, 2005b, Royston, 2008, and Royston et al., 2009), conducting an imputation based on an interchained equations algorithm. The command *mi impute chained* built into STATA 12.1 (Statacorp, 2011) was used for this purpose.

To impute the values for our dataset we first assume that the missing variables are missing at random. For instance, this assumption implies that we do not expect that respondents of a specific income level systematically refused to place themselves into the corresponding income bracket. The same applies to education and age categories and the question about health status.

The variables that were used to predict income (using an ordered logit model that took into account the ordered categorical structure of that variable) were *Avalon, childrennumber, SUV, driver, male, newcar, KM, smoker, mathscore, hunt*, and *NLander*. For the imputation of the missing values of variable education we employed an interval regression model with *Avalon, childrennumber, driver, male, KM, smoker, mathscore, hunt, NL*. Variable *age* was also predicted with an interval regression model that included *health, Avalon, childrennumber, driver, male, KM, smoker, mathscore, hunt, and NLander* as predictors, while variable health was considered continuous and explained by *Avalon,*

childrennumber, SUV, driver, male, newcar, KM, smoker, mathscore, hunt, and NLander.

11.7 Stacking of multiple response observations

Respondents who received Version E of the questionnaire answered two pairs of payment questions, namely one about a privately provided risk reduction and one about a publicly provided one, or *vice versa*. Therefore, before analysing all the data, we stacked together their responses so they would become two observations. We also further stacked the initial and follow-up responses from the DBDC questions for every individual in the sample (which would yield two observations from each respondent in Versions A to D and, in the end, quadruple the observations obtained from each available Version E respondent) for the purpose of running random-effects binary models based on the pseudo-panel structure of the thus stacked dataset. The results of this analysis are reported in Section 12.2.

11.8 Clustering by respondent

Additionally, since respondents who received Version E of the questionnaire answered two pairs of payment questions, our estimates are corrected for clustering by individual, according to variable *QUEST*, which identifies each individual. This does not make a difference for respondents who received the other questionnaire versions, since they only answered a single pair of questions.

After all the transformations and eliminations described in this section, we ended up with a sample of 1099 complete usable observations. These 1099 observations were the ones we used for the analyses whose results are included in the remainder of this report.

Part VII

Results

Chapter 12

Willingness to pay

The main contribution of our analysis is to provide an estimate of the expected benefits derived by the population in Newfoundland from the average reduction in the risk associated with MVCs proposed as part of the hypothetical risk reduction policies we included in our survey. More specifically, given the characteristics of our data and the assumptions behind our econometric specifications, our main output consists of an estimate of the distribution of WTP values in the population for the average level of risk. In this section we present the results of the analysis aimed at producing that output.

As explained in Chapter 11, for our final regressions we use a reduced sample ($N =1099$) after eliminating responses that were "clearly" suspected to be protest responses, as well as the cases of those respondents with extreme, implausible, perceptions of their death and injury risk rates.[1] Chapter 11 also

---

[1] Additionally a few incomplete observations were eliminated, because they had missing

details how we treated the original observations by stacking them when necessary to consider the fact that respondents were asked two payment questions (four in the case of Version E). We report standard errors of our estimates corrected for clustering by individual when applicable, as well as weighting the observations using sampling weights that account for the slight deviations from representativeness we detected in our sample in terms of age, gender, and education attaintment levels.[2]

Summary descriptives, based on the final subsample used, of the variables used in the analysis plus some additional ones that are mentioned in different sections of the report and its appendices are included in Table 12.1.

Table 12.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
age	51.833	13.598	19	85	1099
agegroup	2.688	0.841	1	4	1099
ageinterval1	0.067	0.251	0	1	1099
ageinterval2	0.356	0.479	0	1	1099
ageinterval3	0.399	0.49	0	1	1099
ageinterval4	0.178	0.383	0	1	1099
atemoose	0.67	0.47	0	1	1097
Avalon	0.539	0.499	0	1	1099
baseline	6.959964	6.281803	1	80	1099
baselineI	171.905	138.583	1	1000	1062
bid1	71.838	41.232	15	150	1099
bidh	143.676	82.464	30	300	1099
bidl	35.919	20.616	7.5	75	1099
cartype1	0.425	0.495	0	1	1014
case	550	317.398	1	1099	1099
childrenany	0.321	0.467	0	1	1099
childrennumber	0.579	1.051	0	11	1099

Continued on next page...

Variable	Mean	Std. Dev.	Min.	Max.	N
college	0.507	0.5	0	1	1099
comprehensive	0.346	0.476	0	1	1099
COST	71.838	41.232	15	150	1099
COST2	143.676	82.464	30	300	1099
COSTH	35.919	20.616	7.5	75	1099
diffI	38.535	77.263	0	750	1099
diffIsq	7449.17	26047.929	0	562500	1099
diffM	4.455	4.017	0.5	53.33	1099
diffMsq	35.966	128.538	0.25	2844.089	1099
driver	0.926	0.261	0	1	1099
drives30towork	0.167	0.374	0	1	1099
education	4.906	2.292	1	8	1099
federal	0.516	0.5	0	1	514
firstbid	71.749	41.221	15	150	1099
firstresponse	0.58	0.494	0	1	1099
Foreigner	0.022	0.146	0	1	1099
health	78.962	21.568	1	100	1099
hitmoose	0.549	0.498	0	1	1099
howlonginNL	4.086	1.136	1	7	1098
howsure	7.515	2.695	1	10	1099
hunt	0.291	0.455	0	1	1099
income	4.015	2.355	1	8	1099
incomehat3	0.145	0.174	0.003	0.911	1046
inpuage	0.094	0.292	0	1	1099
inpueducation	0.013	0.112	0	1	1099
inpuhealth	0.011	0.104	0	1	1099
inpuincome	0.183	0.387	0	1	1099
job12to6am	0.108	0.311	0	1	1099
KMyear	19835.789	22622.755	0	300000	1099
knowselse	0.758	0.428	0	1	1097
Mainlander	0.082	0.274	0	1	1099
male	0.49	0.5	0	1	1099
math1OK	0.722	0.448	0	1	1099
math2OK	0.623	0.485	0	1	1099
math3OK	0.196	0.397	0	1	1099
math4OK	0.025	0.158	0	1	1099
mathchild	0.559	0.973	0	4	1099
mathscore	1.567	0.939	0	4	1099
monthofbirth	6.652	3.355	1	12	1099
MULTI	3.033	0.818	2	4	1099
newcar	0.527	0.5	0	1	1099
NLander	0.896	0.305	0	1	1099
nn	0.318	0.466	0	1	1099

Variable	Mean	Std. Dev.	Min.	Max.	N
ny	0.103	0.304	0	1	1099
over	0.278	0.448	0	1	1099
privatefirst	0.164	0.37	0	1	1099
protest	0.056	0.231	0	1	1099
protestclear	0	0	0	0	1099
publicgood	0.468	0.499	0	1	1099
Q12	6.087	8.968	1	99	899
Q13	105.185	170.323	1	2500	904
QUEST	19559.226	11403.667	37	39725	1099
referendumreminder	0.475	0.5	0	1	514
secondbid	95.168	79.412	7.5	300	1099
secondresponse	0.453	0.498	0	1	1099
seenmoosecross	0.858	0.349	0	1	1099
smoker	0.182	0.386	0	1	1099
SUV	0.392	0.488	0	1	1099
Version	3.23	1.518	1	5	1099
version1	0.185	0.388	0	1	1099
version2	0.187	0.39	0	1	1099
version3	0.161	0.368	0	1	1099
version4	0.146	0.354	0	1	1099
version5	0.32	0.467	0	1	1099
verysure	0.682	0.466	0	1	1099
weight	0.98	0.879	0.261	22.445	1099
yn	0.229	0.421	0	1	1099
yy	0.35	0.477	0	1	1099

Since the main objective of our analysis is to provide some indication of the magnitude of the benefits of a policy consisting of reducing the risk of colliding with a moose in the highways of Newfoundland, we focus on the estimation of the mean WTP and standard deviation of the WTP distribution for the average-sized reduction in risk. Therefore, we will first consider regressions that allow us to estimate the distribution of an unconditional mean WTP, that is, one which does not take into account the effects on that expected WTP of any covariates. That is, these regressions assume that all respondents share the same systematic

component of WTP and, therefore, have the very same expected WTP, leaving all the remaining variation about the mean in the random error. This notion of a mean WTP surrounded by a collection of positive and negative deviations from that mean implies that we report our results using a single (normal) distribution map of WTP across respondents. We identify that distribution by a location parameter (mean WTP),[3] given by the estimated coefficient α, as described in Section 7.2 and by a measure of its scale, the estimated standard deviation of the WTP distribution, denoted σ in Section 7.2.

Afterwards, however (in Section 12.4), we will also consider more flexible models with covariates, which will not only help us establish the credibility of our WTP results but also provide interesting insights on how this WTP varies depending on several factors related to respondent characteristics and the particular policy described in each version of the questionnaire.

Throughout, we will also provide details about the uncertainty surrounding our central estimates of WTP. In particular, and as explained in Section 7.2, we calculate confidence intervals around the estimated mean WTP values using the Krinsky-Robb procedure (Krinsky and Robb, 1986; Krinsky and Robb, 1990) with randomly draws of 10,000 new values. We also report the ratio of the width of the confidence intervals thus calculated to the point estimate of the mean WTP, which can be used as a measure of the precision of the estimate of mean WTP (Jeanty, 2007; Ekstrand and Loomis, 1998; Loomis and Ekstrand, 1998; Chang et al., 2007; Broberg and Brännlund, 2008a; Martínez-Espiñeira

---

[3]Given that we adopt the assumption of a normal distribution of WTP, the location parameter is also the median of the distribution, since the distribution is symmetric.

and Lyssenko, 2012). This strategy assumes that, if a certain modeling choice biases the mean WTP measure, it will cause approximately the same bias on the point estimate and on the limits of the 95% confidence interval. Otherwise, one could not meaningfully choose among estimators without running the risk of deeming an estimator as "more precise" simply because of the combination of its higher point estimate, potentially due to the effect of upward bias, and a given size of the confidence interval (Martínez-Espiñeira and Lyssenko, 2012).

In Sections 12.1 to 12.3 we provide the regression results obtained from univariate probit and (both flexible and restricted) bivariate probits, random-effects probit and logit models, and interval models, with and without allowing for question effects affecting the responses to the follow-up payment question. Through most of the analysis, we adopt the assumption of the normality of the errors, a testable assumption which cannot be rejected in the case of our data and, furthermore, lends itself better to comparisons across models used to deal with DBDC data, and that the underlying WTP follows a linear function. However, we also consider alternative specifications based on different distributions. Some of these specifications appear to outperform the ones based on assuming linearity in terms of goodness of fit. However, the differences in terms of welfare measures are not substantial once we account for the uncertainty surrounding the values of central estimates. For this reason, most of the results below are based of models that rely on this assumption.

In a less exhaustive format, we report the regression results with covariates in Section 12.4, before reporting welfare measures extrapolated to the population

of the province and values of the value of a statistical life for different levels of risk reduction in Chapter 13.

12.1   Univariate and bivariate probit analysis

In this section we report the regression results corresponding to the univariate probit and bivariate probit regression formats that model the probability of a positive response (in the univariate case) and the different probabilities of a given sequence of responses[4] (in the bivariate case). They allow us to estimate values of mean WTP from the responses to the first payment question (in the univariate case) and from both the initial and follow-up questions (in the bivariate case). Additionally, they permit us to analyse to which extent the assumptions needed to fully exploit the availability of the follow-up questions to improve the precision of the estimates are met in the case of our dataset.

Table 12.2 shows the results of six models based on probit regressions. The first two (Models *probitF* and *probitS*) are univariate models that simply consider the responses to the first and second payment questions independently. These models, as expected, work reasonably well for the case of the initial question, yielding a negative and highly significant coefficient estimated for the bid variable *firstbid*, which falls is line with the expectation that a smaller proportion of respondents will agree, everything else the same, to pay an increasingly high bid value. This negative relationship between the size of the bid and the

---

[4] And thus falling in one of the four regions of the WTP distribution described in Section 7.2.

proportion of respondents willing to pay it is the basis of our estimation of the expected value of the WTP for the policy. Similarly, the constant is positive and significant. The ratio of the constant to the slope variable in this basic model with no other covariates yields the estimate of mean WTP. As shown in the first column of Table 12.2, the estimated mean WTP is \$132.58, surrounded by a, rather wide, 95% confidence interval, [\$92.26, \$332.67].

Model *probitF* uses only the information from the first response, basically using a single-bounded approach to modeling the initial response and ignoring the second response altogether. This approach, therefore, yields an inefficient estimator but, since there is no room for potential question-effects on the second response, any bias associated with these effects is avoided. In intuitive terms, this single-bounded analysis of the initial response leads to a safe but imprecise estimate of mean WTP.

It does not make much sense to analyse the responses to the follow-up independently of the first response in the case of a double-bounded dichotomous choice question format. This can be seen by looking at the results for Model *probitS*, which show a negative constant and a positive slope, which preclude the calculation of a meaningful estimate of mean WTP.[5]

In sum, there should be no concerns about question effects affecting the point estimates of mean WTP based on a the single-bounded model (Model *probitF*), since they are based on only the information from the first response, before the

---

[5] Similarly, alternative strategies to calculate the mean WTP, such as using the alternative parametrization used by the STATA command *singleb* (López-Feldman, 2011) resulted in non-convergence of the maximum likelihood estimation process.

Table 12.2: Univariate and Bivariate Probit Analyses.

	probitF	probitS	biprobit	biprobit0	biprobit1	biprobit2	biprobit3
firstresponse							
firstbid	-0.0032**		-0.0028*	-0.0033**	-0.0033**	-0.0066**	-0.0066**
constant	0.4202**		0.3899**	0.3103**	0.5975**	0.6211**	0.6188**
secondresponse							
secondbid		0.0023**	-0.0023**	-0.0033**	-0.0066**	-0.0066**	-0.0066**
constant		-0.3429**	0.1124	0.3103**	0.6146**	0.6182**	0.6188**
$\widehat{\sigma}$	315.50	432.69					
athrho			0.8203**	0.8804**	6.1030	6.1030	6.1030
N	1099	1099	1099	1099	1099	1099	1099
mean WTP1	132.58		141.53	95.26	100.98	94.66	94.30
mean LB1	92.26		94.29	66.40	83.79	79.89	79.86
mean UB1	332.67		384.87	130.41	120.97	110.37	109.67
CI to WTP ratio	1.81		2.05	0.67	0.37	0.32	0.32
mean WTP2		NA	48.55	95.26	93.19	94.21	94.30
mean LB2		NA	-35.22	66.40	78.70	79.77	79.86
mean UB2		NA	92.71	130.41	108.72	109.58	109.67
CI to WTP ratio		NA	2.64	0.67	0.32	0.32	0.32
log-likelihood	-728.00	-729.49	-1388.75	-1406.84	-1478.08	-1479.67	-1479.73
pseudo R^2	0.01	0.02					
ρ			0.68**	0.71**	1.00^a	1.00^a	1.00^a

$^+p < 0.10$, * $p < 0.05$, ** $p < 0.01$. a: The ρ parameter was restricted to take the value of 1. Confidence intervals (LBi−UB$_i$) around the estimated mean WTP values were caculated using the Krinsky-Robb procedure (Krinsky and Robb, 1986; Krinsky and Robb, 1990) with 10,000 draws.

respondent had an opportunity to even learn that there would be a follow-up question (except in the case of Version 5 respondents, who would have likely expected it in the second round of questioning). However, since the valuation questions were designed using the DBDC format, the estimates from the single-bounded model are likely to be very imprecise, particularly in the case of the second response (Model *probitS*).

The rest of the models reported in Table 12.2 consider the correlation between the two responses, which is due to the fact that they are provided by the same respondent, so they are both affected by unobserved characteristics of that

respondent. We make different assumptions about the strength of that correlation, though. The first model (Model *biprobit*) is a flexible bivariate probit that allows for any given intra-respondent correlation value between -1 and 1. Our estimate of this correlation is $\hat{\rho} = 0.68$ and the null that it is equal to zero can be clearly rejected ($\chi^2(1) = 64.52$ with the Prob $> \chi^2 = 0.0000$).

Comparing the double-bounded dichotomous choice models presented in Section 12.3 below and these bivariate models in Table 12.2 would also require testing whether the correlation coefficient ρ in the bivariate model is equal to 1. However, a classical test of the null hypothesis that ρ equals 1 could not be conducted, because the distribution of the test statistic is not known when the correlation coefficient is equal to one, that is, under the null hypothesis (Alberini et al., 1997).[6] For the same reason, and although the double-bounded model is a limiting case of the random-effects specification for which $\rho = 1$, it is not straightforward to discriminate between the double-bounded model and the random-effects models reported in Section 12.2 by comparing their respective log-likelihood values (Alberini, 1995a).

The structure of the bivariate probit model does, however, make it possible to test restrictions about the equality of the distribution of WTP across responses. That is, we can test whether the two sets of (constant and slope) coefficients could actually be the same in the equations for both responses, an assumption that, together with the assumption that the correlation is perfect between the errors of the two equations, would allow us to fully exploit the efficiency gains

---

[6] León and León (2003) suggest a Bayesian approach to this test, which is beyond the scope of this study.

given by the DBDC format.

First, allowing the correlation coefficient to freely take the estimated value of $\rho = 0.68$, we test the joint null hypothesis that both the constant and the slope coefficient are the same in both response equations. This hypothesis can be confidently rejected ($\chi^2(\ 2) = 24.43$ with Prob $> \chi^2 = 0.0000$). This, together with the fact the estimated correlation coefficient ($\widehat{\rho} = 0.68$) is, although clearly different from zero, relatively far from 1, suggests that caution should be used before using a model that assumes that the responses to the two questions are not affected by question effects. However, we cannot reject the equality of the two slopes by themselves ($\chi^2(1) = 0.18$, with Prob $> \chi^2 = 0.6683$), which suggests that the WTP distribution obtainable from the second response might be different from the WTP distribution given by the second response only to the extent that the constants of the linear component of the underlying WTP function differ. That is, perhaps there is just a (negative, as expected) *shift* in the distribution, as described in Section 7.3, between the first and the second responses. This is confirmed by the estimated mean WTP values from the first and the second responses (\$141.53 and \$48.55, respectively), which are surrounded by quite wide confidence intervals. Note, in particular, that, without imposing any further restrictions on the estimation of the second response, the coefficient for the constant is not significant and the confidence interval for the estimated mean WTP includes a sizable proportion of negative values.[7]

---

[7] That is, some of the possible estimated distributions of WTP would be, according to this model, centered around a negative mean value. This should not be confused with the notion that, in all models, we should expect the normal distribution of WTP values (even around relatively large means) to include a sizable proportion of negative values.

In contrast to the flexible Model *probit* just described, Models *biprobit0,* *biprobit1, biprobit2,* and *biprobit3* constitute increasingly restrictive bivariate probit models. First, we impose the restriction that the intercept and slope coefficients are the same in both response equations (Model *biprobit0*), while allowing for free correlation (which remains high at 0.71, but still much less than 1). Model *biprobit1* imposes instead a correlation between responses equal to 1 ($\rho = 1$), while leaving the size of the regression coefficients unrestricted. These two constraints, as mentioned above, are simultaneously imposed by the DBDC models that make the most of the efficiency gains afforded by the availability of the responses to the follow-up question.

In the fifth column of Table 12.2 we report the results of Model *biprobit1,* showing that the differences in estimated mean WTP vary much less across responses ($100.98 versus $93.19). We can also see that the ratio of the width of the 95% confidence interval to the size of these point estimates of mean WTP is much smaller (0.37 and 0.32, which suggest a much more precise estimation) and crucially that, by constraining ρ to take the value of 1, we cannot reject the equality of neither the slopes across equations ($\chi^2(1) = 2.40$, with Prob $>$ $\chi^2 = 0.1214$) nor, now, the constants ($\chi^2(1) = 0.85$, with Prob $> \chi^2 = 0.3555$) across response equations. The latter means that, under the assumption that the correlation of the error across responses equals one, we could further constraint the values of the coefficients to be equal across responses, in order to maximize the efficiency gains afforded by the DBDC question format.

This we do in Models *biprobit2* and *biprobit3*, the latter being the most

Table 12.3: Likelihood-ratio test comparing bivariate probit models.

Comparison	Test statistic	P-value	Conclusion
biprobit1 vs. biprobit	$\chi^2(1) =178.65$	0.0000	Reject H$_0$: $\rho = 1$
biprobit2 vs. biprobit	$\chi^2(2) = 181.84$	0.0000	Reject H$_0$: $\rho = 1$ and constant slopes
biprobit2 vs. biprobit1	$\chi^2(1) = 3.19$	0.0742	Cannot reject H$_0$: constant slopes (if $\rho = 1$)
biprobit3 vs. biprobit2	$\chi^2(1) = 0.10$	0.7465	Cannot reject H$_0$: constant intercepts (if constant slopes and $\rho = 1$)

restrictive model and equivalent to the interval model initially suggested by Hanemann et al. (1991) to analyse DBDC data. These models are very similar and, due to the restrictions they embody, yield very similar estimates of mean WTP from both responses, all around \$94.

Table 12.3 shows the results of testing how restrictive the constraints imposed on the bivariate probit model are. These tests[8] confirm that, although the assumption that the correlation coefficient ρ is equal to one might be indeed too restrictive in our case, the hypothesis that the regression coefficients are equal across equations under unitary correlation could not be rejected.

Therefore, it is only with caution that one should accept the potentially biased estimates from the restricted bivariate probit model, equivalent to the DBDC interval model, but there is certainly enough of a link between responses to allow us to explore the possibility of taking advantage of the availability of two responses per individual to the WTP question.

---

[8] The test results reported should be in principle considered with caution, though, since they ignore the clustering by respondent (through variable $QUEST$). However, the corresponding Wald test (valid after models with clustering) with a p-value of 0.1214 (for the comparison of Models *biprobit2* versus *biprobit1*) and p-value 0.2719 (for the comparison of Models *biprobit3* versus *biprobit2*) show that the qualitative results of the testing remains unaffected by this issue.

All of the models in Table 12.2 exploit the assumption that the errors follow (either a univariate or a bivariate) normal distribution. We can test this assumption, using Murphy's (2007) Rao score test,[9] on the unrestricted models (univariate Models *probitF* and *probitS*, and the unrestricted bivariate Model bi*probit*). The assumption of normality cannot be rejected for Model *probitF* ($\chi^2(2) = 2.29$, with Prob $> \chi^2 = 0.3183$). In the case of Model *biprobit* the test even more clearly suggests that the assumption of normality is indeed reasonable ($\chi^2(9) = 4.86$, with Prob $> \chi^2 = 0.8465$). On the other hand, the test rejects the null in the case of Model *probitS* ($\chi^2(2) = 14.77$, with Prob $> \chi^2 = 0.0006$).[10] This confirms, once more that it is only reasonable to exploit the information from the second response jointly with the information from the first response.

12.2 Random-effects models

As explained in Section 7.2, DBDC data can also be analysed using random-effect models, since the two responses are provided by the same individual (Alberini et al., 1997). Table 12.4 shows the results of applying this approach, after stacking the data from the initial and follow-up responses from each respondent in to a pseudo-panel with N=2,198, using both a model based on the normal distribution of the errors and one based on their logistic distribution. As ex-

---

[9] We used Stata's command *scoregof*, developed by Chiburis (2012), for the implementation of the test after the probit and bivariate probit regressions (Chiburis, 2009; Chiburis et al., 2011).

[10] Although, as it usually happens, using the variable *secondbid* in logs would lead to the non-rejection of the null.

Table 12.4: Random-effects models of WTP.

	xtprobit	xtprobitsh	xtlogit	xtlogitsh
response				
bid	-0.0060**	-0.0042**	-0.0104**	-0.0073**
follow-up indicator		-0.4300**		-0.7405**
constant	0.5730**	0.6420**	0.9961**	1.1092**
$\ln \sigma_v^2$	0.8793**	0.7503**	1.9753**	1.8398**
N	2198	2198	2198	2198
mean WTP	95.26	101.19[a]	95.40	101.55[b]
mean LB	74.76	72.82	74.83	72.88
mean UB	119.09	137.79	119.31	138.34
CI to WTP ratio	0.47	0.64	0.47	0.64
log-likelihood	-1406.84	-1388.88	-1406.92	-1388.93
ρ	0.71	0.68	0.69	0.66

[+] $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

a. The mean WTP from the first response would be \$152.14 [\$116.31, \$236.31]

a. The mean WTP from the second response would be \$50.24 [\$-13.46,\$82.10\$]

b. The mean WTP from the first response would be \$152.42 [\$116.52, \$237.16]

b. The mean WTP from the second response would be \$50.67 [\$-13.52,\$82.64\$]

pected, the results are very similar. In both cases a mean WTP of about \$95 is obtained with a 95% confidence interval of [\$75, \$119].

In these models, the larger the correlation ρ, the more suitable the random effects model (Greene, 1997, 896–899) relative to the pooled probit. Instead, if ρ is zero, the panel-level variance component is unimportant, and the panel estimator is equivalent to the pooled estimator. The significance of the test statistic of a likelihood-ratio that tests this null hypothesis is included in Table 12.4 as the significance of the estimate of ρ. It is clear from Table 12.4 that the intra-respondent correlation is strong enough to discourage the modeling of the responses to each payment question ignoring the fact that they were provided by the same respondent.[11]

---

[11]Furthermore, the pooled regression (not reported) yields, anomalously, a negative intercept and a positive slope coefficient.

Additionally, when the two error components of the random-effects models are assumed to be normally distributed, they become comparable to the bivariate probit models (Alberini et al., 1997). In the extreme, if the first and second individual WTP^* amounts are assumed identical ($\rho = 1$), a model equivalent to the interval double-bounded model (such as Model *biprobit3* above) obtains. However, the estimated values of ρ shown in Table 12.4 are substantially lower than 1. In fact, Model *xtprobit* is just the equivalent to Model *biprobit0,* that is, the bivariate probit model with the *bid* and constant coefficients constrained to be equal across response equations and a free correlation coefficient ρ (only with a different parametrization). The results in terms of mean WTP and goodness of fit (log-likelihood) are the same but the different parametrization results in a slightly tighter confidence interval.[12]

Following Whitehead (2002), who also use the pseudo-panel approach to estimate WTP using DBDC data with a random-effects probit model, we consider the modeling of a shift effect on WTP between responses in Model *xtprobitsh*. Models *xtprobitsh* and *xtlogitsh* suggest that a downward shift might be at play between the first and the second question. Aadland and Caplan (2004), however, argued that such an estimation procedure leads to inconsistent parameters and Whitehead (2004) confirmed that considering DBDC responses as a pseudo-panel is not as straightforward as they had originally assumed, so we will deal more explicitly with question effects using the approach based on the direct

---

[12] Part of the reduced size of the confidence intervals in the random-effects models is due to the fact that it is not feasible to use them with a covariance matrix robust to the clustering of responses that affect the observations from respondents who received Version E of the questionnaire. This effect, however, accounts for a small fraction of the difference.

estimation of WTP below.[13]

12.3   SBDC model and DBDC interval and question effects models

Although we have so far reported WTP estimations based on univariate and bivariate probits and random-effects probits and logits, under different sets of constraints, the first-proposed and simplest way to analyse responses to DBDC questions is to use the *Interval Model* approach, whose results we report in this section. Additionally, most studies also assume a normal distribution.[14] Before settling on our using the version of this model based on the normal distribution, we also considered a logarithmic distribution, as well as a lognormal distribution. Unfortunately, the models are not nested, so it is difficult to use conventional testing strategies (based on likelihood-ratio tests, for instance) to discriminate among them. However, the differences in terms of goodness of fit did not appear substantial, particularly, as expected, between the normal and the logarithmic models,[15] which could in this case be considered equivalent for practical purposes. The differences between the normal model and the log-

---

[13] The alternative estimation approach used by Flachaire et al. (2007) and Aprahamian et al. (2008) avoids this issue, since they specifically impose the restrictions implied by the structural model on the error terms on the empirical model.

[14] The initial formulation of Interval Data Model is based on logistic and log-logistic distributions (Hanemann et al., 1991; Hanemann and Kanninen, 1999) but, as noted by Cameron and Quiggin (1994) this model does not allow for the non-zero correlation across response equations, which constitutes a disadvantage when it comes to testing for question effects.

[15] The logarithmic model yields a log-likelihood of -1476.75 and an estimated mean WTP of $90.22 [$75.69, $104.75], which is not significantly different from the $94.31 [$79.55, $109.05] obtained assuming a normal distribution of WTP.

normal model[16] are larger, again as usually found. The log-likelihood value for the log-normal model was -1434.81, suggesting a somewhat better fit, and the estimated mean WTP value would be $387.53 [$274.68,$613.14], with the median WTP being $66.29 [55.24,79.17].[17]

The log-normal distribution is often used, because, unlike the normal distribution, it does not imply a range of negative WTP values,[18] which sometimes do not make much sense in the case of the valuation of a public good, unless it happens to be a non-disposable good that is a bad for a proportion of the target population. Instead, the estimates of WTP values arising from log-normal distribution will all end up strictly positive (Riddel, 2001). One disadvantage of the log-normal distribution, though, is that its heavy right tail is the primary determinant of the estimate of its mean, so the mean estimates are much higher than the median estimates and can result in nonsensical estimates of mean WTP. In our case, the mean WTP falls within a reasonable range but is still more than three times larger than the estimated mean WTP obtained

---

[16] The log-normal model, which would be very close to a log-logistic model (not reported), is equivalent to a bivariate probit model that uses the logarithm of the bid values instead of the levels and that constraints the intercept and slope coefficients to be equal across the two equations for the two responses and restricts the correlation coefficient to take the value of one. It can also obtained using a basic command for interval regression (such as Stata's *intreg*) using the logarithmic transformations of the interval limits.

[17] Although in our case, we ended up reporting all of our results about welfare measures with reference to the normal distribution, for which the mean and the median are equal (See Section 13), it should be noted that "The choice of welfare measure – mean, median, or some other quantile of the WTP distribution – also calls for a judgment by the analyst that can involve both ethical and statistical considerations. The mean is the conventional measure in benefit-cost analysis and reflects the Kaldor-Hicks potential compensation criterion; the median may be more realistic in a world where decisions are based on voting and there is concern for the distribution of benefits and costs. From a statistical point of view, the mean is generally far more sensitive than the median to the choice of a response probability model or the method of estimation" (Hanemann and Kanninen, 1999, p. 330).

[18] In Chapter 13 we address the issue of having estimated a negative WTP for a proportion of respondents.

from the other specifications, which makes them, in this sense, a much more conservative choice.

The results shown in Table 12.5 include a single-bounded analysis of the first response only, which corresponds to Model *probitF*, as reported in Table 12.2, yielding a mean WTP of about \$133. Similarly, what we denote Model *doubleb* Table 12.5 corresponds to Model *biprobit3* in Table 12.2. Therefore, those models yield exactly the same measures of welfare as those previously reported. Additional results are obtained from Model *mlshift,* which allows for a shift $(\widehat{\delta})$ in mean WTP between the initial and the follow-up responses,[19] and Models *anchorshift* and *anchor* consider homogeneous anchoring $(\widehat{\gamma})$, with and without a shift, respectively.

The parametrization used in these models directly yields an estimate of the location parameter, that is, the mean (and, since we assume WTP to be a linear function of the bid bid values and to be normally distributed, also the median), of the WTP distribution and also an estimate of its scale,[20] in the form of its standard deviation σ. Following the notation of Section 7.2, we can refer to these parameters as α and σ.

Note that, as explained in Cameron (1988), the estimates reported in Table 12.2 can be recovered from the estimates in Table 12.5. For example, the intercept parameter (the *constant*) in Model *probitF* is given by the ratio $\widehat{\alpha}/\widehat{\sigma}$, while the bid coefficient is given by -1/$\widehat{\sigma}$. The values of $\widehat{\sigma}$ reported in Table 12.2 were

---

[19] Note that Model *shift* is equivalent to Model *probit2* in Table 12.1.

[20] Not to be confused with the *scale* or scope of the risk reduction proposed as part of the policy scenario in the valuation survey.

Table 12.5: Single bounded, interval, and interval models with question effects.

	singleb	doubleb	shift	anchorshift	anchor
meanWTP ($\widehat{\alpha}$)	132.5783**	94.3048**	94.6553**	166.4507*	165.0907*
$\widehat{\sigma}$	315.4964**	152.3992**	152.4058**	603.8251	604.2460
$\widehat{\delta}$(shift)			-0.4447	-0.4051	
$\widehat{\gamma}$(anchoring)				0.7758**	0.7759**
N	1099	1099	1099	1099	1099
meanWTP	132.58	94.30	94.66	166.45	165.09
meanLB	92.26	79.56	79.58	1.16	1.10
meanUB	332.67	109.05	109.73	331.81	329.12
CI to WTP ratio	1.81	0.31	0.32	1.99	1.99
log-likelihood	-728.00	-1479.73	-1479.67	-1460.21	-1460.25

$^{+}p < 0.10$, * $p < 0.05$, ** $p < 0.01$

obtained following these relationships in reverse.

The estimates from the model (Model *shift*) that controls for shift (δ) effects suggest that the shift parameter is not statistically significant, something also suggested by the results of the model that control for both anchoring (γ) and shift (δ) effects. The anchoring parameter is, however, highly significant and, as expected, positive ($\widehat{\gamma} = 0.78$ whether we include also a shift coefficient or not), suggesting that there is a quite strong anchoring effect influencing the responses to the follow-up questions.

In sum, when it comes to comparing the policy-relevant estimates of welfare measures, we can see that the estimated mean WTP obtained from a model that accounts for the double-bounded nature of the question format under the most restrictive assumptions needed to exploit the associated efficiency gains (that is, the interval models Models *biprobit3* and *doubleb*) is smaller than the one obtained from the analysis of the initial question only using a single single-bounded approach (Models *probitF* and *singleb*). This is because most of the

question effects in the double-bounded model turn out in practice to unduly lower the proportion of positive responses to the follow-up bid, making the mean WTP estimate from the basic double-bounded models more conservative than the value estimated from the first question only.

The typical finding of a reduced mean WTP is also accompanied by a sharp increase in the precision of the estimation in going from the single-bounded analysis to the double-bounded analysis.[21] However, since the gain in efficiency that results from exploiting all the information from the responses to both the initial and the follow-up questions might come at the cost of the substantial bias that the question effects might introduce, we also consider the welfare measures obtained from the models that allow for shifts and anchoring between questions.

Since the shift effect seems to exert no significant influence,[22] the mean WTP remains basically unaffected by its inclusion. However, and according to expectations, this estimate increases substantially when anchoring is introduced. In fact, mean WTP takes values over $165 and thus exceeds the initial estimate from the single-bounded model when anchoring is accounted for. On the flip side, and once again in line with our *a priori* expectations, the precision of the WTP estimates is much poorer when anchoring is modelled, which results in very wide confidence intervals. This is because we cannot get much in terms of additional information from the follow-up question if respondents,

---

[21]In particular, note that, apart from the wider interval around the location parameter, Model *singleb* also results in a flatter (more spread) distribution with 33% respondents with predicted negative WTP, while Model *doubleb* only predicts a 27% of respondents with negative WTP.

[22]As we already observed when comparing (in Table 12.3) Model *probit2* (equivalent to Model *shift*) with the more restrictive Model *probit3*.

Figure 12.1: Estimated distributions of willingness to pay based on the point estimates of α and σ obtained from Models *singleb* and *doubleb*.

to a substantial extent, anchor their WTP on the bid proposed for the initial payment question. Additionally, the estimate of the standard deviation of the distribution of WTP across respondents, which in this model's parametrization is not artificially bounded to be strictly positive, is not even significantly different from zero (which implies that certain values of σ within the 95% confidence interval are estimated to take a value less than zero). Therefore, because it yields a not very conservative estimate of the mean WTP, because this mean WTP is surrounded by a wide confidence interval, and because the estimate of σ is anomalous, we will center the discussion on the estimates obtained from Models *singleb* and *doubleb*. The latter is relatively imprecise but unbiased and

the former is somewhat biased but much more efficient.

Figure 12.1 shows graphically the estimation results in terms of the distribution of the WTP across respondents, which follows, by assumption,[23] a normal density function, with mean given by the estimated $\widehat{\alpha}$ and the variance given by the square of the estimated $\widehat{\sigma}$, for the case of the analysis of only the first question using a single-bounded approach and for the basic analysis of the two questions within a double-bounded framework. The difference in terms of mean WTP between the density functions obtained from Model *singleb* and from Model *doubleb* are barely observable in the graph but it can be easily seen that the double-bounded model provides a more precise estimate of that mean.

Adding, in Figure 12.2, the distribution based on the point estimates from Model *anchor* would highlight even more the notion that the mean WTP does not differ much in relative terms depending on the specification but it can also be seen that question effects add a great deal of variance to the estimated distribution of WTP. Although the graphs shown do not illustrate this, the less efficient models also lead to very wide intervals for both the location and scale parameters (as shown in Table 12.5).

In Figures 12.1 and 12.2 we are considering only the point estimate of the parameters of the WTP distribution. In fact, we know that the unknown true parameters remain unobservable and we can only make probabilistic assumptions about them, given the results of the estimation. That is, we face uncertainty about the parameters. However, the estimation results do provide some idea

---

[23] Since the error in Equation 7.4 in Section 7.2 was assumed to follow a normal distribution in our baseline specification.
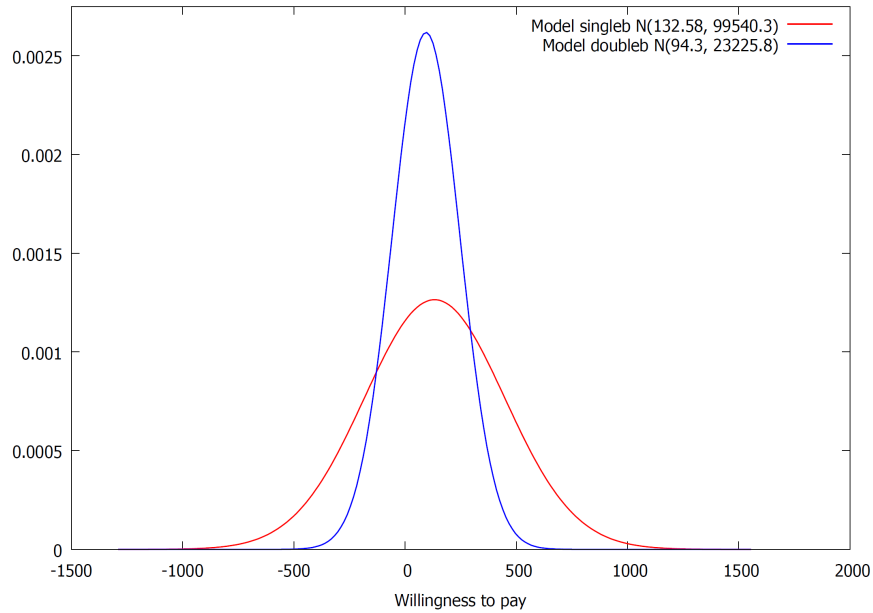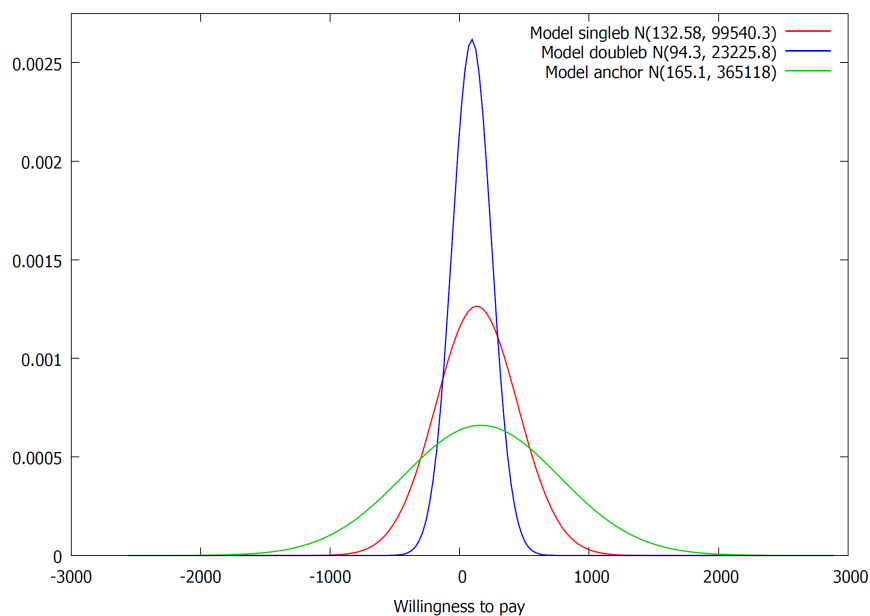
Figure 12.2: Estimated distributions of willingness to pay based on the point estimates of α and σ obtained from Models *singleb*, *doubleb*, and *anchor*.

about how uncertain we are about the values of the parameters, since we have confidence intervals, as shown in Tables 12.5, surrounding the point estimate of mean WTP and, although not shown in the tables of estimation results, around the point estimate of the standard deviation of the WTP distribution ($\widehat{\sigma}$). The estimated width of those intervals is also related to this estimated dispersion of the error, that is, to the estimated dispersion of the latent WTP^*. In order to give the reader an idea of the degrees of parameter dispersion affecting our results, we plot in Figure 12.3 nine distributions based on the results from Model *singleb*. These include the distribution already plotted in Figures 12.1 and 12.2 and additional normal distributions using combinations of the limits of

the 95% confidence intervals around the mean WTP (our estimate of parameter α, which provides a measure of the location of the distribution of WTP) and around the standard deviation (our estimate of σ, which provides a measure of the dispersion of the WTP distribution about its mean). Note that we have abstracted, for simplicity, from the infinite number of distributions in between these extremes. It can be seen that the values of WTP in some of these distributions (the ones based on the smallest value of σ in the interval) are very tightly clustered around the mean, while there is a lot of dispersion in the ones based on the largest values of σ. For policy purposes, this brings up the implication that for the combinations of highest values of mean α and lowest values of σ, there is little need to worry about respondents whose WTP is estimated to be negative, since the portion of the density that falls to the left of zero is negligible. On the other hand, for the lowest values of mean WTP and the largest values of the dispersion coefficient, there is a sizable proportion of respondents whose WTP is estimated as negative. We will deal with this issue more explicitly in Chapter 13.

The density plots in Figures 12.3 and 12.4 also show that the combination of uncertainty due to the stochastic component of WTP about a given mean and the uncertainty about the true value of the parameters identifying the distribution of WTP results in much less overall uncertainty when we use the double-bounded approach, like in the case of Model *doubleb*. In Figure 12.4 all plausible values of WTP (as defined by the 95% confidence intervals estimated for α and σ) fall within a relatively small range.
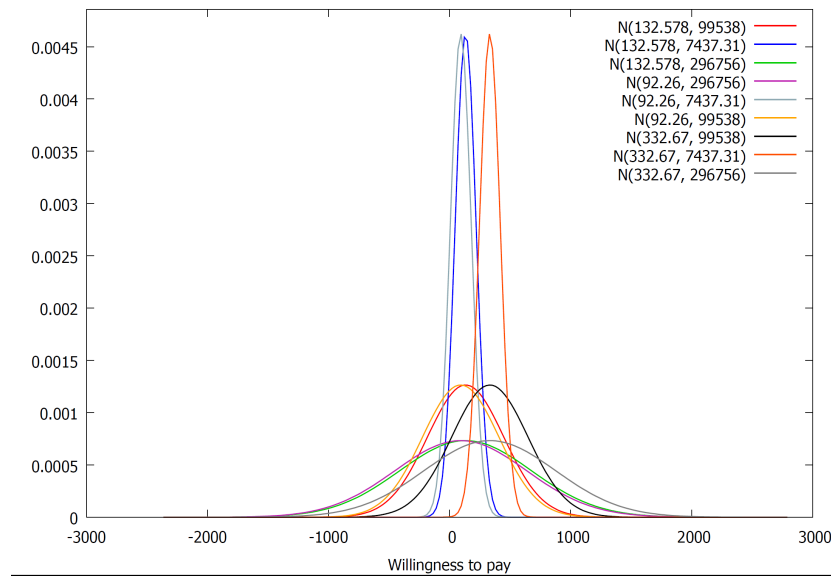
Figure 12.3: Distribution densities based on both point estimates and 95% interval limits for mean WTP (α) and dispersion parameter (σ) from Model *singleb*.
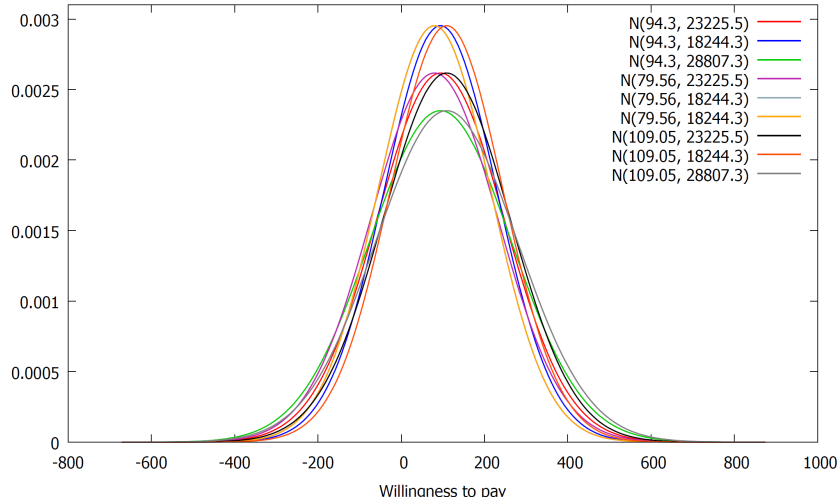
Figure 12.4: Distribution densities based on both point estimates and 95% interval limits for mean WTP (α) and dispersion parameter (σ) from Model *doubleb*.

We have illustrated for the case of the single-bounded approach and the basic double-bounded model the effect on our uncertainty about the predictions about the unconditional value of WTP of two sources of variation, namely variation of α and σ (whose true value is unobservable) about their point estimates and variability of WTP due to the inherent randomness of individual preferences across respondents. Another source of dispersion of WTP that we can consider, though, is that due to observed characteristics of the respondent and the payment scenario in their version of the questionnaire. So far, we have left that source of variation as part of the variation given by the errors in the model. However, by conditioning the mean WTP on the values of variables constructed from the information obtained through the survey, we can make the mean WTP

a function of these variables and leave in the residual only unobserved factors. Therefore, the next section considers the effect of different variables on the expected WTP of each individual.

12.4 Models with covariates

As suggested by Sections 4.2 and 6.6, although the ultimate objective of a CV study might be to obtain a welfare measure for the average respondent, it can also be useful to analyse how the mean WTP is affected by different factors. In our case, we expect the value of mean WTP to be driven by characteristics of the respondent (age, income, self-perceived baseline risk, exposure to risk, etc.) and also by several aspects of the version of the questionnaire the respondent received (proposed level of baseline risk, type of good proposed - public or private, level of risk reduction suggested, etc.). For example, our data made it possible to estimate the WTP for different levels of risk reductions measured in a continuous fashion and also to calculate the value of avoiding a statistical fatality controlling for the effects of variables commonly used in valuation studies of risk reductions.

When it comes to variables about individual characteristics, our respondents answered several questions about a series of sociodemographic characteristics and their experience with the risk valued. In particular, we had available both direct experimental information about and proxies for cognitive scales (as in Andersson and Svensson, 2008), educational attaintment levels, age, income,

gender, perceived health status (as in Andersson, 2007), family composition (Svensson, 2009), experience of the risk (having collided with a moose, variable *hitmoose,* or knowing about someone close who did, variable *knowelse*), degree of exposure to the risk (residence location, Km driven annually, whether the respondent's job involved night-driving or commutting more than 30 Km to work, and the type of car owned), perceived own level of risk (Andersson, 2007), degree of risk aversion (proxied by smoker status), and degree of certainty about the response to the the willingness to pay questions (Alberini et al., 2004).

Apart from these individual respondent characteristics, some of which were used as explanatory variables in our WTP model, we also introduced several controls in the survey instrument. For example, different respondents were asked to value different levels of risk reduction and different scopes of the proposed risk reduction policy (whether the policy considered promised to reduce mortality only or it is more comprehensive, reducing both mortality and morbidity risks). The proposed policy was also randomly varied across respondents in terms of the rivalness of the good provided (whether only the individual or the general population would enjoy the additional safety), the agency in charge of implementing the policy, and the specific means of achieving the risk reduction. In particular, a subsample of our respondents received a question based on the provision of a public good; some others received a question about a privately purchased good that would only protect the driver of a car in the case of a collision, and a third subsample were asked to value both types of goods in sequence. Therefore, we could also compare the effect of the type of good provided (Johannesson et al.,

1996; Hultkrantz et al., 2006; Andersson and Lindberg, 2009; Svensson and Vredin Johansson, 2010; Dekker et al., 2011; Alberini and Ščasný, 2013).

Table 12.6: Interval model with publicgood as only covariate.

	Model doubleb
wtp	
publicgood	48.0214**
constant	70.4115**
σ	149.6178**
N	1099
log-likelihood	-1469.00

$^{+}p < 0.10,\ ^{*}p < 0.05,\ ^{**}{}^{>}p < 0.01$

As a simple illustration, we can thus condition the mean WTP only on an indicator of whether the proposed risk reduction policy involved would be the public good rather than the private good (as described in Section 10.1). The positive and significant estimated coefficient of *publicgood* in this simple model suggests that the mean[24] WTP is much larger for the public good than for the private good in our case. This is not surprising, since we described the private good as something that would only protect the driver of one's vehicle from dying[25] if suffering a collision with a moose (not other passengers, not other drivers, not the moose, and not one's car or anyone else's), while the public good policy involved a the protection of all drivers in the province and involved the reduction in the risk of the collision itself, not only the risk of suffering its impact should the collision take place.

Several studies have found the WTP for a private risk reduction to be much

---

[24] A further extension will investigate whether the variability of the WTP about its mean could also be different depending on whether the respondent received a questionnaire with the public good or the private good scenario or other factors.

[25] And/or being injured, depending on the survey version.

higher than for a public risk reduction of the same nature and magnitude (Svensson and Vredin Johansson, 2010; Hultkrantz et al., 2006; Johannesson et al., 1996). Thus, we might expect the private provision of the good to elicit a higher WTP. However, the nature of the difference in the private and public scenarios may in fact cause the WTP for the publicly provided risk reduction to be higher than for the private one. As in, for instance, Alberini and Ščasný (2013), who themselves found the WTP for the publicly provided risk reduction much higher than for the privately provided one, the public good in our payment scenario involves a risk reduction of the risk of collision for everyone on the highway, while the private good is a risk reduction applied exclusively to the driver and it just involves the death or injury risks following a collision, not the collision itself (so other financial, material, and emotional costs of a collision would remain subject to the same risk rate). If we assume that Newfoundlanders value the safety of others, are willing to pay for it and account for these additional costs, then the publicly provided good should elicit a higher WTP, all else being equal.

Additionally, and as explained in Section 10.1, we assume that the tangible financial and material costs are either fully covered by the drivers' insurance policies or fully taken into account by driver respondents when expressing their WTP for a reduction in risk. This is another reason why we expected the WTP for reducing the risk of suffering a collision altogether to exceed the WTP for a reduction in the risk of dying or getting injured *should the collision occur*, so *publicgood* should have a positive and significant effect on mean WTP.

238

Indeed, the mean WTP can be easily obtained from the results shown in Table 12.6. Although it is estimated as $92.87/year for the average respondent, no such respondent exists, since any given respondent would have received a survey with a payment question about a public good or not. The typical respondent who was asked about the public good policy is expected to be willing to pay $118.43/year, while the private policy is valued on average at $70.41. This is in line with the results of Alberini and Ščasný (2013) and the results obtained in Canada by Adamowicz et al. (2011) for the case of health risk reductions (microbial illnesses/deaths and bladder cancer illnesses/deaths) in the context of drinking water quality treatment by public systems. The difference between WTP for the publicly provided risk reduction and the privately provided one can also be illustrated by considering the approximation to the distribution of WTP across respondents obtained after running the regression shown in Table 12.6. As shown in Figure 12.5, the obviously bimodal kernel density estimate suggests the underlying effect of the two substantially separate distributions that emerge once we condition on *publicgood*. According to these simplified analysis, we would expect only a few respondents to be willing to pay for the private reduction in risk more than what similar respondents are willing to pay for a public risk reduction.

Using a regression with other covariates masks somewhat the double mode given by the effect of *publicgood,* as shown in Figure 12.6. In contrast, using no covariates completely ignores the effects of any potentially influential factors and yields something very close to a normal distribution of WTP across respondents,
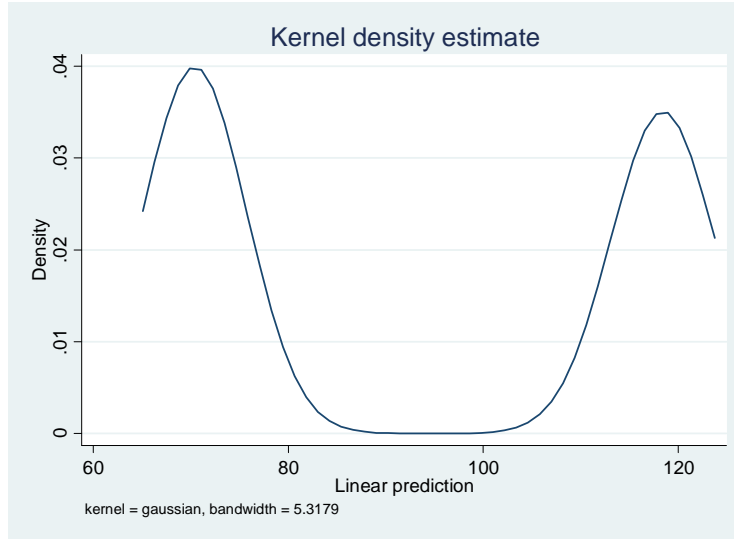
Figure 12.5: Smoothed estimated distribution of WTP when *publicgood* is used as the only covariate.

since the normal distribution was assumed for the WTP construct (Figure 12.7).

We generalised this notion of conditioning on covariates and modelled the WTP as a function of a series of drivers using the three benchmark models from our previous sections, Model *singleb* and Model *doubleb*. We can see in Table 12.7 that Model *singleb* is definitely too imprecise to help us draw strong conclusions about the independent effect of individual variables.[26] We can also see that Model *singleb* seems to suggest that reductions in the risk of injury have a stronger effect on WTP than reductions in the risk of dying from a moose-vehicle collision. Most of the estimated coefficients, however, have the same

---

[26] Although several of the coefficients that appear as not significant even at the 10% level of significance would likely become significant if we had a slightly larger sample, since their associated p-values often fall around 0.15.
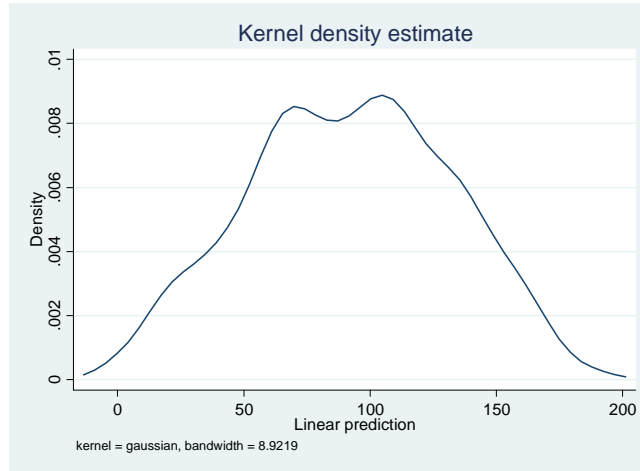
Figure 12.6: Smoothed estimated distribution of WTP with *publicgood* and other covariates (*male, hitmoose, age, income,* and *diffM*).

sign as those from Model *doubleb*, which is much more efficient, although, as described above, could be biased, which explain why their magnitude is smaller than under Model *singleb*. With this caveat in mind, we will focus on the signs of the estimated coefficients, rather than their sizes, throughout most of the discussion below.

We can see that, by and large, the estimated effects of the different covariates in the model fall within *a priori* expectations. First we consider whether respondents exhibited scope sensitivity (having more WTP for larger reductions in risk). The results suggest that respondents are significantly sensitive to the scope of the risk reduction involved on their version of the payment scenario they were allocated both in terms of death risks and, to a lesser extent, injury risks. The effect of increasing the risk reduction promised by the hypotheti-
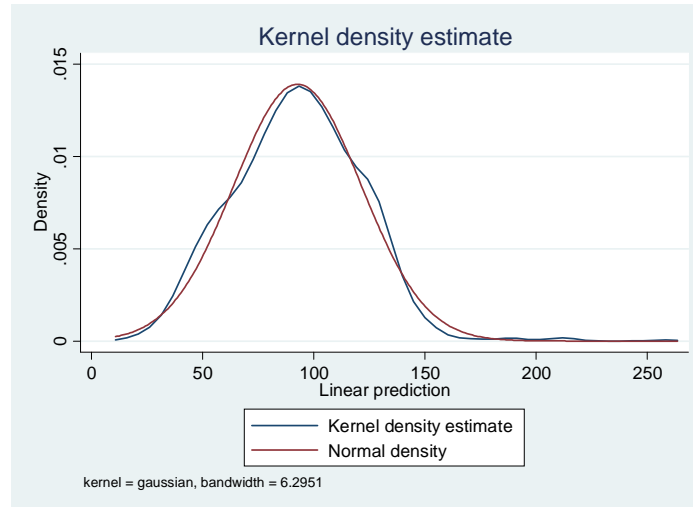
Figure 12.7: Smoothed WTP distribution, with no covariates in the model, showing a mean WTP of about $92.

cal policy is, as the theory would predict, decreasing at the margin. However, although there is a significant sensitivity to scope, this is not enough to allow for near-proportionality, a result that is commonly encountered in this type of studies. This can be seen by considering that the elasticity of the WTP, evaluated at the average values of all the variables, with respect to *diffM* is about 0.18 and about 0.14 with respect to *diffI*. This means that the estimated value of a statistical life associated with this risk reduction policy will depend on the scope of the risk reduction, as shown in Chapter 13.

Another issue that we should mention in relation with the scope variables is that we considered the valuation of both reductions in the risk of death from a MVC and on the risk of getting injured, at least in Versions A and C of the questionnaire. This raises the issue of how to compare the value of a policy that

promises to reduce only the mortality risk with one that promises to reduce both. One way to address this issue would involve translating the injury risk reduction levels (variable *diffI*) into their "death-risk equivalents" (DRE). These DREs can be calculated as the ratio between the unit value of an injury and thus express the outcome of an accident, whether it results in death or not, in a common metric, namely in units of death risk (Hultkrantz et al., 2006). We have not defined or suggested to the respondent what type of injury a MVC would cause. However, an accident caused by a collision with a large-antled animal with the physiological characteristics of the moose is likely perceived by most respondents as relatively dangerous. Therefore, we expected that a DRE for our study would not be one falling far from the estimates used in previous literature for "severe injuries".

Jones-Lee, Loomes, and Philips (1995) suggest that the DRE for a serious non-fatal road injury is between 0.086 and 0.122 with a best estimate of 0.095. This led them to conclude that:

> Given the present state of the art, we feel as confident as it seems prudent to be that the prevention of a typical serious non-fatal injury should be accorded approximately one-tenth of the value placed upon the prevention of a fatality for the purpose of road project appraisal. (Jones-Lee et al., 1995, p. 692-693)

Although they cautioned the reader by adding:

> But it is clear that the present state of the art is beset by a con-

siderable measure of doubt and uncertainty, to which it has to be

conceded that our study may have added as much as it has resolved.

(Jones-Lee et al., 1995, p. 693)

Persson et al. (2000) find a DRE for severe injuries in Sweden around 0.16, while the official values used for benefit-cost assessments by Swedish government agencies imply a DRE of 0.19 (SIKA, 2002), a coefficient also imposed by Hultkrantz et al. (2006) to facilitate a comparison of their results with the current official values applying in Sweden.

In light of these precedents, we would not find it unreasonable to consider the scope of a hypothetical policies leading to, say, a 2 in 100,000 reduction in mortality risk and a 60 in 100,000 reduction in injury risk comparable to a policy reducing the mortality risk only in $2 + 60/10 = 8$ in 100,000. However, the format of our questionnaire involved a reduction in the risks of death and injury that was proportional to their initial incidence. In fact, the reduction in both death and injury risks was simply the result of dividing the baseline by the parameter MULTI (randomly taking a value 2, 3, or 4).[28] Therefore, the we did not need to choose a DRE *a priori*, since, as pointed out by Hultkrantz et al. (2006), the value of a safety improvement that reduces deaths and injuries proportional to their relative frequencies will not be affected by the choice of DRE. This presents the advantage, furthermore, that we would be able to esti-

---

[28]This leads to perfect multicollinearity between *diffI* and *diffM* in the case of the public good versions, since it was systematically assumed that RI was was 30 times RM. Multicollinearity in the whole sample was avoided by having a sizable component of the sample with *diffI*=0, that is, having some respondents value a policy that would only reduce death risks and also by allowing a ratio other than 30 between the death and injury baseline risks in the case of the private versions, given by the self-perceived risk levels.

mate the DRE coefficient by looking at the ratio between the coefficient of diffI and diffM. In the full model (which actually includes also the squared levels of the risk reductinons that form the quadratic form) we obtain that ratio as $0.3940/4.8276 = 0.082$, hence very close to the value recommended by Jones-Lee, Loomes, and Philips (1995). If we used a simplified model with only a constant and *diffI* and *diffM* in levels, the ratio of their estimated coefficients would be about 0.13, so very close to the recommendations given in Hultkrantz et al. (2006).

Apart from the already described positive and significant effect of indicator *publicgood*, we can see that other factors affect the mean WTP for the risk reduction. For example, it is commonly found that males are willing to pay less for risk reductions and we consistently find some weak evidence of this effect. The effect of *income* is positive, although not significant. The same applies to variables *college* and *verysure*, which indicate whether respondents have a university degree and to which extent they were certain about their answers to the payment questions. The WTP appears, as it is usually found in this type of studies, to increase with age until reaching a maximum of about 34 years (in the double-bound model, about 28 in the single-bound model) and then decreasing with age.

Having suffered an accident involving a moose, much as expected, increases also the WTP for a MVC risk reduction policy but just having seen a moose cross the highway has a positive but not significant additional effect. However, even after controlling for both *seenmoosecross* and *hitmoose*, variable *knowselse*

strongly and positively affects WTP. In principle, it seems puzzling that it is the vicarious experience of a collision (or *near miss*) with a moose that has a stronger effect on WTP to reduce the risk of collisions with moose. However, it could well be that someone who already experienced (and survived) a collision considers it more of a familiar risk and now dreads it less, while someone who heard of someone else suffering an accident has a higher WTP, perhaps partly because only the more ghastly experiences of this type are widely shared and more easily recalled.[29]

A similarly strong effect is given by the variable that identifies respondents who commute 30 Km or more to work, likely forcing them to expose themselves to the dangers imposed by moose around the highways in the early morning and late evening, when the risk is highest. On the contrary, the number of Km driven per year actually has a negative effect. This could be explained by the fact that those who use the roads more are better equipped themselves already to deal with the risk of a collision (particularly in the case of long-haul professional drivers). Indeed, if we just remove those 12 respondents in the sample who drive over 100,000 Km a year (which most likely means that they drive a large intercity transportation truck) the effect of *KMyear* is no longer significant.

Another puzzling result is that respondents with children are willing to contribute significantly less than others to reduce the risk of colliding with moose. However, as the significant sign of interaction variable *mathchild* shows, this

---

[29]That is, many respondents might have relatively harmless *near miss* encounters with moose on the road, while the encounters they are most likely to remember from other people's experience would be the most damaging accidents instead.

general effect is driven by those respondents who had the most trouble answering the mathematical and probability questions we used to try and proxy their level of cognitive skills to deal with the questions about risk reductions.

The estimated coefficient of the indicator that the respondent received Version D is significant and positive. Additionally, having been asked about the private good first (in the case of Version E respondents) also significantly increases the expected WTP. In order to understand the meaning of these effects, we refer the reader to recall the information shown in Table 10.1 in Section 10.1. Version D includes a payment scenario based on a public good and a comprehensive risk reduction policy, that is, one that is supposed to reduce both death and injury risks associated with MVCs. We already controlled for the first feature of Version D by introducing variable *publicgood* in our model, while variables *diffI* and *diffIsq,* which take their corresponding value in Versions A and C and a zero in the other versions that only propose a reduction in the death risk, controls for the comprehensiveness of the policy proposed. The only remaining feature that the indicator Version D captures must be the fact that, contrary to the case of Version C, no specific mention was made in Version D of the use of 'fences' to reduce the risk of MVCs. That is, the average respondent, even after controlling for the other aspects of the payment scenario is willing to pay more for a public reduction in the general risk of collisions with moose when the way to achieve that reduction is not specified. This make sense, both because some respondents might have a relative aversion to the fences or doubt their efficacy and because, by not mentioning any particular risk reduction strategy,

the respondent could have assumed that an efficient an effective strategy would be effected.

Understanding the intuition behind the negative and significant sign of the coefficient of Version E might benefit from recalling the type of questioning sequence within that version of the survey instrument. There are two possibilities: either a payment scenario like the one in Version B was presented first, followed by a scenario like the one in Version D or *vice versa.* The effect of this ordering is controlled for by variable *privatefirst,* which is positive but not significant. We would expect *privatefirst* to exert a be positive and significant effect on WTP, because the private good is valued less than the public good (as confirmed by intuition and the positive and significant sign of *publicgood*) and because respondents would likely anchor on their response to the payment question initially faced. That is, someone responding about the private good policy first would have responded without any knowledge that some *better* policy would be proposed afterwards. After answering the dichotomous-choice question with a given maximum WTP in mind, it is likely that they revised their WTP upwards when responding to the payment question about the public good. In general, and particularly because the same bid value was used for both types of goods for each respondent, one would expect a positive effect on WTP for the publicgood when presented second and a negative effect on the WTP for the private good when presented second, leading to a positive effect overall of the variable *privatefirst.* This effect, however, even when we restrict ourselves to the subsample of respondents who received Version E, is not significant. This

would constitute quite good news, since, because these ordering effect is not significant, we do not have to worry about having 'generated' respondents' WTP values rather than just 'eliciting' them.

However, the negative sign on the coefficient of Version E itself suggests that, as the respondents learned how to answer the questions, they tended to start providing more negative responses, which resulted in more conservative estimates of WTP.

In sum, our results reveal that, in general, although not always with great strength or great precision, the predicted mean level of WTP for the proposed risk reduction changes with the values of different variables in the expected direction. This adds to the validity of our estimation exercise and lends support to the idea that the welfare estimates obtained from our analysis are reasonable, in spite of the fact that they have been obtained, as in any non-market valuation exercise, by analysing hypothetical choices rather than actual ones.

The next step consists of using the central estimates of mean WTP for the average risk reduction to derive values of a statistical life (VSL) and population-level welfare measures.

Table 12.7: Single-bounded model and interval model with covariates.

	singleb	doubleb
WTP		
publicgood	56.2996	24.9257$^+$
diffM	0.0899	4.8276$^+$
diffMsq	0.0571	-0.1051$^+$
diffI	1.0345*	0.3940*
diffIsq	-0.0019	-0.0006$^+$
income	7.7713	3.3026
college	34.4044	3.6173
age	3.9808	3.2648
agesq	-0.0715	-0.0480*
verysure	42.3588	18.2835
male	-50.1968$^+$	-19.1359
hitmoose	44.5674	28.8734*
seenmoosecross	71.5756	21.3661
knowselse	106.8238*	53.2393**
drives30towork	102.4664*	53.7000**
KMyear	-0.0014$^+$	-0.0006*
childrenany	-128.2775$^+$	-67.7255*
mathchild	35.5476	25.4963$^+$
Version D	139.6192*	60.0380*
Version E	-55.5023	-38.1844*
privatefirst	76.7000	21.8318
constant	-125.9609	-57.1862
σ	276.0957**	137.3991**
γ		
N	1099	1099
meanWTP	132.58	95.66[27]
meanLB	92.26	84.00
meanUB	332.67	107.03
CItoWTPratio	1.81	0.24
log-likelihood	-658.52	-1390.98

$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Chapter 13

Welfare estimates

13.1 Estimates of aggregate WTP

Before we extrapolate to the population of insular Newfoundland and Labrador the welfare measures obtained for our sample (N=1099), we need to consider a couple of features of the estimated distribution of WTP across respondents. The first one is that, given the estimated mean of about \$96/year and the estimated standard deviation of \$137 of the distribution of WTP across the sample of respondents, some 24% of the respondents are predicted to have a negative WTP. This is because we did not restrict the functional form of the distribution of the underlying WTP to rule out negative values of WTP.[1]

In this case, a choice must be made about how to deal with these proportion

---

[1] See Haab and McConnell (2002, p.85) for a more complete discussion related to the possibility of negative ranges in distribution of WTP.

of respondents. If it is assumed that their WTP is actually negative in the sense that they would need to be compensated for the risk reduction (which, in the case of the public good version of the policy, would thus be a *non-disposable*[2] public bad), \$96/year is the value that should be extrapolated to the population of the island part of Newfoundland and Labrador.[3]

Assuming instead that a negative WTP does not make sense, the estimated WTP can be reinterpreted as "desired WTP" instead (Verbeek, 2008, p. 194). In this case, and using an interpretation similar to that applied in the case of Tobit models (Tobin, 1958), we would consider the actual WTP of these proportion of respondents as zero, since in this case we would not need to worry about compensating them for any negative WTP. The WTP now, given that it is positive, would follow a truncated normal distribution with a (higher, as expected) mean of about \$153. Therefore, setting the WTP values of the 24% of respondents with negative "desired" WTP to zero, the overall mean WTP for the two groups would be about $(1-0.24) \cdot \$153 = \$116$. This would be the mean WTP that we would extrapolate to the number of households in Newfoundland (Cameron and Quiggin, 1994; Verbeek, 2008). According to Statistics Canada's 2011 household counts and assuming that Newfoundland and Labrador have the same average household size, the number of households in the island part of the province of Newfoundland and Labrador would be approximately 197,992.[4]

---

[2]This means that it is not possible to avoid or ignore the benefits (or in this case the negative impacts) of the policy.

[3]Haab and McConnell (2002, p. 86) cite the case of in which individuals with strong views about animal rights might need to be compensated for a policy of deer control based on hunting the deer (Curtis, 2001).

[4]This calculation is based on information on population by economic zone

Therefore, the total annual WTP in the island part of the province for the average risk reduction of 4.455 in 100,000 in the 10-year risk of dying from a MVC (from the average *baseline* used in our sample of 6.960 in 100,000) would be equal to $22,769,080 or close to 23 million dollars per annum.[5]

If, instead, we were to consider that those whose WTP was estimated as negative had to be compensated (rather than just counting their willingness to contribute as zero), perhaps because we conceive of them as being actually made worse-off by the proposed risk reduction policy,[6] we would obtain a total annual WTP of 197,992 times $96. This would amount to something close to 19 million dollars. This would be a relatively conservative estimate, since it is based on the possibly quite biased downwards estimate of mean WTP from the most restrictive double-bounded models. Using the "safer but less precise" estimator based on the responses to the initial question only, we would find a much larger value. However, that value would have been estimated which such imprecision that the associated 95% confidence interval would also encompass the values above.

This analysis of welfare measures also assumes that our sample is (after weighting to adjust for disparities between the proportions of age, gender, and

---

for the whole province of Newfoundland and Labrador from the 2011 census (http://www.stats.gov.nl.ca/Statistics/Census2011/PDF/Pop_EconZone_Total%20Pop_1991-2011.pdf).

[5] This central estimate would be surrounded by a confidence interval with a width of around a couple of million dollars, based on the most precise estimates we can obtain from the analysis of the double-bounded dichotomous choice questions.

[6] Note that this assumption would only make sense in the case of having the risk reduced by the public policy, since the private device for one's car is obviously, as a private good, also a *disposable* good, in the sense that one would not have to rent it if it provided a negative utility (Johannesson et al., 1996).

education levels in the sample versus the population) representative of the population and also that what we deemed to be clear protest responses could justifiably be removed from the analysis.[7]

It should be noted also that the estimates of mean WTP that we extrapolated to the population were the ones corresponding to the average values of all the explanatory variables. For example, if we considered a larger scope of risk reduction, we would have estimated a larger aggregate benefit for the population.[8]

In particular, we should remember that our estimates above come from the joint analysis of the observations from respondents that received the questionnaire based on the public good version of the risk reduction policy and the observations coming from those other respondents who were asked instead about their WTP for a private and individual risk reduction device. Now, it is important to decide which type of measure we would like to extrapolate in order to use it in decision making. In principle, cost-benefit analyses are supposed to be based on the aggregation of individual preferences, considering aggregate WTP and aggregate WTA. In the case of public goods (that is, *non-rival* goods) that are not excludable, there is a joint provision of the good, such that if it is made available to one individual others can enjoy for free, without diminishing anyone else's ability to enjoy the very same units of the good (Samuelson, 1954; Samuelson, 1955). This makes the issue of aggregation of individual preferences

---

[7]Leaving them as rejections of the bid would have substantially decreased the resulting welfare estimates.

[8]In Section 13.2 we show how mean WTP values change with levels of scope (variable *diffM*).

difficult, since when valuing a public good it is perfectly plausible that each individual is considering her altruistic preferences given by the benefit she derives from others' enjoyment of that same good. In that case, the question arises as how the analyst should account for those altruistic values.

In the case of valuations of risk reduction policies, as mentioned in Section 3.1.3, it is often the case that only WTP for reductions in private risks is derived, although the resulting estimates of the VSL are more often than not used to value risk reductions provided as public goods (Brady, 2008). This approach is valid assuming that there is no difference between the WTP for reductions in private versus public risks or that any difference between them does not matter. In our case, we have presented two types of payment scenario, so we can see that individuals show a higher WTP for a risk reduction that affects not only themselves but also other occupants of their vehicle, Newfoundland drivers in general, and also the moose. One possibility would be to aggregate these higher values by multiplying the mean WTP obtained from the sample by the number of households in the island part of the province, assuming that the decision unit is the household (as we have assumed so far),[9] and include in the benefits of the public risk reduction program (whether based on the installation of fences or something unspecified) the safety of others. After all, it is reasonable, when using the WTP approach to the valuation of risk reductions to assume that someone's benefit from a safer road is given by her individual WTP for it as

---

[9]For example, Beattie et al. (1998) follow the same strategy, considering individual WTP when the good valued is a private car safety device and household WTP when the good valued is a public good.

well as how others value that individual's safety, just like, as pointed out by Mishan (1971), the value of a person's life is equal to her private WTP to prevent her own death plus all others' WTP to prevent her death.

The other possibility would be to consider the mean WTP for the reduction of only one's own individual risk of dying (or getting injured) in the event of hitting a moose. It is not surprising that this mean WTP is estimated as much lower, since, even if altruism played no role, the good is not exactly the same. In the case of the public good, the overall risk of a collision is reduced for everyone (and the WTP for this should include the fact that when no collision at all occurs, the vehicle does not get damaged either, there is no time lost, no psychological shock, etc.). This mean estimate of WTP could then be extrapolated by multiplying, not times the number of households, but by the number of individuals in the population or, in our case, the number of individual *drivers* in the province.

Assuming that the variability of the WTP is the same regardless of the type of good (something that we plan to relax in an extension to this work), we can calculate, based on the different means of the WTP distribution, the proportion of respondents expected to have a negative "desired WTP". The second column in Table 13.1 shows that the proportion of respondents with nonnegative WTP values is substantially larger for the case of the public good. Similarly, both the unconditional mean WTP (which is a weighted average that considers all consumers but attaches a WTP value of zero to those falling in the negative range of the WTP distribution) and the raw mean WTP (which implicitly as-

	publicgood=1	publicgood=0	Total
%WTP>0	79	73	76
conditional mean WTP (WTP>0)	$160	$147	$153
unconditional mean WTP (WTP≥0)	$126	$107	$116
mean WTP	$109	$84	$96

Table 13.1: Truncated (conditional) mean WTP; unconditional mean WTP if negative values are counted as zero, and raw mean WTP, assuming that negative WTP values are valid.

sumes that that proportion of respondents would require compensation for the implementation of the policy) are substantially greater in the case the public policy.

In the case of the public good scenario, compensating those who appear to stand to lose from the policy might make more sense, so we use $109 as a conservative measure of mean WTP per household that we multiply times the number of households. It is perfectly plausible that some individuals feel that a risk reduction policy yields, after the added taxes have been accounted for, a negative level of utility. We also found some evidence that some consumers might disagree to a greater or lesser extent with the specific way in which the policy would be effected, since the version of the questionnaire that explicitly mentioned fencing as the specific strategy led to a smaller mean WTP value. The aggregate value of the benefit would then be $109 times 197,992 households,[10] so close to 26 million dollars per annum.

In the case of the private device, however, it certainly makes no sense to account for non-disposability, so we can use $107 multiplied times the number

---

[10]See Lindhjem and Navrud (2009) for a study that explicitly considers the issues involved in choosing between the individual and the household as the decision unit in valuation works.

of individual drivers. The number or relevant drivers could be obtained from official statistics on "licensed drivers". However, in order to maintain consistency with the rest of our analysis, we apply the sample proportion of frequent drivers we found in our data to estimate the number of adults in the population that fall into this category. We found that 92.63% of our respondents said to be frequent drivers. Out of the about 387,000 adults[11] (over 19, the cutoff eligibility age used in our survey), we could assume that there would be then around $358,500$ frequent drivers. This would yield an aggregate benefit $107 times $358,500$, or $38,360,000 from a policy that reduced the average *individual* risk of death for drivers from a MVC by 4.455 in 100,000.

A final note: this calculation of the welfare benefit based on the individual WTP for the private safety device abstracts from the fact that Version A proposed that the device would only reduce the death risk should the respondent suffer a collision. That is, the private device would be similar to an air bag, reducing individual mortality risks but no other personal financial or material costs affecting the individual. It is conceivable to think about a policy that required and subsidized the installation of that type of safety device. However, a more realistic policy would likely involve the reduction in the likelihood of a collision in the first place. Therefore, even when measuring benefits at the strictly individual level, that more realistic policy would yield a higher estimate of aggregate welfare. In this sense, we view this figure of $38,360,000 as a lower bound estimate of the likely benefit that a public reduction strategy of moose-

---

[11]Based on 2011 population data. More current population estimates would be larger, leading to a slightly less conservative measure of aggregate welfare.

vehicle collisions (something like the installation of fences along the highways, or the reduction of the moose herd) would generate.

13.2 Calculation of values of a statistical life

Since one very common approach to describing the benefits of a risk reduction policy involves the estimation of VSL values, we calculated several sets of such values, which depend (apart from on other variables) on the value of $diffM$.[12] This is because, although the estimated level of mean WTP increases with the proposed size of the risk reduction, this increase is not near-proportional, as most other empirical applications find. And just as we had different approaches to the aggregation of welfare benefits, we will have different ways to calculate VSLs depending on the type of mean WTP that we choose. If we ignore the proportion of respondents with a predicted negative WTP and truncate the estimated distribution of WTP at zero, we find a VSL than ranges from over 14 million dollars for a $diffM$ (the reduction in the 10-year death risk rate) of 1 in 100,000 to less than 1 million dollars for a risk reduction of 20 in 100,000. Some intermediate results are shown on the first column of Table 13.2, while in Figure 13.1 we plotted in blue a smoothed scatter of these values and the corresponding

---

[12] Throughout this section, we illustrate how the VSL changes with the death risk reduction $diffM$. We are thus abstracting from changes in $diffI$ that are concomitant with changes in $diffM$. We, instead, just use the mean values of $diffI$ (and its square) throughout for simplicity, in order to follow most similar studies, and to avoid the need to assume a *death rate equivalent* (Hultkrantz et al., 2006; Veisten et al., 2013). or rely on the estimated ratio we found from our data (see Section 12.4). The latter would be very imprecise, since we did not specify the type or seriousness of the injury in the payment scenarios. In a real-life policy setting, we would need to take into account that a policy that reduced the death risk in a certain proportion would likely also reduce the risk of injury, however defined, in some proportion.

Table 13.2: Values of a Statistical Life (millions of CAN$) at different levels of risk reduction ($diffM$), including mean value of 4.47/100,000.

$diffM$	VSL truncated WTP	VSL unconditional WTP	VSL mean WTP
1	14.526	10.524	8.193
2	7.380	5.427	4.322
3	4.995	3.724	3.025
4	3.801	2.870	2.371
4.46	**3.435**	**2.607**	**2.169**
5	3.083	2.355	1.974
6	2.603	2.009	1.706
7	2.259	1.761	1.512
8	1.999	1.572	1.364
9	1.797	1.424	1.246
10	1.633	1.304	1.150
20	0.867	0.721	0.659

values of $diffM$. Note that for the average level of death risk reduction proposed ($diffM = 4.455$)[13] the VSL value based on this truncated distribution would be 3.435 million dollars.

However, it is more reasonable to consider that those who are expected to have a zero WTP for the risk reduction should be included in the calculation of VSL, although with a value of zero WTP. This leads to the second set of estimates of VSL (reported on the second column of Table 13.2 and plotted in green in Figure 13.1) ranging from about 10.5 million dollars for a risk reduction of the death rate of 1 in 100,000 to about $721,000 for $diffM = 20$. The VSL at the mean of diffM = 4.455 would be in this case $2,607,097$.

Adopting a much more restrictive approach based on the notion that those expected to be negatively impacted by the risk reduction policy would need to be compensated for their negative WTP, we obtain the smaller values shown in

---

[13] Note that the average of variable *baseline,* indicating the curret death risk rate is 6.960 in 100,000, so our average risk reduction is rather substantial in relative terms.
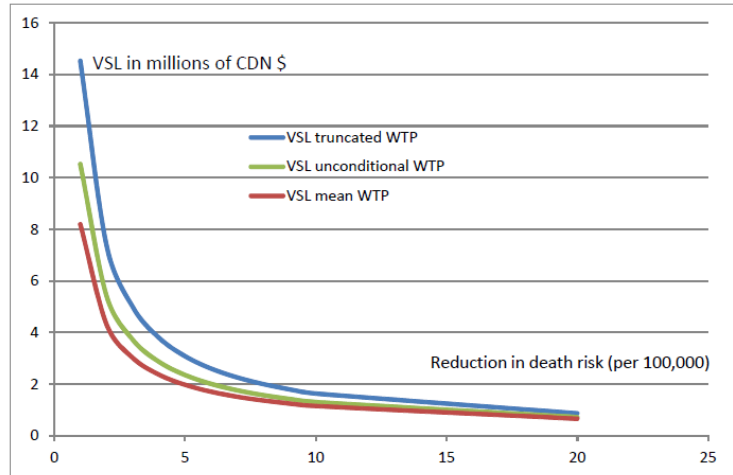
Figure 13.1: VSLs calculated at different levels of risk reduction (*diffM*).

the last column of Table 13.2 and plotted in red in Figure 13.1. These values are based on the untransformed mean WTP obtained from Model *doubleb*. In this case an average level of risk reduction would yield a VSL of about 2.16 million dollars.

The notion that the estimated value of VSL in our study varies substantially with the size of the risk reduction considered (because, although there is a significant degree of sensitivity of WTP to the scope of the proposed risk reduction policy, there is not enough sensitivity to achieve near-proportionality) but takes values very close to what the literature would suggest for the most reasonable ranges of risk reductions is illustrated in Figure 13.1 for the average case. It can be seen that values of VSL corresponding to reductions in the risk of death around 4 to 6 in 100,000 fall roughly within the interval given by 2 and 4 million dollars.
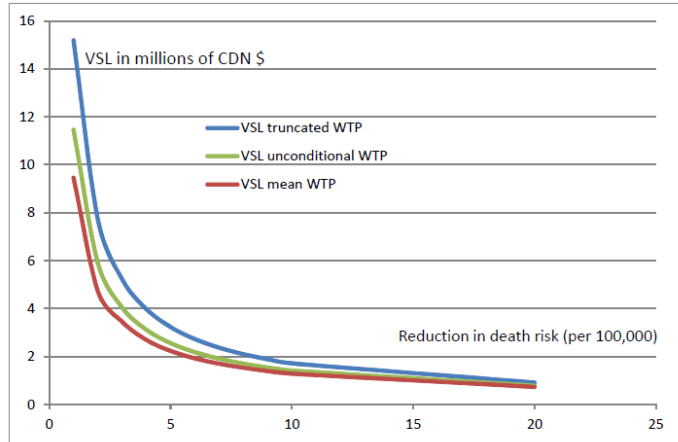
Figure 13.2: VSL values calculated at different levels of risk reduction (*diffM*), evaluated at *publicgood*=1.
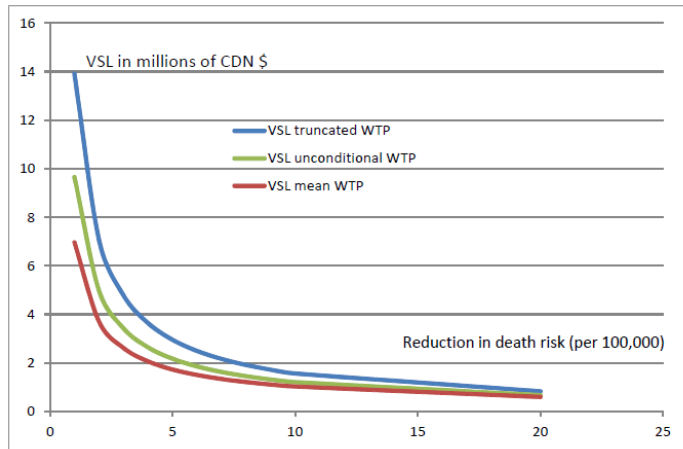
Figure 13.3: VSL values calculated at different levels of risk reduction (*diffM*), evaluated at *publicgood*=0.

Table 1
Willingness-to-pay for reduced road accident fatality risk and implied value of a statistical life.

Risk reduction (per million)	Willingness to pay (US dollars 2005)	Marginal willingness to pay	Value of a statistical life (US dollars 2005)
1	63.38	15.15	63,376,490
5	93.11	4.45	18,621,386
10	109.88	2.63	10,988,241
15	121.06	1.93	8,070,914
20	129.68	1.55	6,484,020
50	161.43	0.77	3,228,583
100	190.51	0.46	1,905,146
200	224.84	0.27	1,124,202

Figure 13.4: Table 1 in Elvik (2013, p. 63). Reproduced with permission from Elsevier.

Similarly, Figures 13.2 and 13.3 illustrate the same type of effect, although in this case we show separately the estimates obtained from the public good version of the policy scenario and the private good version. It is quite remarkable how these estimates of the relationship between VSL and risk reduction levels resemble the equivalent calculations shown by Table 1 and Figure 1 in Elvik (2013, p. 62 and 63) based on the meta-analysis of stated-preference road safety valuation studies by Lindhjem et al. (2011). Their analysis refers to general road safety, so the absolute values of their baseline risks would be higher than in our case and their figures are quoted in USD of 2005. However, most of the risk reductions they consider for the calculations leading to the relationships tabled (see our Figure 13.4) and graphed (see our Figure 13.5) in Elvik (2013, p. 63 and 64) are comparable to ours, ranging from 1 in 1,000,000 to 200 in 1,000,000. Their 10 in 1,000,000 reduction in risk (resulting in a VSL of
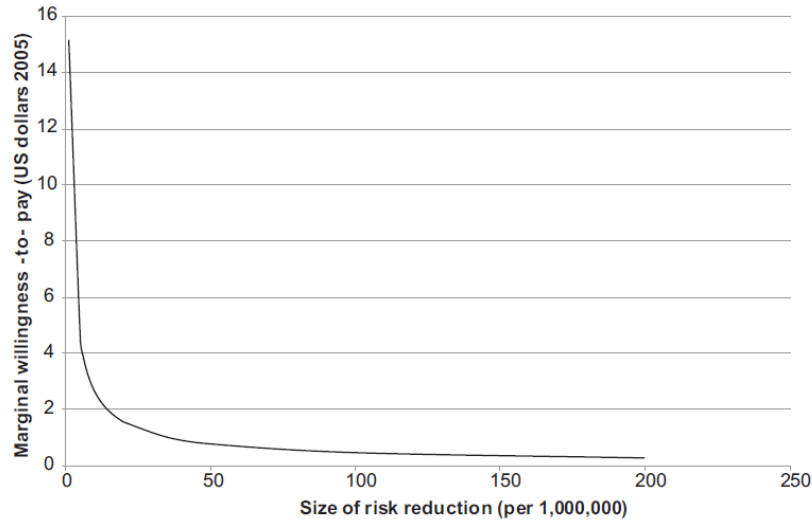
Fig. 1. Demand function for reduced risk of road accident fatality.

Figure 13.5: Figure 1 in Elvik (2013, p. 64). Reproduced with permission from Elsevier.

2005 USD 10,988,241) is equivalent to our 1 in 100,000 risk reduction, which as shown in our Table 13.2 yielded a roughly comparable VSL. Similarly, their 20 in 1,000,000, 50 in 1,000,000, 100 in 1,000,000, and 200 in 1,000,000 risk reduction follow a pattern of VEL decay (6,484,020, 3,228,583, 1,905,146, 1,124,202, all in 2005 USD) that appears to be remarkably comparable to the pattern we found for the corresponding risk reduction values (namely, 2, 5, 10, and 20 in 100,000). In fact, it is quite remarkable, in our view, that not only the figures of VSL are so close to ours for the average level of risk reduction we proposed but also that the pattern of VSL decay between extremely low values of risk reduction and relatively large ones is also very similar.

264

13.3 How do our VSL estimates compare with earlier estimates?

In order to further put our VSL results into context, we can compare them with the results obtained in the literature, which suggest that, although there exist a wide range of estimates available, a most reasonable range would be between 3 and 7 million dollars in individual studies and meta-analytical studies that deal with risks reasonably comparable to the ones we considered and that use the CVM[14] both in different countries (Elvik, 1995; De Blaeij et al., 2003; Kochi et al., 2006) and Canada (Lanoie et al., 1995; Dionne and Lanoie, 2004; Chestnut and De Civita, 2009; Adamowicz et al., 2011).

When it comes to estimates and recommendations for best practices in the Canadian context, we can see that, although our mean estimates of VSL lie on the low side (a point that we revisit in Chapter 14 below), they fall close to what earlier works have found. For example, Dionne and Lanoie (2004), recommend that Canadian Federal and Provincial transport authorities use a VSL of 2000 CAD 5 million as the mean in for cost-benefit analyses, with a band between 3 to 7 million dollars for sensitivity analysis. Their conclusions are the result of reviewing more than 85 VSL studies but based primarily on seven "best" studies in the transport sector (Jones-Lee et al., 1985; Atkinson and Halvorsen, 1990; Kip Viscusi et al., 1991; Dreyfus and Viscusi, 1995; Johannesson and Johansson, 1996; Corso et al., 2001a; Persson et al., 2001). As they show in the summary in

---

[14]Hedonic wage methods tend to yield substantially larger estimates, though.

their Table 5 (Dionne and Lanoie, 2004, p. 264), four of these seven studies are from the US and five of them are based on the CVM, while two are consumer market studies. Their average VSL is 2000 CAD 5,183,000 and their median VSL is 2000 CAD 5,369,000. Dionne and Lanoie (2004) note that the average VSL of Canadian only studies is also about 2000 CAD 5 million.

Further, Zhang et al. (2005) after their comprehensive review of recent VSL studies,[15] which identified several works with a significant component of information pertinent to Canada (Chestnut et al., 1999; Miller, 2000; Boardman et al., 2001; Krupnick et al., 2002; Mrozek and Taylor, 2002; Dionne and Lanoie, 2004; Viscusi and Aldy, 2003), observed that those studies suggested that the VSL in Canada ranged between 2002 CAD 1.0 million and 2002 CAD 7.5 million. They themselves further narrowed this interval suggesting that a reasonable point estimate of the VSL for policy purposes in Canada would be 2002 CAD 4.25 million. They also noted that this VSL was more than twice the figure used by Transport Canada ($1.76 million) but lower than the corresponding values used by Environment Canada ($4.46 million) and Health Canada ($4.47 million) and also slightly higher than the US Office of the Secretary of Transportation 2002 USD 3.0 (USDOT, 2011, p. 6) or 2002 CAD 3.63 million. The latter value has since then been updated to USD 5.8 million (USDOT, 2011, p. 7).

When Zhang et al. (2005) adopt an alternative method consisting of adjusting best US estimates for income differences, which they found in the range of

---

[15] Note that these studies and meta-analyses include both studies based on stated-preference methods and on revealed-preference methods. The reader is referred to Table 3.8 in Zhang et al. (2005, p. 201) for further details on each individual source.

2002 CAD 1.5 million – 2002 CAD 8.5 million, and using an income elasticity ranging between 0.5 and 1.0, they arrive at a best point estimate of the VSL in Canada between 2002 CAD 4.25 million and 2002 CAD 4.63 million. This alternative calculation is thus quite close to their initial one, if slightly higher.

The most recent Canada-based estimates of VSL obtained through a WTP approach that we are aware of are those by Zhang et al. (2013). They explicitly distinguish between paternalistic and non-paternalistic estimates of VSL. Their non-paternalistic public-good VSL for a microbial death (CAD 4.47 million) is reasonably close to our VSL estimates, if, again, higher and their non-paternalistic public-good estimate of VSL for a cancer death is CAD 1.5 million. The same authors, using the same Canadian data and also in the context of risks affecting drinking water, had found a higher VSL for the same type of risk when not differentiating between the private and the public aspects of the risk reduction (Adamowicz et al., 2011).

13.4   VSL in other transport safety studies

So far we have considered comparisons with an "average" VSL for use in Canada. However, the VSL have been consistently found to vary according to individual characteristics (such as income/wealth, age, and culture) and transportation mode or risk context (due in particular to differences in risk levels and degrees of control or dread) or policy dimension characteristics (Zhang et al., 2005). One question beyond the scope of this study is whether the "average" VSL should be

adjusted for such factors.[16] However, since the VSL also depends on the type of risk considered, we focus in this section on the comparison of our VSL estimates with other estimates found in the context of road safety or recommended by governmental agencies in charge of policies related to transportation safety. We find it that it is also the case that our estimates fall close to those that are used or have been used by different agencies that deal with policies related to safety, although perhaps somewhat on the low side.

As explained in detail by Zhang et al. (2005), Transport Canada used for years a VSL in all modes of transportation in the region of 2002 CAD 1.8 million. In 1994, they used CAD 1.5 million in 1991 CAD, equivalent to 2002 CAD 1.762 million, with a a sensitivity analysis in a range between 1991 CAD 0.5 million and 1991 CAD 2.5 million, following a WTP approach and as recommended in the Guide to Benefit Cost Analysis in Transport Canada (Transport Canada, 1994, p. 36) and an evaluation of international studies and best practices. Additionally, Transport Canada (2003, p. 2) works on the assumption that the cost of a life lost in a wildlife vehicle accident is CAD 2 million.[17]

At the time of the report by Zhang et al. (2005), the VSL used by Transport Canada was lower than the one used by Health Canada, which is based on Chestnut et al. (1999). Indeed, on the basis of based on mean VSLs in Chestnut et al. (1999), the Treasury Board of Canada (Jenkins et al., 2007, p. 24)

---

[16]Economic theory would suggest that such adjustments be made in order to improve the efficient allocation of resources. However, ethical, political, and pragmatic most often suggest the contrary. Using an average VSL implicitly adjusts to nullify the influence of income and other factors (Zhang et al., 2005).

[17]In the European Union the value of 1998 Euro 1.5 million has been recommended, being a more readily acceptable estimate than the best scientifically-based estimate of scientifically based one of 1998 Euro 2.4 ± 1 million (Wijnen et al., 2009).

recommends the use of a VSL of \$6.11 million 2004 CAD and an adjustment for inflation using the Canadian CPI. Transport Canada does not select the VSL but just uses this VSL established by the Treasury Board Secretariat (TBS) for use in its regulatory related CBA *(Transportation Safety Board of Canada, 2013)*. It should be stressed that this VSL was selected on the basis of a report for Environment Canada and Health Canada based on a review conducted last century of VSL studies (Chestnut et al., 1999).

This exercise, however, brings to the forefront the fact that, although often-times officially recommended VSL figures end up acquiring an aura of authoritativeness of sorts, they can be chosen on relatively weak grounds[18] and are often transferred across policy contexts based on *ad hoc* adjustments.[19]

For example, in Transport Canada (1994), we find the cautionary note:

> "It is difficult to establish a value of a fatality avoided with objectivity and precision. There are wide variations in the value used by analysts in Canada and in other countries for project evaluation purposes" (Transport Canada, 1994, p. 36)

In further detail, Hauer (2011) highlights these concerns when he the evolution of officially recommended VSL values in the US. The U.S. Department of

---

[18] One example described by Hauer (2011) illustrates this point:

"In this manner an estimate given in 2005–2006 to the last dollar comes from a chain that hinges on a round number provided by administrative guidance in 2002, which was inherited from similar guidance in 1994 because the "newer estimates did not converge on a consensus value or range" and which, in turn, evolved from a research report tabled in 1991" (Hauer, 2011, p. 152)

[19] Wijnen et al. (2009) also point out that the choices of VSL made by governmental agencies tend to be lower than what would be regarded as the best scientific evidence.

Transportation (DOT) followed in the early 1990s the recommendation made by the Office of the Secretary of transportation that ". . .those agencies that use a dollar value of life in economic analysis should use USD 1.5 million." (FHWA, 1994, p. 2), although In 1993 a VSL of USD 2.5 million was recommended (FHWA, 1994). Based on the meta-analysis by Miller et al. (1991) the VSL was updated to USD 2.6 million (FHWA, 1994, p. 2).[20] In 2002 the U.S. DOT adjusted the VSL value noting that: "Recent years have seen considerable expansion in the number of studies published and refinement in analytical techniques. However, it does not appear that newer estimates converge on a consensus value or range that would justify modification of our established standard, and significant estimates continue to lie well below it. . . .we now recommend the use of a value of $3.0 million in all DOT analyses." (Hauer, 2011, p. 152) This recommendation was again based on meta-analyses. This approach involves a kind of averaging of a variety of results; "finding the 'center' of many research results" (Hauer, 2011, p. 152), which can be, and in the case of VSL meta-analyses often are, wildly different (Mrozek and Taylor, 2002; Viscusi and Aldy, 2003; De Blaeij et al., 2003). For example, even within the context of road safety de Blaeij et al. (2003) found estimates ranging from USD 200,000 to USD 30,000,000, concluding that:

> "The assumption that 'life' can be summarized in a single numerical value ("the" VSL), as is often suggested by scholars as well as policy makers, is neither sound from a theoretical perspective, nor

---

[20] This value has been further updated since then, as detailed in USDOT (2011).

warranted on the basis of empirical analysis." (De Blaeij et al., 2003, p. 984)

It is in light of these considerations that we tread lightly when comparing our results with those used as official guidelines. That is, governmental agencies face serious difficulties when choosing a given VSL for policy analysis and their choice is often the result of having to compromise on a central value found on the basis of available meta-analyses that might not even focus on the policy context at hand, while in many other cases, it is simply the result of adopting the recommendation made earlier by some other governmental agency, without considering how it was decided upon. Therefore, although finding VSLs in the case of reductions in the risk of MVCs close to what transportation agencies use in the US or Canada might seem to add a layer of credibility to our results, we suggest that these comparisons be taken with extreme caution. On the other hand, knowing that there is often a great many degrees of separation (in terms of policy context, baseline risk levels, geographical are, timing, sociodemographics, and so on) between the original studies on the basis of which the official VSL is chosen and the policy issues at hand, makes in our opinion a strong case estimate afresh a VSL for the problem of reducing the risk of MVCs in Newfoundland.

Chapter 14

Discussion

The results reported in the previous chapters offer a preliminary picture of what a reduction in the risk associated with suffering a MVC are worth to the people of Newfoundland. We have also found that the task of coming up with benefit values for risk reduction policies presents the researcher with a long series of issues, particularly when one tries to exploit the theoretical efficiency afforded by a DBDC question format. In the end, we chose to base our discussion on the results obtained from the DBDC model, judging that it was the best (or least bad) choice in terms of facing the tradeoff between precision and bias when analyzing responses to iterated payment questions.

Although we could not find any previous studies based in this province for a comparison, our results fall roughly within the expectations generated by previous findings in the context of road safety from other jurisdictions. However, our VSL estimates fall on the lower side of the spectrum. It is likely that

this is because of a combination of factors. Some of these are technical. For example, we feel that even after the refinement of the bid vector (which added two additional higher bid values to the top of the vector), we might not have included enough high bid values to fully capture the whole distribution of WTP. In this sense, our estimates of VSL might be somewhat conservative. This issue will be ameliorated when we analyse a fuller sample resulting from the addition of a further subsample that uses a bid vector with higher bids. Apart from that, we chose to conservatively report the results of the DBDC model. This type of format tends to lead to lower estimates of WTP than the SBDC format. Our estimates of WTP (and therefore VSL) are larger when calculated from the SBDC model.

Additionally, we suspect that the risk of suffering a MVC might be, understandably, considered by many Newfoundland residents as lower to start with than other traffic-related risks. To the extent that one should expect a lower WTP to reduce more remote risks, it is not surprising to find that the estimated VSL from reducing the risk of a MVC is less than the one from reducing other risks. Finally, MVCs might be seen as a relatively familiar concept to most respondents and, presumably, also make them feel less dread than other types of accident. Both factors would contribute to a lower VSL.

On the other hand, some of the research choices needed to analyze the data might have led to a less conservative estimate of VSL than if a different decision had been made. In particular, it should be noted that we eliminated protest responses from the analysis, as per conventional practice in CVM studies. Leaving

those protest responses as negatives to pay the proposed bid would have resulted in smaller estimates of WTP and therefore VSL. Moreover, although for we estimated the measures of monetary benefits that a conventional CBA requires, if the decision about the reduction policy were to be made on the basis of a referendum,[1] we would have to account for the fact that protest responses should count as 'no' votes. That is, protest responses should be eliminated if the decision making is based on the criterion of efficiency, while they should be counted as negative votes if the criterion of majority voting is followed instead.

In Chapter 13, we describe in detail that, given that the distribution function we assumed for the WTP variable, negative WTP values were not ruled out. This is most plausible in the case of the public good version of the proposed policy, since it would deliver a non-disposable good. The estimates of aggregate WTP and the associated estimates of the VSL are affected (if not substantially) by the political decisions made about the potential losers from a policy aimed at reducing the risk of MVCs. This is a particularly relevant point, given that many respondents seemed to feel strongly about the notion that avoiding crashes with moose is a personal responsibility, while others appear to feel, also quite strongly, that the provincial government should do something about the problem.

---

[1] In a referendum the principle of 'one individual one vote' is followed, rather than the principle of 'one dollar one vote'.

Part VIII

Conclusions and suggestions

for further research

This report describes the results of our application of the Contingent Valuation Method to estimate the economic benefits those living on the insular part of Newfoundland and Labrador would derive from a reduction in the risks of death (and injury) associated with moose-vehicle collisions. This type of measure could be used, when conducting a cost-benefit analysis, to inform the calculation of the benefits of any public initiative that would divert taxpayers' money towards the task of protecting the public from the risk of collisions. More generally, these estimates could also inform the assessment not only of road safety policies involving the reduction of moose-vehicle accidents but also of any public initiative aimed at sacrificing of resources in order to protect the public from the risk of collisions.

To our knowledge, this is the first study that estimates how much individuals are willing to pay for death risk reductions and the value of a statistical life in Newfoundland and Labrador, let alone in the specific context of moose-related accidents or even road safety in general. This type of empirical evidence is also relatively scarce and outdated in the wider Canadian context. Therefore, we expect that our analysis will help decision-makers not only in this province but also perhaps in other Canadian regions by providing them with a first-order approximation to the value of prolonging life, often a key ingredient in the assessment of public regulations and projects in the areas of infrastructure development, environment, health, labour, etc.

Our estimated mean willingness to pay for the average reduction in mortality risk associated with moose-vehicle accidents we presented to our respondents

(namely a reduction of 4.46 in 100,000 in the 10-year mortality risk rate from an average baseline of 6.96 in 100,000)[2] is dependent on the specification used to exploit the information obtained from the double-bounded dichotomous choice payment questions. However, we judge that a reasonable lower bound would fall around \$100 per person and year. When this individual mean benefit obtained from our sample is extrapolated to the relevant population, we estimated that the benefit of a policy that delivered a change in risk of this magnitude would total about 20 million dollars per year, slightly more or slightly less depending the treatment of the proportion of respondents whose willingness to pay we predict to be negative. These aggregate estimates would, in principle, be the ones to compare to the costs of implementing a policy or set of regulations expected to deliver a risk reduction of this size from the current level of risk, since the latter is reasonably well approximated by the average value of the baseline proposed in our survey instrument.

Our results also illustrate how the estimated value of a risk reduction policy would quite sensitive to how the risk reduction were to be implemented and who is supposed to benefit. In particular, we focus on the distinction between a policy protecting only individual drivers from death and injury in the event of a crash with a moose and another type of policy that would decrease the province-wide risk of that type of accident altogether. We find that the mean WTP is lower in the former case but, when aggregated over individual drivers,

---

[2] Once again we remind the reader that our mean WTP estimated are based on the average *baseline* death risk rate of 6.960 in 100,000, so our average risk reduction represents a substantial safety improvement in relative terms.

that policy results in an overall benefit close to 40 million dollars per year. The public policy yields, instead, some 26 million dollars a year, because the higher mean WTP must be aggregated, conservatively, only over households.[3]

Again depending on the treatment of respondents predicted to be hurt by the policy, our associated estimates of the value of a statistical life average between 2 and 3.5 million dollars. These VSLs fall perhaps a bit on the lower side but still reasonably close to the central ranges most commonly found in the literature.

Our estimates of the value of a statistical life depend not only on the type of policy proposed, as described above, but, unfortunately, also on the scope of the risk reduction proposed. Indeed, as it is common in these type of studies, we find that the sensitivity of WTP to the size of the risk reduction, although statistically significant, is not sufficient for near-proportionality.

Although our results are in this respect somewhat better than the majority of previous studies that attempt to value risk reductions through stated-preference methods we propose to exploit further the information in our dataset by analysing to which extent different types of respondents tend to exhibit more sensitivity to scope. In particular, we suspect that those who are more comfortable dealing with questions about proportions and probabilities can also better understand the payment scenarios and their willingness to pay is likely more sensitive to the scope of the risk reduction.

---

[3]We would like to stress that one of the limitations of our study is that our aggregate welfare measures are all based on extrapolation to 2011 population estimates. To the extent that the population in this province might have increased substantially during the last three years, our estimates of aggregate willingness to pay could be substantially biased downwards.

We found that there is a wealth of additional analysis that could be performed on the data collected, focusing on different aspects of the problem of valuing risk reductions using stated-preference data. A taste of the type of work that we plan to do as part of future extensions of the main analysis presented above is provided by the Appendices attached to this report. We plan to study in detail the differences in risk perception and the ability of different types of respondents to comprehend changes in small risks. In particular, we will experiment with the technical effects and the consequences in terms of policy recommendations of eliminating respondents with low values of the cognitive index *mathscore* or weighting the observations according to *mathscore*. We also want to consider in further depth the advantages that could be afforded by using the numerical certainty scale to recode or re-weight the raw data.

Among these other suggestions for further research, we would highlight the analysis of both observed heterogeneity and unobserved heterogeneity. In the former case, we are interested and the effects of allowing a flexible scale parameter σ in the estimations of willingness to pay. That is, we want to see if different types of respondents have not only a different mean in their distribution of willingness to pay values but also (or instead) a different variance. In the latter case, we would like to apply latent class modeling techniques and random parameters techniques to both within a discrete framework and a continuous framework account for systematic differences among respondents because of their unobservable characteristics, in particular because of their degree of risk-aversion.

When it comes to the analysis of the question effects affecting the estimation

of double-bound dichotomous choice payment questions, we feel that there is a lot of scope for further work. We are, therefore, planning to investigate alternative models that allow for heterogeneous shifts and anchoring patterns and variations of these models that make it possible also to deal with unobserved heterogeneity.

In the analysis reported above, we have abstracted from issues of potential endogeneity. For example, one could consider models of sample selection to analyse whether those respondents who provided protest responses are systematically different from those who were kept in the sample. It is also conceivable that variables such as the indicators of the degree of risk perception, the experience with the risk valued, the type of vehicle driven, are also endogenously determined with willingness to pay.

Finally, we would like to note that a second wave of fieldwork is already being planned, which should lead to an increase of the sample size sufficient for a more precise estimation not only of the welfare measures needed for policy-making but also for finding significant effects of variables that with our current sample and up with borderline significant coefficients. In addition, we are including slightly altered payment scenarios. We will include an scenario falling in terms of risk reduction beneficiaries between our current private good scenario and our public good scenario. By proposing the rental of an in-vehicle moose-detection device (like an infrared light detector) rather than an individual self-protection device (like an airbag), we will be able to estimate how much more respondents are willing to sacrifice to reduce the risk of an accident occurring in the first

place (and hence avoiding damages to other occupants in their vehicle and also to the moose) than for a risk reduction for themselves only. We will also run the remaining combination of survey variations, namely the combination of two comprehensive payment scenarios: one using the public good first and one using it second.

This second wave of fieldwork will also include a question about the precise type, age, and model of car most commonly driven, which will permit us to construct a measure of revealed willingness to pay for extra safety and o derive measures of risk aversion for the new respondents. Further policy implications in terms of equity and cost distribution could result from the investigation of how risk perception, risk aversion, willingness to pay, and ability to pay are distributed across different types of respondents (divided according to age, health status, income level, education, etc.). We will then address thorny normative questions such as should the value of a statistical life be different across individuals?

However, for now we wanted to provide a first broad analysis of the data in relative simple terms, so we could present the most general policy conclusions. We would like to stress that our results and conclusions are just intended as an ingredient in a broad decision-making process. We can provide only an approximation to the value that the average respondent (or the average respondent in a group sharing certain observable characteristics) would derive for a hypothetical reduction in risk. Whether a particular risk reduction policy is efficient or not depends also on the costs of that policy.

Furthermore, efficiency might be only one criterion, and perhaps not even the main one, followed by those in charge of deciding how to allocated taxpayers' money to protect the population from moose-vehicle collisions. There are other criteria like fairness or political acceptability that could overshadow strict efficiency considerations. Even if, say, a policy of erecting of moose fences were efficient from a strict cost-benefit analysis, it is unclear whether that policy should be carried out. In particular, the distribution of gains and losses from a policy across individuals matters too. The political decision-making process considers distributional issues as part of the overall effects of the provision of public goods. For example, comparing the value of statistical life associated with that type of project with another project, such a regulation expected to reduce the incidence of cancer or heart disease, might help would not provide the final answer for those in charge of choosing how to allocate public funds. The political process might still prompt the government to take measures to reduce moose-vehicle collisions, even if saving a statistical life were less costly in health care, perhaps because it could be considered unfair that the value of a statistical life is higher in health because it is older and richer people that push the willingness to pay to a higher level.

In sum, our study cannot answer questions such as whether the provincial government should spend more or less on reducing road safety risks such as the risk of suffering a moose-vehicle collision, borne largely by a certain group of taxpayers, or on reducing the health risks by, say, younger individuals (such as children) or perhaps older ones (such those disproportionately affected by

cancer risks). More technically, one must remember that the implications of aggregating risks across the population and multiplying by the average, common, value of a statistical life may involve minimal bias if everyone in the population faces the same risk reduction and everyone's willingness to pay for risk reductions is identical. However, in real applications of policies aimed at reducing the risk of moose-vehicle collisions the risk reductions resulting from a policy might be shared very unevenly across individuals and the marginal WTP for risk reductions likely varies widely.[4]

In particular, we would like to stress that our results can only help decide whether it is socially worthwhile to implement a risk reduction policy under certain circumstances but we have no say on recommending it altogether, particularly without further knowledge about its costs of implementation and about the costs and benefits of alternative policies available to the regulator. To put it in allegorical terms, we have provided and estimate of the *speed of the train* but do not know *whether the train station should be built* (not even whether it is worthwhile to build it) and certainly cannot say anything about whether the government would be expected to or obliged to build it.

---

[4] As explained by Cameron (2010) Risk reductions and marginal WTP amounts would then be, potentially correlated, jointly distributed random variables. When two random variables are correlated, it is not true that the average of their products is equal to the product of their averages. When multiplying an aggregate risk reduction by the common value of a statistical life one implicitly assumes the absence of any correlation.

Bibliography

Aadland, D. and Caplan, A. (2004). Incentive incompatibility and starting-point bias in iterative valuation questions: Comment. *Land Economics*, 80(2):312–315.

Ackerman, F. and Heinzerling, L. (2004). *Priceless: On Knowing the Price of Everything and the Value of Nothing*. The New Press, New York.

Adamowicz, W., Dupont, D., Krupnick, A., and Zhang, J. (2011). Valuation of cancer and microbial disease risk reductions in municipal drinking water: An analysis of risk context using multiple valuation methods. *Journal of Environmental Economics and Management*, 61(2):213–226.

Akter, S., Bennett, J., and Akhter, S. (2008). Preference uncertainty in contingent valuation. *Ecological Economics*, 67(3):345–351.

Akter, S., Brouwer, R., Brander, L., and Beukering, P. V. (2009). Respondent uncertainty in a contingent market for carbon offsets. *Ecological Economics*, 68(6):1858–1863.

Alberini, A. (1995a). Efficiency vs bias of willingness-to-pay estimates: Bivariate and interval-data models. *Journal of Environmental Economics and Management*, 29(2):169–180.

Alberini, A. (1995b). Optimal designs for discrete choice contingent valuation surveys: Single-bound, double-bound, and bivariate models. *Journal of Environmental Economics and Management*, 28(3):287–306.

Alberini, A. (2005). What is a life worth? robustness of vsl values from contingent valuation surveys. *Risk Analysis*, 25(4):783–800.

Alberini, A., Boyle, K., and Welsh, M. (2003). Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty. *Journal of Environmental Economics and Management*, 45(1):40–62.

Alberini, A., Cropper, M., Krupnick, A., and Simon, N. B. (2004). Does the value of a statistical life vary with age and health status? evidence from the US and Canada. *Journal of Environmental Economics and Management*, 48(1):769–792.

Alberini, A., Hunt, A., and Markandya, A. (2006). Willingness to pay to reduce mortality risks: Evidence from a three-country contingent valuation study. *Environmental and Resource Economics*, 33(2):251–264.

Alberini, A., Kanninen, B., and Carson, R. T. (1997). Modeling response incentive effects in dichotomous choice contingent valuation data. *Land Economics*, 73(3):309–324.

Alberini, A. and Ščasný, M. (2013). Exploring heterogeneity in the value of a statistical life: Cause of death v. risk perceptions. *Ecological Economics*, 94(C):143–155.

Aldy, J. E. and Viscusi, W. K. (2007). Age differences in the value of statistical life: Revealed preference evidence. *Review of Environmental Economics and Policy*, 1(2):241–260.

Aldy, J. E. and Viscusi, W. K. (2008). Adjusting the value of a statistical life for age and cohort effects. *The Review of Economics and Statistics*, 90(3):573–581.

Allison, P. D. (1999). Comparing logit and probit coefficients across groups. *Sociological Methods and Research*, 28(2):186–208.

Alvarez, R. M. and Brehm, J. (1995). American ambivalence towards abortion policy: Development of a heteroskedastic probit model of competing values. *American Journal of Political Science*, 39(4):1055–1082.

Andersson, H. (2007). Willingness to pay for road safety and estimates of the risk of death: Evidence from a swedish contingent valuation study. *Accident Analysis and Prevention*, 39(4):853–865.

Andersson, H. (2008). Willingness to pay for car safety: Evidence from Sweden. *Environmental and Resource Economics*, 41(4):579–594.

Andersson, H. (2011). Perception of own death risk: An assessment of road-traffic mortality risk. *Risk Analysis*, 31(7):1069–1082.

Andersson, H., Hammitt, J. K., Lindberg, G., and Sundström, K. (2013a). Willingness to pay and sensitivity to time framing: A theoretical analysis and an application on car safety. *Environmental and Resource Economics*, pages 1–20.

Andersson, H., Hole, A. R., and Svensson, M. (2013b). Valuation of small and multiple health risks: A critical analysis of SP data applied to food and water safety. International Choice Modelling Conference, International Choice Modelling Conference 2013.

Andersson, H. and Lindberg, G. (2009). Benevolence and the value of road safety. *Accident Analysis and Prevention*, 41(2):286–293.

Andersson, H. and Lundborg, P. (2007). Perception of own death risk : An analysis of road-traffic and overall mortality risks. *Journal of Risk and Uncertainty*, 34(1):67–84.

Andersson, H. and Svensson, M. (2008). Cognitive ability and scale bias in the contingent valuation method: An analysis of willingness to pay to reduce mortality risk. *Environmental and Resource Economics*, 39(4):481–495.

Andersson, H. and Svensson, M. (2013). Scale sensitivity and question order in the contingent valuation method. Technical report. forthcoming in the Journal of Environmental Planning and Management.

Andersson, H. and Treich, N. (2011). The value of a statistical life. In de Palma, A., Lindsey, R., Quinet, E., and Vickerman, R., editors, *Handbook in Transport Economics*, pages 396 – 424. Edward Elgar.

Andreoni, J. (1990). Impure altruism and donations to public goods: a theory of warm-glow giving. *Economic Journal*, 100(401):464–477.

Aprahamian, F., Chanel, O., and Luchini, S. (2008). Heterogeneous anchoring and the shift effect in iterative valuation questions. *Resource and Energy Economics*, 30(1):12–20.

Araña, J. E. and León, C. J. (2002). Willingness to pay for health risk reduction in the context of altruism. *Health Economics*, 11(7):623–635.

Araña, J. E. and León, C. J. (2008). Do emotions matter? Coherent preferences under anchoring and emotional effects. *Ecological Economics*, 66(4):700–711.

Araña, J. E. and León, C. J. (2012). Scale-perception bias in the valuation of environmental risks. *Applied Economics*, 44(20):2607–2617.

Araña, J. E. and León, C. J. (2007). Repeated dichotomous choice formats for elicitation of willingness to pay: Simultaneous estimation and anchoring effect. *Environmental and Resource Economics*, 36(4):475–497.

Armantier, O. (2006). Estimates of own lethal risks and anchoring effects. *Journal of Risk and Uncertainty*, 32(1):37–56.

Armantier, O. and Treich, N. (2004). Social willingness to pay, mortality risks and contingent valuation. *Journal of Risk and Uncertainty*, 29(1):7–19.

Arrow, K. and Leamer, E. (1994). Comment no. 87. submitted to NOAA in response to advanced notice of proposed rulemaking (59 FR 1062).

Arrow, K. J. (1971). *Essays in the Theory of Risk-Bearing.* Markham Publishing Company, Chicago, IL.

Arrow, K. J., Solow, R., Portney, P., Leamer, E., Radner, R., and Schuman, H. (1993). Report of the NOAA panel on contingent valuation. Technical Report 58: 4601-4614, Federal Register.

Atkinson, S. E. and Halvorsen, R. (1990). The valuation of risks to life: Evidence from the market for automobiles. *The Review of Economics and Statistics*, 72(1):133–136.

Baker, R., Chilton, S., Jones-Lee, M., and Metcalf, H. (2008). Valuing lives equally: Defensible premise or unwarranted compromise? *Journal of Risk and Uncertainty*, 36(2):125–138.

Baker, R., Chilton, S. M., Jones-Lee, M. W., and Metcalf, H. R. T. (2009). Valuing lives equally in a benefit-cost analysis of safety projects: A method to reconcile theory and practice. *Safety Science*, 47(6):813–816.

Balcombe, K. and Fraser, I. (2009). Dichotomous choice contingent valuation with "dont know" responses and misreporting. *Journal of Applied Econometrics*, 24(7):1137–1152.

Barbier, E. B. (1994). Valuing environmental functions: Tropical wetlands. *Land Economics*, 70(2):155–173.

Bashir, M. O. and Abu-Zidan, F. M. (2006). Motor vehicle collisions with large animals. *Saudi Medical Journal*, 27(8):1116.

Bateman, I. J., Carson, R. T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Özdemiroglu, E., Pearce, D. W., Sugden, R., and Swanson, J. (2002). *Economic Valuation with Stated Preferences: A Manual.* Edward Elgar, Ltd., Cheltenham, UK.

Bateman, I. J., Langford, I. H., Jones, A. P., and Kerr, G. N. (2001). Bound and path effects in double and triple bounded dichotomous choice contingent valuation. *Resource and Energy Economics*, 23(3):191–213.

Bateman, I. J., Langford, I. H., and Rasbash, J. (1999). Willingness-to-pay question format effects in contingent valuation studies. In (Bateman and Willis, 1999), pages 511–539.

Bateman, I. J., Langford, I. H., Turner, R. K., Willis, K. G., and Garrod, G. D. (1995). Elicitation and truncation effects in contingent valuation studies. *Ecological Economics*, 12(2):161–179.

Bateman, I. J. and Willis, K. G. (1999). *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EC and Developing Countries.* Oxford University Press, Oxford, UK.

Beattie, J., Covey, J., Dolan, P., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N., Robinson, A., and Spencer, A. (1998). On the contingent valuation of safety and the safety of contingent valuation: Part 1-caveat investigator. *Journal of Risk and Uncertainty*, 17(1):5–25.

Bellavance, F., Dionne, G., and Lebeau, M. (2009). The value of a statistical

life: A meta-analysis with a mixed effects regression model. *Journal of Health Economics*, 28(2):444–464.

Benjamin, D. K. and Dougan, W. R. (1997). Individuals' estimates of the risks of death: Part i - a reassessment of the previous evidence. *Journal of Risk and Uncertainty*, 15(2):115–133.

Benjamin, D. K., Dougan, W. R., and Buschena, D. (2001). Individuals' estimates of the risks of death: Part ii - new evidence. *Journal of Risk and Uncertainty*, 22(1):35–57.

Bergstrom, T. C. (1982). When is a man's life worth more than his human capital? In Jones-Lee, M. W., editor, *The Value of Life and Safety*, pages 3–26. North-Holland, Amsterdam, Netherlands.

Bergstrom, T. C. (2006). Benefit-cost in a benevolent society. *American Economic Review*, 96(1):339–351.

Bernstein, P. L. (1996). *Against the gods: the remarkable story of risk*. John Wiley & Sons, New York.

Bishop, R. C. and Heberlein, T. A. (1979). Measuring values of extramarket goods: Are indirect measures biased? *American Journal of Agricultural Economics*, 61(5):926–930.

Bissonette, J. A. and Rosa, S. (2012). An evaluation of a mitigation strategy for deer-vehicle collisions. *Wildlife Biology*, 18(4):414–423.

Bleichrodt, H. and Eeckhoudt, L. (2006). Willingness to pay for reductions in health risks when probabilities are distorted. *Health Economics Letters*, 15(2):211–214.

Blomquist, G. C. (2004). Self-protection and averting behavior, values of statistical lives, and benefit cost analysis of environmental policy. *Review of Economics of the Household*, 2(1):89–110.

Blomquist, G. C., Blumenschein, K., and Johannesson, M. (2009). Eliciting willingness to pay without bias using follow-up certainty statements: Comparisons between Probably/Definitely and a 10-point certainty scale. *Environmental and Resource Economics*, 43(4):473–502.

Blumenschein, K., Johannesson, M., Blomquist, G. C., Liljas, B., and O'Conor, R. M. (1998). Experimental results on expressed certainty and hypothetical bias in contingent valuation. *Southern Economic Journal*, 65(1):169–178.

Blumenschein, K., Johannesson, M., Yokoyama, K. K., and Freeman, P. R. (2001). Hypothetical versus real willingness to pay in the health care sector: Results from a field experiment. *Journal of Health Economics*, 20(3):441–457.

Boardman, A., Greenberg, D., Vining, A., and Weimer, D. (2001). *Cost-Benefit Analysis: Concepts and Practice.* Prentice Hall, Upper Saddle River, NJ.

Boardman, A. E., Greenberg, D. H., Vining, A. R., Weimer, D. L., and Boardman, A. E. (1997). "Plug-in" shadow price estimates for policy analysis. *The Annals of Regional Science*, 31(3):299–324.

Boholm, A. (1998). Comparative studies of risk perception:a review of twenty years of research. *Journal of Risk Research*, 1(2):135–163.

Boman, M. and Bostedt, G. (1999). Valuing the wolf in Sweden: Are benefits contingent upon the supply? In Boman, M., Brännlund, R., and Kriström, B., editors, *Topics in Environmental Economics*, pages 157–176. Kluwer Academic Press, Dordrecht.

Bond, A. R. F. and Jones, D. N. (2013). Wildlife warning signs: Public assessment of components, placement and designs to optimise driver response. *Animals*, 3(4):1142–1161.

Bosetti, V. and Pearce, D. W. (2003). A study of environmental conflict: The economic value of grey seals in south west england. *Biodiversity and Conservation*, 12(12):2361–2392.

Boyle, K. (2003). Contingent valuation in practice. In Champ, P., Boyle, K., and Brown, T., editors, *A Primer on Nonmarket Valuation*. Kluwer Academic Publishers, Amsterdam.

Boyle, K. and Bergstrom, J. C. (1999). Doubt, doubt, and doubters:the genesis of a new research agenda? In Bateman, I. and Willis, K., editors, *Valuing Environmental Preferences. Theory and Practice of the Contingent Valuation Method in the US, EU,and Developing Countries*, pages 183– 206. Oxford University Press, Oxford.

Boyle, K. J. (1989). Commodity specification and the framing of contingent-valuation questions. *Land Economics*, 65(1):57–63.

Brady, K. (2008). The value of human life: A case for altruism. *Natural Resources Journal*, 48(3):541–564.

Broberg, T. and Brännlund, R. (2008a). An alternative interpretation of multiple bounded WTP data: Certainty dependent payment card intervals. *Resource and Energy Economics*, 30(4):555–567.

Broberg, T. and Brännlund, R. (2008b). On the value of large predators in Sweden: A regional stratified contingent valuation analysis. *Journal of Environmental Management*, 88(4):1066–1077.

Brouwer, R., Langford, I. H., Bateman, I. J., and Turner, R. K. (1999). A meta-analysis of wetland contingent valuation studies. *Regional Environmental Change*, 1:47–57.

Burton, A. C., Carson, K. S., Chilton, S. M., and Hutchinson, W. G. (2003). An experimental investigation of explanations for inconsistencies in responses to second offers in double referenda. *Journal of Environmental Economics and Management*, 46(3):472–489.

Byrne, A. (2012). Introduction of moose to the island of newfoundland. A Research Paper submitted to Provincial Historic Sites as part of the Newfoundland and Labrador Provincial Historic Commemorations Program.

Calia, P. and Strazzera, E. (2001). A sample selection model for protest votes in contingent valuation analyses. *Statistica*, 61(3):473–485.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics Using Stata. Rev. ed.* Stata Press, College Station, TX.

Cameron, T. (2010). Euthanizing the value of a statistical life. *Review of Environmental Economics and Policy*, 4(2):161–178.

Cameron, T., DeShazo, J., and Stiffler, P. (2010). Demand for health risk reductions: A cross-national comparison between the u.s. and canada. *Journal of Risk and Uncertainty*, 41(3):245–273.

Cameron, T. A. (1988). A new paradigm for valuing non-market goods using referendum data: Maximum likelihood estimation by censored logistic regression. *Journal of Environmental Economics and Management*, 15(3):355–379.

Cameron, T. A. (1991). Interval estimates of non-market resource values from referendum contingent valuation surveys. *Land Economics*, 67(4):413–421.

Cameron, T. A. and DeShazo, J. R. (2013). Demand for health risk reductions. *Journal of Environmental Economics and Management*, 65(1):87–109.

Cameron, T. A. and Englin, J. (1997). Respondent experience and contingent valuation of environmental goods. *Journal of Environmental Economics and Management*, 33(3):296–313.

Cameron, T. A. and James, M. D. (1987). Efficient estimation methods for "closed-ended" contingent valuation surveys. *The Review of Economics and Statistics*, 696(2):269–276.

Cameron, T. A., Poe, G. L., Ethier, R. G., and Schulze, W. D. (2002). Alternative non-market value-elicitation methods: Are the underlying preferences the same? *Journal of Environmental Economics and Management*, 44(3):391–425.

Cameron, T. A. and Quiggin, J. (1994). Estimation using contingent valuation data from a "dichotomous choice with follow-up" questionnaire. *Journal of Environmental Economics and Management*, 27(3):218–234.

Carlsson, F., Daruvala, D., and Jaldell, H. (2010). Preferences for lives, injuries, and age: A stated preference survey. *Accident Analysis and Prevention*, 42(6):1814–1821.

Carlsson, F. and Johansson-Stenman, O. (2010). Scale factors and hypothetical referenda: A clarifying note. *Journal of Environmental Economics and Management*, 59(3):286–292.

Carlsson, F., Johansson-Stenman, O., and Martinsson, P. (2004). Is transport safety more valuable in the air? *Journal of Risk and Uncertainty*, 28(2):147–163.

Carson, R., Groves, T., and Machina, M. (1999). Incentive and informational properties of preference questions. Plenary Address to the European Association of Resource and Environmental Economists, Oslo, Norway.

Carson, R. and Mitchell, R. (1995). Sequencing and nesting in contingent valuation surveys. *Journal of Environmental Economics and Management*, 28(2):155–173.

Carson, R., Wilks, L., and Imber, D. (1994a). Valuing the preservation of Australia's Kakadu Conservation Zone. *Oxford Economic Papers*, Vol. 46, Special Issue on Environmental Economics (Oct., 1994),:727–749.

Carson, R. T. (1985). *Three Essays on Contingent Valuation*. PhD thesis, University of California, Berkeley.

Carson, R. T. (1997). Contingent valuation and tests of insensitivity to scope. In Kopp, R., Pommerhene, W., and Schwartz, N., editors, *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*. Kluwer, Amsterdam.

Carson, R. T. (2013). Contingent valuation: A practical alternative when prices aren't available. *Journal of Economic Perspectives*, 26(4):27–42.

Carson, R. T., Flores, N. E., and Meade, N. F. (2001). Contingent valuation: Controversies and evidence. *Environmental and Resource Economics*, 19(2):173–210.

Carson, R. T. and Groves, T. (2007). Incentive and informational properties of preference questions. *Environmental and Resource Economics*, 37:181–210.

Carson, R. T., Hanemann, W. M., Kopp, R. J., Krosnick, J. A., Mitchell, R. C., Presser, S., Ruud, P. A., Smith, V. K., Conaway, M., and Martin, K. (1998). Referendum design and contingent valuation: The NOAA panel's no-vote recommendation. *The Review of Economics and Statistics*, 80(2):335–338.

Carson, R. T. and Mitchell, R. C. (1993). The issue of scope in contingent val-

uation surveys. *American Journal of Agricultural Economics*, 75(5):1263–1267.

Carson, R. T., Wright, J., Alberni, A., Carson, N., and Flores, N. (1994b). A bibliography of contingent valuation studies and papers. Technical report, Natural Resource Damage Assessment, Inc., La Jolla, CA.

Champ, P. and Bishop, R. (2001). Donation payment mechanisms and contingent valuation: An empirical study of hypothetical bias. *Environmental and Resource Economics*, 19(4):383–402.

Champ, P., Flores, N. E., Brown, T. C., and Chivers, J. (2002). Contingent valuation and incentives. *Land Economics*, 78(4):591–604.

Champ, P. A., Bishop, R. C., Brown, T. C., and McCollum, D. W. (1997). Using donation mechanisms to value nonuse benefits from public goods. *Journal of Environmental Economics and Management*, 33(2):151–162.

Champ, P. A., Moore, R., and Bishop, R. C. (2009). A comparison of approaches to mitigate hypothetical bias. *Agricultural and Resource Economics Review*, 38(2):166–180.

Chang, J., Yoo, S., and Kwak, S. (2007). An investigation of preference uncertainty in the contingent valuation study. *Applied Economics Letters*, 14(9):691–695.

Chapman, D. J., Bishop, R. C., Hanemann, W. M., Kanninen, B. J., Krosnick, J. A., Morey, E. R., and Tourangeau, R. (2009). Natural resource dam-

ages associated with aesthetic and ecosystem injuries to Oklahoma's Illinois River System and Tenkiller Lake. Expert Report for State of Oklahoma. Case No. 05-cv-0329-gkf-saj. Technical report, State of Oklahoma v. Tyson Foods, et al. In the United States District Court for the Northern District of Oklahoma I.

Chestnut, L., Mills, D., and Rowe, R. D. (1999). Air quality valuation model version 3.0 (aqvm 3.0) report 2: Methodology. Technical report. Prepared for Environment Canada, Health Canada by Status Consulting Inc., Boulder, CO.

Chestnut, L. G. and De Civita, P. (2009). Economic valuation of mortality risk reduction: review and recommendations for policy and regulatory analysis. Technical report, Government of Canada Policy Research Initiative (Canada).

Chiburis, R. (2012). SCOREGOF: Stata module to perform score test of normality for probit and bivariate probit. Statistical Software Components, Boston College Department of Economics.

Chiburis, R. C. (2009). Score tests of normality in bivariate probit models: Comment and implementation. Working paper, University of Texas at Austin.

Chiburis, R. C., Das, J., and Lokshin, M. (2011). A practical comparison of the bivariate probit and linear IV estimators. Policy Research Working Paper Series 5601, The World Bank.

Chien, Y.-L., Huang, C. J., and Shaw, D. (2005). A general model of starting point bias in double-bounded dichotomous contingent valuation surveys. *Journal of Environmental Economics and Management*, 50(2):362–377.

Cho, S., Yen, T., Bowker, J., and Newman, D. (2008). Modeling willingness to pay for land conservation easements: Treatment of zero and protest bids and application and policy implications. *Journal of Agricultural and Applied Economics*, 40(1):267–285.

Christie, J. and Nason, S. (2003). Analysis of vehicle collisions with moose and deer on New Brunswick arterial highways. volume 2003, pages 1548–1557. Proceedings, Annual Conference - Canadian Society for Civil Engineering, June 4-7.

Clevenger, A. P. (2011). Moose-vehicle collisions and their mitigation in Newfoundland. Technical report. Available at http://sopacnl.com/sopac/wp-content/uploads/2011/05/Moose-vehicle-collisions-and-their-mitigation-in-Newfoundland.pdf.

Clevenger, A. P., Chruszcz, B., and Gunson, K. E. (2001). Highway mitigation fencing reduces wildlife-vehicle collisions. *Wildlife Society Bulletin*, 29(2):646–653.

Collins, A. R. and Rosenberger, R. S. (2007). Protest adjustments in the valuation of watershed restoration using payment card data. *Agricultural and Resource Economics Review*, 36(2):321–335.

Condon, B. S. and Adamowicz, W. L. (1995). The economic value of moose hunting in Newfoundland. *Canadian Journal of Forest Research*, 25(2):319–328.

Corso, P. S., Hammitt, J. K., and Graham, J. D. (2001a). Valuing mortality-risk reduction: Using visual aids to improve the validity of contingent valuation. *Journal of Risk and Uncertainty*, 23(2):165–184.

Corso, P. S., Hammitt, J. K., and Graham, J. D. (2001b). Valuing mortality-risk reduction: Using visual aids to improve the validity of contingent valuation. *Journal of Risk and Uncertainty*, 23(2):165–184.

Cropper, M., Hammitt, J. K., and Robinson, L. A. (2011). Valuing mortality risk reductions: Progress and challenges. *Annual Review of Resource Economics*, 3:313–336.

Curtis, J. A. (2001). The use of follow-up questions to no responses in dichotomous choice contingent valuation surveys. *Agricultural and Resource Economics Review*, 30(2).

Czajkowski, J. R. (2009). Modeling shifts in willingness to pay from a bayesian updating perspective. *Land Economics*, 85(2):308–328.

Czajkowski, M. and Hanley, N. (2008). How to 'sell' an environmental good: Using labels to investigate scope effects. University of Stirling, Economics Discussion Paper 2008-16.

Danks, Z. D. and Porter, W. F. (2010). Temporal, spatial, and landscape habitat

characteristics of moose vehicle collisions in western Maine. *Journal of Wildlife Management*, 74(6):1229–1241.

Davidson, D. J. and Freudenburg, W. R. (1996). Gender and environmental risk concerns: a review and analysis of available research. *Environment and Behavior*, 28(3):302–339.

De Blaeij, A., Florax, R., Rietveld, P., and Verhoef, E. (2003). The value of statistical life in road safety: A meta-analysis. *Accident Analysis and Prevention*, 35(6):973–986.

Dekker, T., Brouwer, R., and Hofkes, M.and Moeltner, K. (2011). The effect of risk context on the value of a statistical life: A bayesian meta-model. *Environmental and Resource Economics*, 49(4):597–624.

DeShazo, J. R. (2002). Designing transactions without framing effects in iterative question formats. *Journal of Environmental Economics and Management*, 43(3):360–385.

Desvousges, W., Mathews, K., and Train, K. (2012a). Adequate responsiveness to scope in contingent valuation. *Ecological Economics*, 84:121–128.

Desvousges, W., Mathews, K., and Train, K. (2012b). The Incremental Adding Up Test for Contingent Valuation. *Working Paper, University of California Berkeley*.

Desvousges, W. H., Johnson, F., Dunford, R., Hudson, S., and Wilson, K. (1993). Measuring natural resource damages with contingent valuation:

Tests of validity and reliability. In Hausman, J. A., editor, *Contingent Valuation: A Critical Assessment.* Elsevier, Amsterdam.

Diamond, P. and Hausman, J. (1994). Contingent valuation: Is some number better than no number? *Journal of Economic Perspectives*, 8(4):45–64.

Diamond, P., Hausman, J., Leonard, G., and Denning, M. (1993). Does Contingent Valuation Measure Preferences? Experimental Evidence. In Hausman, J. E., editor, *Contingent Valuation, A Critical Assessment*, pages 41–89. Elsevier, Amsterdam.

Dionne, G. and Lanoie, P. (2004). Public choice about the value of a statistical life for cost-benefit analyses: The case of road safety. *Journal of Transport Economics and Policy*, 38(2):247–274.

DOECNL (2011). Animal facts: Moose. Department of Environment and Conservation for the Government of Newfoundland. Available at http://www.env.gov.nl.ca/env/snp/programs/education/animalfacts/mammals/moose.html.

DOECNL (2013). Moose management in Newfoundland and Labrador. Department of Environment and Conservation for the Government of Newfoundland. Available at http://www.env.gov.nl.ca/env/wildlife/moosemanage.html.

Dolan, P., Metcalfe, R., Munro, V., and Christensen, M. (2008). Valuing lives and life years: Anomalies, implications, and an alternative. *Health Economics, Policy and Law*, 3(3):277–300.

303

Donfouet, H. P. P., Mahieu, P.-A., and Jeanty, P. W. (2011). Dealing with internal inconsistency in double-bounded dichotomous choice: An application to community-based health insurance. Economics Working Paper Archive (University of Rennes 1 & University of Caen) 201130, Center for Research in Economics and Management (CREM), University of Rennes 1, University of Caen and CNRS.

Dreyfus, M. K. and Viscusi, W. K. (1995). Rates of time preference and consumer valuations of automobile safety and fuel efficiency. *Journal of Law and Economics*, 38(1):pp. 79–105.

Drèze, J. H. (1962). L'utilité sociale d'une vie humaine. *Revue Française de Reserche Opérationnelle*, 6:83–118.

Dubourg, W., Jones-Lee, M., and Loomes, G. (1997). Imprecise preferences and survey design in contingent valuation. *Economica*, 64:115–133.

Duncan, S. (2004). Mainland moose: Status, potential impacts and mitigation considerations of proposed highway 113.

Dussault, C., Poulin, M., Courtois, R., and Ouellet, J.-P. (2006). Temporal and spatial distribution of moose-vehicle accidents in the laurentides wildlife reserve, quebec, canada. *Wildlife Biology*, 12(4):415–425.

Dziegielewska, D. A. and Mendelsohn, R. (2007). Does "no" mean "no"? a protest methodology. *Environmental and Resource Economics*, 38(1):71–87.

Eeckhoudt, L. and Hammitt, J. (2001). Background risks and the value of a statistical life. *Journal of Risk and Uncertainty*, 23(3):261–279.

Eeckhoudt, L. R. and Hammitt, J. K. (2004). Does risk aversion increase the value of mortality risk? *Journal of Environmental Economics and Management*, 47(1):13–29.

Ekstrand, E. R. and Loomis, J. (1998). Incorporating respondent uncertainty when estimating willingness to pay for protecting critical habitat for threatened and endangered fish. *Water Resources Research*, 34:3149–3155.

Eldegard, K., Lyngved, J. T., and Hjeljord, O. (2012). Coping in a human-dominated landscape: trade-off between foraging and keeping away from roads by moose (alces alces). *European Journal of Wildlife Research*, 58(6):1–11.

Elvik, R. (1995). An analysis of official economic valuations of traffic accident fatalities in 20 motorized countries. *Accident Analysis and Prevention*, 27(2):237–247.

Elvik, R. (2013). Paradoxes of rationality in road safety policy. *Research in Transportation Economics*, 43(1):62–70.

Ethier, R. G., Poe, G. L., Schulze, W. D., and Clark, J. (2000). A comparison of hypothetical phone and mail contingent valuation responses for green-pricing electricity programs. *Land Economics*, 76(1):54–67.

Evans, M. and Smith, V. c. (2006). Do we really understand the age-vsl relationship? *Resource and Energy Economics*, 28(3):242–261.

Evans, M. F. and Smith, V. K. (2010). Measuring how risk tradeoffs adjust with income. *Journal of Risk and Uncertainty*, 40(1):33–55.

Farmer, M. and Belasco, E. (2011). A finite mixture model of heterogeneous anchoring with distinct anchoring patterns. *Applied Economics Letters*, 18(2):137–141.

Farrell, T. M., Jr., S., E., J., Clark, D. E., Horner, W. R., Morris, K. I., Finison, K. S., Menchen, G. E., and Cohn, K. H. (1996). Moose-motor vehicle collisions: An increasing hazard in northern new england. *Archives of Surgery*, 131(4):377–381.

FHWA (1994). Motor vehicle accident cost. Technical report, U.S. Department of Transportation, Federal Highway Administration. Technical Advisory T 7570.2.

Finn, P. and Bragg, B. W. (1986). Perception of the risk of an accident by young and older drivers. *Accident Analysis & Prevention*, 18(4):289 – 298.

Finucane, M. L., Slovic, P., Mertz, C. K., Flynn, J., and Satterfield, T. A. (2000). Gender, race, and perceived risk: The 'white male' effect. *Health, Risk and Society*, 2(2):170–172.

Flachaire, E. and Hollard, G. (2006). Controlling starting-point bias in double-bounded contingent valuation surveys. *Land Economics*, 82(1):103–111.

Flachaire, E. and Hollard, G. (2007a). Model selection in iterative valuation questions. *Revue d'Economie Politique*, 117(5):853–865.

Flachaire, E. and Hollard, G. (2007b). Starting point bias and respondent uncertainty in dichotomous choice contingent valuation surveys. *Resource and Energy Economics*, 29(3):183–194.

Flachaire, E., Hollard, G., and Luchini, S. (2007). Heterogeneous anchoring in dichotomous choice valuation framework. *Recherches Economiques de Louvain*, 73(4):369–385.

Ford, A. T., Clevenger, A. P., Huijser, M. P., and Dibb, A. (2011). Planning and prioritization strategies for phased highway mitigation using wildlife-vehicle collision data. *Wildlife Biology*, 17(3):253–265.

Frederick, S. and Fischoff, B. (1998). Scope (in)sensitivity in elicited valuations. *Risk Decision and Policy*, 3:109–123.

Freeman III, A. M. (2003). *The Measurement of Environmental and Resource Values: Theory and Methods.* Resources for the Future, Washington D.C., 2 edition.

Garcia, S., Harou, P., Montagné, C., and Stenger, A. (2008). Models for sample selection bias in contingent valuation: Application to forest biodiversity. *Journal of Forest Economics*, 15(1-2):59–78.

Gayer, T., Hamilton, J. T., and Viscusi, W. K. (2000). Private values of risk

tradeoffs at superfund sites: housing market evidence on learning about risk. *The review of economics and statistics*, 82(3):439–451.

Geehan, S. (2011). Supplementary affidavit presented to tha Supreme Court of Newfoundland and Labrador Trial Division (General) in the trial between Ben Bellows and Hugh George and Her Majesty the Queen in right of Newfoundland and Labrador. From a digital web source. Government Code for document:2011 01G 0013CP.

Giraud, K. L., Loomis, J. B., and Johnson, R. L. (1999). Internal and external scope in willingness-to-pay estimates for threatened and endangered wildlife. *Journal of Environmental Management*, 56(3):221–229.

Glendon, A. I., Dorn, L., Davies, D. R., Matthews, G., and Taylor, R. G. (1996). Age and gender differences in perceived accident likelihood and driver competences. *Risk Analysis*, 16(6):755–762.

González-Cabán, A., Loomis, J. B., Rodriguez, A., and Hesseln, H. (2007). A comparison of CVM survey response rates, protests and willingness-to-pay of native americans and general population for fuels reduction policies. *Journal of Forest Economics*, 13(1):49–71.

Goodstein, E. (2005). *Economics and the Environment*. Prentice Hall, Englewood Cliffs, NJ.

Gould, W., Pitblado, J., and Sribney, W. (2003). *Maximum Likelihood Estimation with Stata*. Stata Press., College Station,Texas, 2 edition.

Groothuis, P. A., Groothuis, J. D., and Whitehead, J. C. (2007). The willingness to pay to remove billboards and improve scenic amenities. *Journal of Environmental Management*, 85(4):1094–1100.

Groothuis, P. A. and Whitehead, J. C. (2002). Does don't know mean no? Analysis of 'don't know' responses in dichotomous choice contingent valuation questions. *Applied Economics*, 34(15):1935–1940.

Grosman, P., Jaeger, J., Biron, P., Dussault, C., and Ouellet, J.-P. (2011). Trade-off between road avoidance and attraction by roadside salt pools in moose: An agent-based model to assess measures for reducing moose-vehicle collisions. *Ecological Modelling*, 222(8):1423–1435.

Gunderson, M. and Hyatt, D. (2001). Workplace risks and wages: Canadian evidence from alternative models. *Canadian Journal of Economics*, 34(2):377–395.

Gunson, K. E., Chruszcz, B., and Clevenger, A. P. (2003). Large animal-vehicle collisions in the central canadian rocky mountains: patterns and characteristics. In Irwin, C. L., Garrett, P., and McDermott, K., editors, *Proceedings of the International Conference on Ecology and Transportation*. Raleigh, NC: Center for Transportation and the Environment, North Carolina State University.

Guo, L. (2011). A Canadian symbol brought to court: Victims of moose crashes sue Newfoundland in class action suit. *The Court*. Available at http://www.thecourt.ca/2011/06/13/.

Haab, T. and McConnell, K. (2002). *Valuing Environmental and Natural Resources: Econometrics of Non-Market Valuation.* Edward Elgar, Cheltenham, UK.

Haab, T. C. (1997). Analyzing multiple question contingent valuation surveys: A reconsideration of the bivariate probit. Working Papers 9711, East Carolina University, Department of Economics.

Haener, M. K. and Adamowicz, W. L. (1998). Analysis of "don't know" responses to referendum contingent valuation questionss. *Agricultural and Resource Economics Review*, 27(2):218–230.

Hakes, J. K. and Viscusi, W. (2004). Dead reckoning: Demographic determinants of the accuracy of mortality risk perceptions. *Risk Analysis*, 24(3):651–664.

Halstead, J., Luloff, A., and Stevens, T. H. (1992). Protest bidders in contingent valuation. *Northeast Journal of Agricultural and Resource Economics*, 21(2):160–169.

Hammitt, J. and Treich, N. (2007). Statistical vs. identified lives in benefit-cost analysis. *Journal of Risk and Uncertainty*, 35(1):45–66.

Hammitt, J. K. (2000a). Evaluating contingent valuation of environmental health risks: The proportionality test. *AERE (Association of Environmental and Resource Economics) Newsletter*, 20(1):14–19.

Hammitt, J. K. (2000b). Valuing mortality risk: Theory and practice. *Environmental Science and Technology*, 34(8):1396–1400.

Hammitt, J. K. (2002). Qalys versus wtp. *Risk Analysis*, 22(5):985–1001.

Hammitt, J. K. (2007). Valuing changes in mortality risk: Lives saved versus life years saved. *Review of Environmental Economics and Policy*, 1(2):228–240.

Hammitt, J. K. and Graham, J. D. (1999). Willingness to pay for health protection: Inadequate sensitivity to probability? *Journal of Risk and Uncertainty*, 18(1):33–62.

Hanemann, W. M. (1985). Some issues in continuous- and discrete-response contingent valuation studies. *Northeastern Journal of Agricultural and Resource Economics*, 14:5–13.

Hanemann, W. M. (1991). Willingness to pay and willingness to accept: How much can they differ? *American Economic Review*, 81(3):635–647.

Hanemann, W. M. (1994). Valuing the environment through contingent valuation. *The Journal of Economic Perspectives*, 8(4):19–43. Symposia: Contingent Valuation.

Hanemann, W. M. and Kanninen, B. J. (1999). The statistical analysis of discrete-response CV data. In (Bateman and Willis, 1999), pages 302–442.

Hanemann, W. M., Loomis, J., and Kanninen, B. J. (1991). Statistical efficiency of double-bounded dichotomous choice contingent valuation. *American Journal of Agricultural Economics*, 73(4):1255–1263.

Hanley, N. and Kriström, B. (2008). What's it worth? Exploring value uncertainty using interval questions in contingent valuation.

Hanley, N. D. and Spash, C. L. (1995). *Cost-Benefit Analysis and the Environment.* Edward. Elgar Publishing Ltd., Hampshire, UK.

Hauer, E. (2011). Computing what the public wants: Some issues in road safety cost-benefit analysis. *Accident Analysis and Prevention*, 43(1):151–164.

Hausman, J., editor (1993). *Contingent Valuation: A Critical Assessment.* North-Holland, Amsterdam.

Hausman, J. (2012). Contingent valuation: From dubious to hopeless. *Journal of Economic Perspectives*, 26(4):43–56.

Heberlein, T., Wilson, M., Bishop, R., and Schaeffer, N. C. (2005). Rethinking the scope test as a criterion for validity in contingent valuation. *Journal of Environmental Economics and Management*, 50(1):1–22.

Hedlund, J. H., Curtis, P. D., Curtis, G., and Williams, A. F. (2004). Methods to reduce traffic crashes involving deer: What works and what does not. *Traffic Injury Prevention*, 5(2):122–131.

Hensher, D. A.and Rose, J. M., Ortúzar, J., and Rizzi, L. (2009). Estimating the willingness to pay and value of risk reduction for car occupants in the road environment. *Transportation Research Part A: Policy and Practice*, 43(7):692–707.

Herriges, J. A. and Shogren, J. F. (1996). Starting point bias in dichotomous choice valuation with follow-up questioning. *Journal of Environmental Economics and Management*, 30(1):112–131.

Hole, A. R. (2007). A comparison of approaches to estimating confidence intervals for willingness to pay measures. *Health Economics*, 16(8):827–840.

Hörnberg, S. (2001). Changes in population density of moose (alces alces) and damage to forests in Sweden. *Forest Ecology and Management*, 149(1–3):141 – 151.

Huijser, M. P., Duffield, J. W., Clevenger, A. P., Ament, R. J., and McGowen, P. T. (2009a). Cost-benefit analyses of mitigation measures aimed at reducing collisions with large ungulates in the United States and Canada; a decision support tool. *Ecology and Society*, 14(2):15.

Huijser, M. P., Duffield, J. W., Clevenger, A. P., Ament, R. J., and McGowen, P. T. (2009b). Cost-benefit analyses of mitigation measures aimed at reducing collisions with large ungulates in the united states and canada: a decision support tool. *Ecology and Society*, 14(2):15.

Huijser, M. P., McGowen, P. T., Clevenger, A. P., and Ament, R. (2008). *Wildlife-vehicle collision reduction study: best practices manual*. Available at http://environment.fhwa.dot.gov/ecosystems/wvc/index.asp.

Hultkrantz, L., Lindberg, G., and Andersson, C. (2006). The value of improved road safety. *Journal of Risk and Uncertainty*, 32(2):151–170.

Hung, H.-C. (2009). The attitude towards flood insurance purchase when respondents' preferences are uncertain: A fuzzy approach. *Journal of Risk Research*, 12(2):239–258.

Jeanty, P. W. (2007). Constructing Krinsky and Robb confidence interval for mean and median WTP using Stata. North American Stata Users' Group Meetings 2007 8, Stata Users Group, revised 30 Aug 2007.

Jeanty, P. W., Haab, T. C., and Hitzhusen, F. J. (2007). Willingness to pay for biodiesel in diesel engines: A stochastic double bounded contingent valuation survey. 2007 Annual Meeting, July 29-August 1, 2007, Portland, Oregon TN 9868, American Agricultural Economics Association (New Name 2008: Agricultural and Applied Economics Association).

Jenkins, G., Kuo, C.-Y., and Giraldez, J. (2007). Canadian Regulatory Cost-Benefit Analysis Guide. Development Discussion Papers 2007-03, JDI Executive Programs. Treasury Board Secretariat.

Jennings, S., Yamazaki, S., and Lyle, J. (2010). Testing heterogeneous anchoring and shift effect in double-bounded models: The case of recreational fishing in Tasmania. 2010 Conference (54th), February 10-12, 2010, Adelaide, Australia 59089, Australian Agricultural and Resource Economics Society.

Johannesson, M., Blomquist, G. C., Blumenschein, K., Johansson, P.-O., Liljas, B., and O'Conor, R. M. (1999). Calibrating hypothetical willingness to pay responses. *Journal of Risk and Uncertainty*, 18(1):21–32.

Johannesson, M. and Johansson, P.-O. (1996). To be, or not to be, that is the question: An empirical study of the wtp for an increased life expectancy at an advanced age. *Journal of Risk and Uncertainty*, 13(2):163–174.

Johannesson, M., Johansson, P.-O., and O'Conor, R. (1996). The value of private safety versus the value of public safety. *Journal of Risk and Uncertainty*, 13(3):263–275.

Johannesson, M., Liljas, B., and Johansson, P. (1998). An experimental comparison of dichotomous choice contingent valuation questions and real purchase decisions. *Applied Economics*, 30(5):643–647.

Johansson, P.-O. (2002). On the definition and age-dependency of the value of a statistical life. *Journal of Risk and Uncertainty*, 25(3):251–263.

Johansson-Stenman, O. and Martinsson, P. (2008). Are some lives more valuable? an ethical preferences approach. *Journal of Health Economics*, 27(3):739–752.

Jones, N., Sophoulis, C. M., and Malesios, C. (2008). Economic valuation of coastal water quality and protest responses: A case study in Mitilini, Greece. *Journal of Socio-Economics*, 37(6):2478–2491.

Jones-Lee, M. (1974). The value of changes in the probability of death or injury. *Journal of Political Economy*, 82(4):835–849.

Jones-Lee, M. (1991). Altruism and the value of other people's safety. *Journal of Risk and Uncertainty*, 4(2):213–219.

Jones-Lee, M. and Loomes, G. (1995). Scale and context effects in the valuation of transport safety. *Journal of Risk and Uncertainty*, 11(3):183–203.

Jones-Lee, M. W. (1976). *The Value of Life: An Economic Analysis*. University of Chicago Press, Chicago, IL.

Jones-Lee, M. W. (1979). Trying to value a life. Why broome does not sweep clean. *Journal of Public Economics*, 12(2):249–256.

Jones-Lee, M. W. (2004). Valuing internationla safety externalities: Does the 'golden rule' apply? *Journal of Risk and Uncertainty*, 29(3):277–287.

Jones-Lee, M. W., Hammerton, M., and Philips, P. R. (1985). The value of safety: Results of a national sample survey. *The Economic Journal*, 95(377):49–72.

Jones-Lee, M. W., Loomes, G., and Philips, P. R. (1995). Valuing the prevention of non-fatal road injuries: Contingent valuation vs. standard gambles. *Oxford Economic Papers*, 47(4):676–695.

Jorgensen, B. S. and Syme, G. J. (2000). Protest responses and willingness to pay: Attitude toward paying for stormwater pollution abatement. *Ecological Economics*, 33(2):251–265.

Jorgensen, B. S., Syme, G. J., Bishop, B. J., and Nancarrow, B. E. (1999). Protest responses in contingent valuation. *Environmental and Resource Economics*, 14(1):131–150.

Jorgensen, B. S., Wilson, M. A., and Heberlein, T. A. (2001). Fairness in the contingent valuation of environmental public goods: Attitude toward paying for environmental improvements at two levels of scope. *Ecological Economics*, 36(1):133–148.

Joyce, T. L. and Mahoney, S. P. (2001). Spatial and temporal distributions of moose-vehicle collisions in newfoundland. *Wildlife Society Bulletin*, 29(1):281–291.

Kahneman, D. (1986). Comments. In Cummings, R. G., Brookshire, D. S., and Schulze, W. D., editors, *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*, pages 185–194. Rowman and Allanheld, Totowa, NJ.

Kahneman, D. and Knetsch, J. L. (1992). Valuing public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management*, 22(1):57–70.

Kanninen, B. (1993). Optimal experimental design for double-bounded dichotomous choice contingent valuation. *Land Economics*, 69(2):138–146.

Kanninen, B. (1995). Bias in discrete response contingent valuation. *Journal of Environmental Economics and Management*, 28(1):114–125.

Kaplow, L. (2005). The value of a statistical life and the coefficient of relative risk aversion. *Journal of Risk and Uncertainty*, 31(1):23–34.

Karni, E. (1983). Risk aversion for state-dependent utility functions: Measurement and applications. *International Economic Review*, 24(3):637–647.

Kasperson, R. E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., Kasperson, J. X., and Ratick, S. (1988). The social amplification of risk: A conceptual framework. *Risk Analysis*, 8(2):177–187.

Kip Viscusi, W., Magat, W. A., and Huber, J. (1991). Pricing environmental health risks: survey assessments of risk-risk and risk-dollar trade-offs for chronic bronchitis. *Journal of Environmental Economics and Management*, 21(1):32–51.

Knetsch, J. L. and Sinden, J. A. (1984). Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *The Quarterly Journal of Economics*, 99(3):507–521.

Kniesner, T. J., Viscusi, W. K., and Ziliak, J. (2010). Policy relevant heterogeneity in the value of statistical life: New evidence from panel data quantile regressions. *Journal of Risk and Uncertainty*, 40(1):15–31.

Kochi, I., Hubbell, B., and Kramer, R. (2006). An empirical bayes approach to combining and comparing estimates of the value of a statistical life for environmental policy analysis. *Environmental & Resource Economics*, 34(3):385–406.

Krinsky, I. and Robb, A. L. (1986). On approximating the statistical properties of elasticities. *The Review of Economics and Statistics*, 68(4):715–719.

Krinsky, I. and Robb, A. L. (1990). On approximating the statistical properties of elasticities: A correction. *The Review of Economics and Statistics*, 72(1):189–190.

Krupnick, A. (2007). Mortality-risk valuation and age: Stated preference evidence. *Review of Environmental Economics and Policy*, 1(2):261–282.

Krupnick, A., Alberini, A., Cropper, M., Simon, N., O'Brien, B., Goeree, R., and Heintzelman, M. (2002). Age, health and the willingness to pay for mortality risk reductions: A contingent valuation survey of ontario residents. *Journal of Risk and Uncertainty*, 24(2):161–186.

Krutilla, J. V. (1967). Conservation reconsidered. *The American Economic Review*, 57(4):777–786.

Lanoie, P., Pedro, C., and Latour, R. (1995). The value of a statistical life: A comparison of two approaches. *Journal of Risk and Uncertainty*, 10(3):235–257.

Lavsund, S. and Sandegren, F. (1991). Moose-vehicle relations in Sweden: A review. *Alces*, 27(1):118Ű126.

Leblond, M., Dussault, C., Ouellet, J.-P., Poulin, M., Courtois, R., and Fortin, J. (2007). Electric fencing as a measure to reduce moose-vehicle collisions. *Journal of Wildlife Management*, 71(5):1695–1703.

Leiter, A. M. and Pruckner, G. J. (2009). Proportionality of willingness to pay

to small changes in risk: The impact of attitudinal factors in scope tests. *Environmental and Resource Economics*, 42(2):169–186.

León, R. and León, C. J. (2003). Single or double bounded contingent valuation? A Bayesian test. *Scottish Journal of Political Economy*, 50(2):174–188.

Li, C.-Z. and Mattsson, L. (1995). Discrete choice under preference uncertainty: An improved structural model for contingent valuation. *Journal of Environmental Economics and Management*, 28(2):256–269.

Li, H., Jenkins-Smith, H. C., Silva, C. L., Berrens, R. P., and Herron, K. G. (2009). Public support for reducing US reliance on fossil fuels: Investigating household willingness-to-pay for energy research and development. *Ecological Economics*, 68(3):731–742.

Lichtenstein, S., Fischhoff, B., Layman, M., and Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning & Memory*, 4(6):551–578.

Lindhjem, H. and Navrud, S. (2009). Asking for individual or household willingness to pay for environmental goods?: Implication for aggregate welfare measures. *Environmental and Resource Economics*, 43(1):11–29.

Lindhjem, H., Navrud, S., Braathen, N., and Biausque, V. (2011). Valuing mortality risk reductions from environmental, transport, and health policies: A global meta-analysis of stated preference studies. *Risk Analysis*, 31(9):1381–1407.

Lindsey, G. (1994). Market models, protest bids and outliers in contingent valuation. *Journal of Water Resources Planning and Management*, 120(1):121–129.

List, J. and Gallet, C. (2001). What experimental protocol influence disparities between actual and hypothetical stated values? *Environmental and Resource Economics*, 20(3):241–254.

Loomis, J. (2011). What's to know about hypothetical bias in stated preference valuation studies? *Journal of Economic Surveys*, 25(2):363–370.

Loomis, J. B. and Ekstrand, E. (1997). Economic benefits of critical habitat for the Mexican spotted owl: A scope test using a multiple-bounded contingent valuation survey. *Journal of Agricultural and Resource Economics*, 22(2):356–366.

Loomis, J. B. and Ekstrand, E. (1998). Alternative approaches for incorporating respondent uncertainty when estimating willingness to pay: The case of the Mexican Spotted Owl. *Ecological Economics*, 27(1):29–41.

Ludwig, J. and Cook, P. J. (2001). The benefits of reducing gun violence: Evidence from contingent-valuation survey data. *Journal of Risk and Uncertainty*, 22(3):207–226.

Lyssenko, N. and Martínez-Espiñeira, R. (2012a). 'Been there done that': Disentangling option value effects from user heterogeneity when valuing natural resources with a use component. *Environmental Management*, 50(5):819–836.

Lyssenko, N. and Martínez-Espiñeira, R. (2012b). Respondent uncertainty in contingent valuation: The case of whale conservation in Newfoundland and Labrador. *Applied Economics*, 44(15):1911–1930.

López-Feldman, A. (2010). Doubleb: Stata module to compute contingent valuation using double-bounded dichotomous choice. Statistical Software Components, Boston College Department of Economics.

López-Feldman, A. (2011). Singleb: Stata module to compute contingent valuation using single-bounded dichotomous choice. Statistical Software Components, Boston College Department of Economics.

López-Feldman, A. (2012). Introducción a la valoración contingente utilizando Stata. In *Aplicaciones en Economía y Ciencias Sociales con Stata*. Stata Press, College Station, TX.

MacKinnon, B. (1986). Pricing human life. *Science, Technology, & Human Values*, 11(2):29–39.

Marris, C., Langford, I., and O'Riordan, T. (1996). Integrating sociological and psychological approaches to public perceptions of environmental risks: Detailed results from a questionnaire survey. CSERGE Working Paper GEC 96-07.

Martínez-Espiñeira, R. (2006). A Box-Cox double-hurdle model of wildlife valuation: The citizen's perspective. *Ecological Economics*, 58(1):192–208.

Martínez-Espiñeira, R. and Lyssenko, N. (2011). Correcting for the endogeneity

of pro-environment behavioral choices in contingent valuation. *Ecological Economics*, 70(8):1435–1439.

Martínez-Espiñeira, R. and Lyssenko, N. (2012). Alternative approaches to dealing with respondent uncertainty in contingent valuation: A comparative analysis. *Journal of Environmental Management*, 93(1):130–139.

Mastro, L. L., Conover, M. R., and Frey, S. N. (2008). Deer–vehicle collision prevention techniques. *Human–Wildlife Confllicts*, 2(1):80–92.

Matthews, M. L. and Moran, A. R. (1986). Age differences in male drivers' perception of accident risk: The role of perceived driving ability. *Accident Analysis & Prevention*, 18(4):299 – 313.

McFadden, D. (1994). Contingent valuation and social choice. *American Journal of Agricultural Economics*, 76(4):689–708.

McFadden, D. and Leonard, G. (1995). Issues in the contingent valuation of environmental goods: methodologies for data collection and analysis. In *Contingent Valuation: A Critical Assessment*. North-Holland, Amsterdam.

McLaren, B. E. (2002). Social and natural history of moose introduced to Newfoundland. In *Lessons from the Islands: introduced species and what they tell us about how ecosystems work. Proceedings from the Research Group on Introduced Species 2002 Symposium, Queen Charlotte City, Queen Charlotte Islands, British Columbia.* Canadian Wildlife Service, Environment Canada, Ottawa.

McLaren, B. E., Mahoney, S. P., Porter, T. S., and Oosenbrug, S. M. (2000). Spatial and temporal patterns of use by moose of pre-commercially thinned, naturally-regenerating stands of balsam fir in central newfoundland. *Forest Ecology and Management*, 133(3):179 – 196.

McLaren, B. E., Roberts, B. A., Djan-Chékar, N., Lewis, K. P., et al. (2004). Effects of overabundant moose on the newfoundland landscape. *Alces*, 40:45–59.

McLeod, D. and Bergland, O. (1999). Willingness-to-pay estimates using the double-bounded dichotomous-choice contingent valuation format: A test for validity and precision in a bayesian framework. *Land Economics*, 75(1):115–125.

Meyerhoff, J. and Liebe, U. (2006). Protest beliefs in contingent valuation: Explaining their motivation. *Ecological Economics*, 57(4):583–594.

Milgrom, P. (1993). Is sympathy an economic value? philosophy, economics, and the contingent valuation method. In Hausman, J. A., editor, *Contingent Valuation: A Critical Assessment*, pages 417–441. North-Holland, Amsterdam:.

Miller, T. R. (2000). Variations between countries in values of statistical life. *Journal of Transport Economics and Policy*, 34(2):169–188.

Miller, T. R., Rossman, S., and Viner, J. (1991). Costs of highway crashes. Technical report, The Urban Institute. (Also, Publication FHWA-RD-91-055, Federal Highway Administration, McLean VA.).

Mishan, E. J. (1971). Evaluation of life and limb: A theoretical approach. *The Journal of Political Economy*, 79(4):687–705.

Mishan, E. J. (1982). *Cost-Benefit Analysis : An Informal Introduction*. G. Allen & Unwin, London.

Mitchell, R. C. and Carson, R. T. (1989). *Using Surveys to Value Public Goods*. Johns Hopkins University for Resources for the Future, Baltimore.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, 26(1):67–82.

Moore, R., Bishop, R. C., Provencher, B., and Champ, P. A. (2010). Accounting for respondent uncertainty to improve willingness-to-pay estimates. *Canadian Journal of Agricultural Economics*, 58(3):381–401.

Morrison, M. D., Blamey, R. K., and Bennett, J. W. (2000). Minimising payment vehicle bias in contingent valuation studies. *Environmental and Resource Economics*, 16(4):407–422.

Mrozek, J. R. and Taylor, L. O. (2002). What determines the value of life? A meta-analysis. *Journal of Policy Analysis and Management*, 21(2):253–270.

Murphy, A. (2007). Score tests of normality in bivariate probit models. *Economics Letters*, 95(3):374–379.

Murphy, J., Allen, P., Stevens, T., and Weatherhead, D. (2005). A meta-analysis

of hypothetical bias in stated preference valuation. *Environmental and Resource Economics*, 30(3):313–325.

Needleman, L. (1976). Valuing other people's lives. *The Manchester School of Economic and Social Studies*, 44(4):309–342.

Nielsen, J., Gyrd-Hansen, D., and Kjaer, T. (2012). Valuation of morbidity and mortality risk reductions. does context matter? *Accident Analysis and Prevention*, 48:246–253.

Nielsen, J. and Kjær, T. (2011). Does question order influence sensitivity to scope? Empirical findings from a web-based contingent valuation study. *Journal of Environmental Planning and Management*, 54(3):369–381.

Norwood, F., Lusk, J., and Boyer, T. (2008). Forecasting hypothetical bias: A tale of two calibrations. In Cherry, T., Kroll, S., and Shogren, J., editors, *Environmental Economics, Experimental Methods*, pages 447–465. Routledge Press, New York, NY.

Ojea, E. and Loureiro, M. L. (2008). Consumers versus non-consumers' benefits from recovering an overexploited fish stock. Paper presented at the 10th BIOECON Conference Sidney Sussex College Cambridge September 28-30.

Olaussen, J. O. and Skonhoft, A. (2011). A cost-benefit analysis of moose harvesting in scandinavia. a stage structured modelling approach. *Resource and Energy Economics*, 33(3):589–611.

Olofsson, A. and Rashid, S. (2011). The white (male) effect and risk perception: Can equality make a difference? *Risk Analysis*, 31(6):1016–1032.

Parsons, G. R. (2003). The travel cost model. In Champ, P. A., Boyle, K. J., and Brown, T. C., editors, *A Primer on Nonmarket Valuation*, chapter 9. Kluwer Academic Publishing, London.

Pelton, M., F. vanManen, A. C., Weaver, K., Pederson, J., and Eason, T. (1994). Black bear conservation action plan – north america. Technical report, IUCN/SSC Bear Specialist Group.

Persson, U., Norinder, A., Hjalte, K., and Gralén, K. (2001). The value of a statistical life in transport: Findings from a new contingent valuation study in Sweden. *Journal of Risk and Uncertainty*, 23(2):121–134.

Pfarr, C., Schmid, A., and Schneider, U. (2010). Estimating ordered categorical variables using panel data: a generalized ordered probit model with an autofit procedure. MPRA Paper 23203, University Library of Munich, Germany.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32:122–136.

Pratt, J. W. and Zeckhauser, R. J. (1996). Willingness to pay and the distribution of risk and wealth. *Journal of Political Economy*, 104(4):pp. 747–763.

Putman, R. (1997). Deer and road traffic accidents: Options for management. *Journal of Environmental Management*, 51(1):43–57.

Rattey, T. E. and Turner, N. (1991). Vehicle-moose accidents in newfoundland. *The Journal of Bone and Joint Surgery*, 73(10):1487–1491.

Rea, R. V. (2003). Modifying roadside vegetation management practices to reduce vehicular collisions with moose alces alces. *Wildlife Biology*, 9(2):81–91.

Ready, R. C., Navrud, S., and Dubourg, W. R. (2001). How do respondents with uncertain willingness to pay answer contingent valuation questions? *Land Economics*, 77(3):315–326.

Ready, R. C., Whitehead, J. C., and Blomquist, G. C. (1995). Contingent valuation when respondents are ambivalent. *Journal of Environmental Economics and Management*, 29(2):181–196.

Reed, D. F., Beck, T. D. I., and Woodward, T. N. (1982). Methods of reducing deer-vehicle accidents: Benefit-cost analysis. *Wildlife Society Bulletin*, 10:349–354.

Riddel, M. (2001). Simulated maximum likelihood for double-bounded referendum models. *Journal of Agricultural and Resource Economics*, 26(2):491–507.

Romin, L. and Bissonette, J. A. (1996). Deer-vehicle collisions: Status of state monitoring activities and mitigation efforts. *Wildlife Society Bulletin*, 24(2):276–283.

Roodman, D. (2011). Fitting fully observed recursive mixed-process models with cmp. *The Stata Journal*, 11(2):159–206.

Royston, P. (2004). Multiple imputation of missing values. *Stata Journal*, 4(3):227–241.

Royston, P. (2005a). Multiple imputation of missing values: Update. *The Stata Journal*, 5(2):188–201.

Royston, P. (2005b). Multiple imputation of missing values: Update of ICE. *The Stata Journal*, 5(4):527–536.

Royston, P. (2008). ICE: Stata module for multiple imputation of missing values. *Statistical Software Components*, S446602. Statistical Software Components. Boston College Department of Economics, revised 25 Apr 2008.

Royston, P., Carlin, J. B., and White, I. R. (2009). Multiple imputation of missing values: New features for mim. *Stata Journal*, 9(2):252–264.

Samnaliev, M., Stevens, T. H., and More, T. (2006). A comparison of alternative certainty calibration techniques in contingent valuation. *Ecological Economics*, 57(3):507–519.

Samuelson, P. A. (1954). The pure theory of public expenditure. *The Review of Economics and Statistics*, 36(4):387–389.

Samuelson, P. A. (1955). Diagrammatic exposition of a theory of public expenditure. *The Review of Economics and Statistics*, 37(4):350–356.

Savage, I. (1993). Demographic influences on risk perceptions. *Risk Analysis*, 13(4):413–420.

Schelling, T. C. (1968). The life you save may be your own. In *Problems in Public Expenditure Analysis*, pages 127–162. The Brookings Institution, Washington, D.C.

Schwab Christe, N. G. and Soguel, N. C. (1996). The pain of road-accident victims and the bereavement of their relatives: a contingent-valuation experiment. *Journal of Risk and Uncertainty*, 13(3):277–291.

Schwartz, N. (1997). Cognition, communication, and survey measurement: Some implications for contingent valuation surveys. In Kopp, R. J., Pommerhene, W. W., and Schwartz, N., editors, *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*. Kluwer, Boston.

Schwarzinger, M., Carrat, F., and Luchini, S. (2009). "If you have the flu symptoms, your asymptomatic spouse may better answer the willingness-to-pay question". Evidence from a double-bounded dichotomous choice model with heterogeneous anchoring. *Journal of Health Economics*, 28(4):873–884.

Seiler, A. (2005). Predicting locations of moose-vehicle collisions in sweden. *Journal of Applied Ecology*, 42(2):371–382.

Shaikh, S. L., Sun, L., and Kooten, G. C. V. (2007). Treating respondent uncertainty in contingent valuation: A comparison of empirical treatments. *Ecological Economics*, 62(1):115–125.

Shepard, D. S. and Zeckhauser, R. J. (1984). Survival versus consumption. *Management Science*, 30(4):423–439.

Shiell, A. and Gold, L. (2002). Contingent valuation in health care and the persistence of embedding effects without the warm glow. *Journal of Economic Psychology*, 23(2):251–262.

Short Jr, J. F. (1984). The social fabric of risk: towards the social transformation of risk analysis. *American Sociological Review Review*, 49(6):711–725.

SIKA (2002). Review of cost benefit calculations. methods and valuations in the transport sector. Technical report, The Swedish Institute for Transport and Communications Analysis, SIKA. SIKA Report 2002:4.

Sivak, M., Soler, J., Tränkle, U., and Spagnhol, J. M. (1989). Cross-cultural differences in driver risk-perception. *Accident Analysis and Prevention*, 21(4):355–362.

Sjöberg, L. (2000). Factors in risk perception. *Risk Analysis*, 20(1):1–11.

Sjöberg, L., Moen, B.-E., and Rundmo, T. (2004). Explaining risk perception. An evaluation of the psychometric paradigm in risk perception research. Technical report, Rotunde publikasjoner. Rotunde no. 84.

Sjöberg, L., Kolarova, D., Rucai, A.-A., Bernström, M.-L., and Flygelholm, H. (1996). Risk perception and media reports in Bulgaria and Romania. Technical report, Stockholm: Center for Risk Research.

Slovic, P., Fischhoff, B., and Lichetenstein, S. (1985). Charachterizing perceived risk. In Kates, C. H. C. and Kasperson, J. X., editors, *Perilous Progress: Managing the Hazards of Technology*. Westview Press, Boulder, CO.

Smith, V. K. and Osborne, L. L. (1996). Do contingent valuation estimates pass a "scope" test? a meta-analysis. *Journal of Environmental Economics and Management*, 31(3):287–301.

Stantec (2010). Labrador Ű island transmission link: Moose and black bear component study. Technical report, Stantec Consulting.

Statacorp (2011). *Stata Statistical Software: Release 12*. StataCorp LP, College Station, TX.

StatsCanada (2009). Mortality, summary list of causes, 2009. Statistics Canada. Catalogue no. 84F0209X.

Stevens, T. H., DeCoteau, N. E., and Willis, C. E. (1997). Sensitivity of contingent valuation to alternative payment schedules. *Land Economics*, 73(1):140–148.

Strazzera, E., Genius, M., Scarpa, R., and Hutchinson, G. (2003a). The effect of protest votes on the estimates of WTP for use values of recreational sites. *Environmental and Resource Economics*, 25(4):461.

Strazzera, E., Scarpa, R., Calia, P., Garrod, G., and Willis, K. (2003b). Modeling zero values and protest responses in contingent valuation surveys. *Applied Economics*, 35(2):133–138.

Sullivan, T. L. and Messmer, T. A. (2003). Perceptions of deer-vehicle collision management by state wildlife agency and department of transportation administrators. *Wildlife Society Bulletin*, 31(1):163–173.

Sund, B. (2009). Certainty calibration in contingent valuation - exploring the within-difference between dichotomous choice and open-ended answers as a certainty measure. Swedish Business School, Örebro University, Sweden.

Sunstein, C. R. (2004a). Lives, life-years, and willingness to pay. *Columbia Law Review*, 104(1):205–252.

Sunstein, C. R. (2004b). Valuing life: A plea for disaggregation. *Duke Law Journal*, 54(2):385–445.

Svedsäter, H. (2000). Contingent valuation of global environmental resources: Test of perfect and regular embedding. *Journal of Economic Psychology*, 21(6):605–623.

Svensson, M. (2009). The value of a statistical life in sweden: Estimates from two studies using the "certainty approach" calibration. *Accident Analysis & Prevention*, 41(3):430–437.

Svensson, M. and Vredin Johansson, M. (2010). Willingness to pay for private and public road safety in stated preference studies: Why the difference? *Accident Analysis and Prevention*, 42(4):1205–1212.

Taylor, L. O. (2003). The hedonic method. In Champ, P. A., Boyle, K. J., and

Brown, T. C., editors, *A Primer on Nonmarket Valuation*, pages 331–387. Kluwer Academic Publishers, Netherlands.

Thompson, I. D., Larson, D. J., and Montevecchi, W. A. (2003). Characterization of old" wet boreal" forests, with an example from balsam fir forests of western newfoundland. *Environmental Reviews*, 11(S1):S23–S46.

Timmermann, H. R. and Rodgers, A. R. (2005). Moose: Competing and complementary values. *Alces*, 41:85–120.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26:24–36.

Train, K. E. (2003). *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge.

Transport Canada (1994). Guide to Benefit-Cost Analysis in Transport Canada. Technical report. Transport Canada Report, TP11875E, September.

Transport Canada (2003). Overview of technologies aimed at reducing and preventing large animal strikes. Technical report, Standards Research and Development Branch. Road Safety and Motor Vehicle Regulation Directorate. Transport Canada, Ottawa, Ontario.

Transportation Safety Board of Canada (2013). Assessment of the responses to aviation safety recommendation a06-08. Value of statistical life. Available at http://www.tsb.gc.ca/eng/recommandations-recommendations/aviation/2006/rec_a0608.pdf.

Trenholm, R., Lantz, V., and Little, S. (2013). A benefit-cost analysis of riparian preservation in an Eastern Canadian watershed. *Journal of Environmental Management*, 116(15):81–94.

Trumbull, W. N. (1990). Who has standing in cost-benefit analysis? *Journal of Policy Analysis and Management*, 9(2):201–218.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 4:207–232.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185:1124–1130.

USDOT (2011). Treatment of the economic value of a statistical life in departmental analyses. 2011 interim adjustment. US Department of Transportation Memorandum July 29 2011.

Van Kooten, G. C., Krcmar, E., and Bulte, E. (2001). Preference uncertainty in non-market valuation: A fuzzy approach. *American Journal of Agricultural Economics*, 83(3):487–500.

Veisten, K., Flügel, S., Rizzi, L. I., Ortúzar, J. D. D., and Elvik, R. (2013). Valuing casualty risk reductions from estimated baseline risk. *Research in Transportation Economics*, 43(1):50–61.

Verbeek, M. (2008). *A Guide to Modern Econometrics*. Wiley, Sussex, 3 edition.

Viscusi, W. (2010). The heterogeneity of the value of statistical life: Introduction and overview. *Journal of Risk and Uncertainty*, 40(1):1–13.

Viscusi, W. K. (1992). *Fatal tradeoffs: public and private responsibilities for risk*. Oxford University Press, Oxford, UK.

Viscusi, W. K. (2013). The benefits of mortality risk reduction: Happiness surveys vs. the value of a statistical life. *Duke Law Journal*, 62:1735–1745.

Viscusi, W. K. and Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, 27(1):5–76.

Viscusi, W. K., Hakes, J. K., and Carlin, A. (1997). Measures of Mortality Risks. *Journal of Risk and Uncertainty*, 14:213–233.

Viscusi, W. K. and Magat, W. A. (1987). *Learning about Risk: Consumer and Worker Response to Hazard Warnings*. Harvard University Press, Cambridge, MA, USA.

Vossler, C. A., Ethier, R. G., Poe, G. L., and Welsh, M. P. (2003). Payment certainty in discrete choice contingent valuation responses: Results from a field validity test. *Southern Economic Journal*, 69(4):886–902.

Walsh, R., Johnson, D., and McKean, J. (1992). Benefit transfer of outdoor recreation demand studies: 1968-1988. *Water Resources Research*, 28:707–713.

Wang, H. (1997). Treatment of "don't know" responses in contingent valuation surveys: A random valuation model. *Journal of Environmental Economics and Management*, 32(2):219–232.

Wang, H. and He, J. (2011). Estimating individual valuation distributions with multiple bounded discrete choice data. *Applied Economics*, 43(21):2641–2656.

Watson, V. and Ryan, M. (2007). Exploring preference anomalies in double bounded contingent valuation. *Journal of Health Economics*, 26(3):463–482.

Weimer, D. L. and Vining, A. R. (2005). *Policy Analysis: Concepts and Practice*. Pearson Prentice-Hall, Englewood Cliffs, N.J.

Weinstein, M. C., Shepard, D. S., and Pliskin, J. S. (1980). The economic value of changing mortality probabilities: A decision-theoretic approach. *The Quarterly Journal of Economics*, 94(2):pp. 373–396.

Weir, E. (2002). Collisions with wildlife: the rising toll. *Canadian Medical Association Journal*, 166(6):775–775.

Welsh, M. P. and Poe, G. L. (1998). Elicitation effects in contingent valuation: Comparisons to a multiple bounded discrete choice approach. *Journal of Environmental Economics and Management*, 36(2):170–185.

Whitehead, J. (2004). Incentive compatibility and starting point bias in iterative valuation questions: reply. *Land Economics*, 78(2):285–297.

Whitehead, J. C. (2002). Incentive compatibility and starting point bias in iterative valuation questions. *Land Economics*, 78(2):285–297.

Whitehead, J. C. (2005). Combining willingness to pay and behavior data with limited information. *Resource and Energy Economics*, 27(2):143–155.

Whitehead, J. C. (2006). Improving willingness to pay estimates for quality improvements through joint estimation with quality perceptions. *Southern Economic Journal*, 73(1):100–111.

Whitehead, J. C. and Cherry, T. L. (2007). Willingness to pay for a green energy program: A comparison of ex-ante and ex-post hypothetical bias mitigation approaches. *Resource and Energy Economics*, 29(4):247–261.

Whitehead, J. C., Morgan, A., Huth, W. L., Martin, G. S., and Sjolander, R. (2012). Willingness-to-pay for oyster consumption mortality risk reductions. Working Papers 12-07, Department of Economics, Appalachian State University.

Whittington, D. and MacRae, D. (1986). The issue of standing in cost-benefit analysis. *Journal of Policy Analysis and Management*, 5(4):665–682.

Wijnen, W., Wesemann, P., and de Blaeij, A. (2009). Valuation of road safety effects in cost-benefit analysis. *Evaluation and Program Planning*, 32(4):326–331.

Wildavsky, A. and Dake, K. (1990). Theories of risk perception: Who fears what and why? *Daedalus*, 119(4):41–60.

Williams, R. (2006). Generalized ordered logit/ partial proportional odds models for ordinal dependent variables. *The Stata*

Journal, 6(1):58–82. A pre-publication version is available at http://www.nd.edu/ rwilliam/gologit2/gologit2.pdf.

Williams, R. (2009). Using heterogeneous choice models to compare logit and probit coefficients across groups. *Sociological Methods and Research*, 37(4):531–559.

Williams, R. (2010). Fitting heterogeneous choice models with oglm. *The Stata Journal*, 10(4):540–567.

Wiseth, B. (1991). Moose-train collisions: the effects of vegetation removal with a cost-benefit analysis. *Alces*, 27:93–99.

Yatchew, A. and Griliches, Z. (1985). Specification error in probit models. *The Review of Economics and Statistics*, 67(1):134–39.

Zeckhauser, R. J. and Viscusi, W. K. (1990). Risk within reason. *Science*, 248(4955):559–564.

Zhang, A., Boardman, A. E., Gillen, D., and Waters II, W. G. (2005). Towards estimating the social and environmental costs of transportation in Canada. A report for Transport Canada. Technical report, Center for Transportation Studies, University of British Columbia.

Zhang, J., Adamowicz, W., Dupont, D. P., and Krupnick, A. (2013). Assessing the extent of altruism in the valuation of community drinking water quality improvements. *Water Resources Research*, 49(10):6286–6297.

Appendix A

Abbreviations

CAD: Canadian Dollar

CBA: Cost Benefit Analysis

CI: (95%) Confidence Interval

CPI: Consumer Price Index

CS: Consumer Surplus

CVM: Contingent Valuation Method

CV: Compensating Variation

DBDC: Double-Bounded Dichotomous Choice

DK: 'don't know/no response'

DOT: (US) Department of Transportation

DRE: death risk equivalent

EV: Equivalent Variation

FHWA: (US Department of Transportation) Federal Highway Administra-

tion

GPS: Geographic Positioning System

MMA: Moose management area

MVC: Moose-Vehicle Collision

NL: Newfoundland and Labrador

NOK: Norwegian Kronor

NCS: Numerical certainty scale

NOOA: (United States) National Oceanic and Atmospheric Administration

OLS: Ordinary Least Squares

QALY: Quality-Adjusted Life Year

SBDC: Single-Bounded Dichotomous Choice

LB: Lower Bound (of Confidence Interval)

SUV: Sport-Utility Vehicle

TBS: Treasury Board Secretariat

UB: Upper Bound (of Confidence Interval)

USD: US dollar

USDOT: United States Department of Transportation

VSL: Value of Statistical Life

VOLY: Value of Life Year

WTA: Willingness to Accept

WTP: Willingness to Pay

Appendix B

Modeling of risk perception

As reported in Chapters 8 and 10, we were interested in examining the differences in terms of risk perception held by different respondents. After providing them with estimates of baseline "average" risk rates for the whole province (variables RM and RI), we asked them, therefore, to provide us with their own estimate of their individual level of risk, taking into account their particular circumstances.[1]

We used a conditional modelling process approach (Roodman, 2011), in order to account for the likely correlation between the two values provided (variables $Q12$ and $Q13$). This tool makes it possible to model potentially correlated variables regardless of the type of variable. In our case, we treated both variables as continuous, since although they allow only for positive values, their averages were relatively large.

---

[1] The full text of Questions $Q12$ and $Q13$ is available in Appendix E.

Our results, reported in Table B.1, show that indeed there is a significant degree of correlation between the errors of the two equations involved. This means that, even after we controlled for the observable variables included in the models, there are unobservable factors that affect the self-perceived rates of mortality and morbidity in a similar way for a given respondent.

We can see that only a few variables have a significant effect on the level of perceived own risk. As expected, the variable *RM* influences very strongly both variables *Q12* and *Q13*. This is because *RM* represents the objective measure of risk we suggested in the questionnaire before asking each respondent to come up with their own estimate about their individual level of risk.[2] Additionally, having experienced encounters with moose, also as expected, increase the level of perceived own risk (variable *hitmoose*, in particular, is highly significant, variable *seenmoose* less so).

We find a significantly negative effect of gender (variable *male)*, corresponding to the so-called *white male effect*, only in the case of the equation for the mortality risk. Similarly, we found weakly significant effects for other variables that would in principle be expected to related to the level of risk faced by the respondent (variables about smoking status, driving habits, etc.). However, there are no other variables with a consistent effect on both types of perceived risk. While this preliminary attempt at modelling self-perceived risk levels is somewhat disappointing, we expect that the additional observations collected

---

[2]The variable measuring the morbidity risk counterpart was not included in the model, because it would lead to exact multicollinearity. This is because it was simply constructed as 30 times RM.

343

Table B.1: Conditional modelling process of self-perceived risk variables Q12 and Q13 (N=968).

	Q12	Q13
RM	0.5083**	0.2881**
male	-1.7120*	1.6460
ln(age)	-1.9637	-45.3429*
income	-0.0360	3.4151
huntedmoose	-0.1664	-20.6844
college	-0.6120	12.6314
hitmoose	1.4857*	36.1069**
NLander	-0.8338	25.3695$^+$
seenmoosecross	1.2674$^+$	28.5122*
newcar	-0.0290	14.0161
smoker	-1.7929*	-6.2331
smoker×male	4.1955*	
health	-0.0220	-0.2874
mathscore	-0.6500$^+$	-14.8015
childrennumber	0.0407	18.5228*
childrenany	0.2382	-21.4743
SUV	1.4729*	-3.1119
verysure (howsure>6)	-1.1769	19.6502$^+$
drives30towork	0.5741	46.6628*
constant	12.6029$^+$	144.6681
$ln(\sigma_1)$	2.1734**	
$ln(\sigma_2)$	5.0931**	
atanhrho$_{12}$	0.3265**	
log-likelihood	-9072.65	
ρ_{12}	0.3137**	

$^+p < 0.10$, $^*p < 0.05$, $^{**}p < 0.01$

through the planned second phase of the fieldwork will allow us to unveil further relationships between the level perceived risk and the individual circumstances of our respondents.

Appendix C

Modeling of respondent uncertainty

In this appendix, we show the results of trying to model the values of the respondents' numerical certainty scale (variable *howsure*), which measures how confident they were about their answers to the payment questions. In Table C.1 we show the percentual frequency distribution of variable *howsure* cross-tabulated with the values of variable *depvar*, that is, the polytomous variable that indicates the pattern of responses to the payment question and its follow-up in our DBDC payment format. We can see that, for the combination of responses NN and YY, around half of the responses show the highest level of certainty (*howsure*=10). In fact, only about 3% of YY respondents had less than 5 in their certainty scale, and only about 15% in the case of NN respondents.

The counterparts for the NY and YN respondents are 24% and 17%. When we just consider the values of *howsure* 8, 9, and 10, we can see that the proportions of respondents they encompass are much higher for the cases of the NN (70%) and YY (69%) patterns than for the cases of the YN (28%) and NY (43%) patterns. This confirms the notion that, after responding to the second payment question differently than to the first one, respondents face in general much more uncertainty about their responses.[1] This is not surprising, since answering NY or YN means that their true maximum WTP lies in between the two bids presented to them, so, particularly for low bid values (low values of variable COST), they will have faced a close proximity between their maximum WTP and the offered bids. This means, paradoxically, that the observations that, technically, offer the most information to the researcher about the respondents' WTP are in a sense the least reliable, because the respondents are more ambivalent about whether to answer 'yes' or 'no' to the proposed bid. Similarly, it could be shown that for the observations with a higher value of *howsure* the spread of the WTP is larger too, confirming one of the competing hypothesis about the effect of bid size on response variability (Alberini et al., 1997).[2]

We show in Figures C.1 and C.2 the distribution of variable *howsure* in graphical terms, first in the case of the whole sample and then only for the case of the NY and YN responses. Figure C.1 reflects the high degree of skewness in the distribution given by the fact that respondents are very certain about

---

[1] In terms of just the mean of the distribution, we find that the mean howsure for YY and NN respondents (8.03) is significantly higher than for NY and YN ones (6.47).

[2] The heteroskedasticity of the WTP model will be explored fully in an extension to the analysis presented in this report.

Table C.1: Distribution of values of numerical certainty scale (*howsure*), by response to the payment question (%).

	depvar				
howsure	No-No	No-Yes	Yes-No	Yes-Yes	Total
1	10.03	7.08	5.16	0.52	5.28
2	2.29	4.42	3.17	0.52	2.09
3	1.43	4.42	3.17	1.56	2.18
4	1.43	7.96	5.56	0.78	2.82
5	8.02	25.66	19.05	10.65	13.28
6	3.72	12.39	6.75	6.23	6.19
7	3.15	9.73	13.89	10.65	8.92
8	9.46	7.96	16.67	17.66	13.83
9	6.3	3.54	4.76	8.57	6.46
10	54.15	16.81	21.83	42.86	38.94
	100	100	100	100	100

their uncensored (in the language of the interval model) responses NN and YY. On the other hand, Figure C.2 shows that for responses NY and YN the distribution of variable *howsure* is much closer to a normal distribution. this is because, although in general most of our respondents said to be very certain about their answers, in the censored cases we have a relatively high proportion of intermediate values of *howsure*. In particular the mode for NY cases is 26% for *howsure* = 5 (followed by 17% for *howsure*=10), while for YN the 19% proportion of *howsure*=5 is a close second for the 22% for *howsure*=10. Indeed, a normality test[3] clearly rejects the null hypothesis in the case of the full sample, while that null hypothesis cannot be rejected (Prob>z = 0.14248) for the case of the censored responses (NY and YN).

In general, the great majority of respondents chose to state a value of 10 for *howsure* about their answers to the payment questions. However, the values

_____

[3] We used a Shapiro-Wilk W normality test in this case.
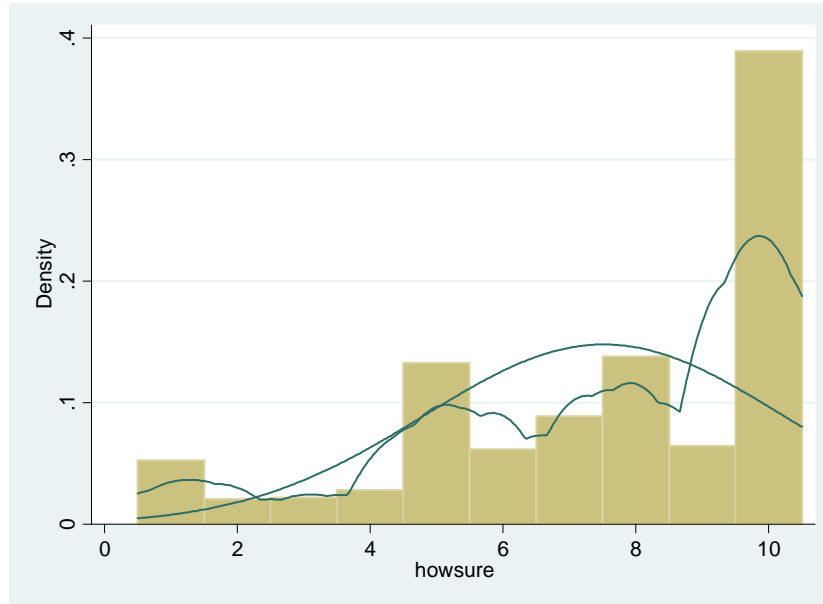
Figure C.1: Histogram of *howsure* with normal density and kernel density plots added. Full sample.

of 1, 5, and 8 were also chosen rather frequently. In fact, just as in Martínez-Espiñeira and Lyssenko (2012b), it seems that 5 may have acted as a focal point for those who wanted to express an intermediate level of certainty, while 8 was probably the choice for those who were very sure but still wanted to leave some room for uncertainty. The non-normality of the distribution of values of howsure is further reflected by the very infrequent choice of values 2, 3, and 4. It is noteworthy that both Lyssenko and Martínez-Espiñeira (2012b) and Champ and Bishop (2001) also found about 13% of their respondents choosing the intermediate value 5 for the NCS. Our study also agrees with these previous works in finding that, in general, the most frequent choices for the numerical
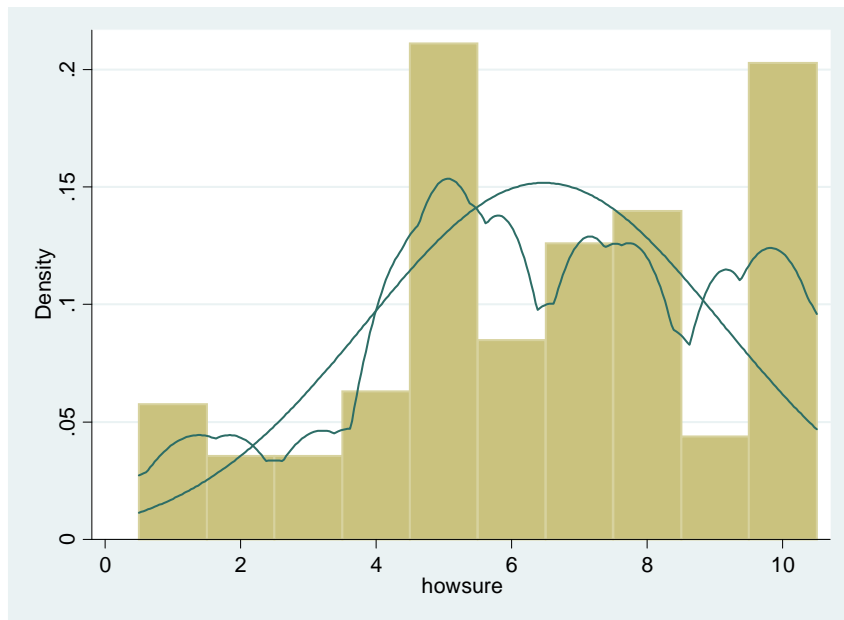
Figure C.2: Histogram of *howsure* with normal density and kernel density plots added. Only NY and YN responses.

certainty scale are 5, 7, 8, and 10.

When it came to developing a model to explain the values of variable *how-sure*, we noticed that there is very little guidance in the literature about what drivers could help explain the different values taken by this type of numerical certainty scale (NCS). However, Loomis and Ekstrand (1998) suggested that, theoretically, the level of certainty should be higher for both really high and really low bids, while the highest uncertainty would be associated with inter-mediate bids closer to the respondent's maximum willingness to pay.[4] We tried introducing the initial bid level (variable COST) in a quadratic form but ended up finding that the effect predicted by Loomis and Ekstrand (1998) was better captured by introducing binary indicators for whether the respondent switched from answering 'yes' to 'no', or vice versa, when responding to the initial and follow-up questions. This means that, as explained above, the respondents were most uncertain when their WTP was bracketed by the two offered bids. In the OLS and Tobit regressions (whose results are reported in Table C.2) both indicators ny and yn end up with significantly negative estimated coefficients. Other drivers that were found to affect *howsure* in a significant manner are the binary indicators of Version B, which has a positive effect, and the correspond-ing indicator for version C, which has a negative sign. The second version of the survey (Version B) presented the simplest payment scenario. Version B re-spondents only had to consider reductions in individual private risks of death,

---

[4] Other contributions find contradictory results about the relationship between the bid size and the value taken by the numerical certain scale (Champ and Bishop, 2001; Akter et al., 2009).

so it is not very surprising that they said to be relatively certain about their WTP , as compared with those respondents who had to consider both public and collective reductions in death risks and in injury risks as well. That is, the more elements included in the policy valued, the higher the higher the respondent uncertainty about WTP. Additionally, we found that smokers and those who stated to be healthier also reported a higher level of certainty.

On the other hand, we found no significant effect of variables such as *male* (which Samnaliev et al. 2006 found more certain than females about their responses, although Lyssenko and Martínez-Espiñeira 2012b found the opposite); *age* (which Sund 2009 and Svensson 2009 found positively correlated with response certainty) or *college*.

We also expected, given both our intuition and earlier results (Loomis and Ekstrand, 1998; Lyssenko and Martínez-Espiñeira, 2012b) to find that those respondents with previous experience of the risk would be more certain about their answers. However, although variable *drives30towork* is weakly significant and positive in the Tobit model, *hitmoose* is not significant, although also positive.

We find that further work must be done on the analysis of the numerical certainty scale given by variable howsure. It is likely very important to address the fact that many respondents cluster artificially around a few focal values.

351

Table C.2: OLS and Tobit models of numerical certainty scale (*howsure*).

	OLS	Tobit
lnav	0.1059	-0.0144
ny	-2.0199**	-3.1458**
yn	-1.5401**	-2.5153**
male	0.3055	0.3520
age	-0.0107	-0.0066
hitmoose	0.1637	0.3161
college	0.1884	0.1812
smoker	0.5759*	1.0376*
drives30towork	0.3850	0.8736$^+$
health	0.0130*	0.0178*
childrennumber	0.1787*	0.3523$^+$
Version B	-0.5942*	-1.0285**
Version C	0.5848$^+$	1.2324$^+$
diffM	-0.0216	-0.0377
diffI	-0.0026$^+$	-0.0042
constant	6.6963**	7.8103**
σ		3.8240**
N	1099	1099
ll	-2595.71	-2115.75
pseudo-R^2		0.04
R^2	0.1461	

$^+p < 0.10$, $^*p < 0.05$, $^{**}p < 0.01$

Appendix D

Modeling of mathscore

The *mathscore* variable was based on a series of four questions that were asked at the end of the survey. The questions assessed math skills such as numerical computation and understanding small fractions and decimals, both of which are skills used when assessing various mortality risk scenarios. The values taken by *mathscore* were simply the number of correctly answered questions by the respondent. For the complete list of questions, please refer to Appendix E.

We modelled by first merging the top two categories into one and then using an ordered logit regression to assess what covariates affect cognitive ability as measured here by *mathscore*. An ordered logit model is a model that is designed in assessing categorical dependent variables that have a meaningful order. Variable *mathscore* has a meaningful order here and it is assumed that more correctly answered questions is a good proxy measure of increasing cognitive ability in terms of numerical skills.

The model in Table D.1 shows high overall significance as the Wald test has a p-value of less than 0.000. The proportionality of odds test is a test whose null hypothesis is that one of the model assumptions (that the parameters have a consistent effect over all categorical values of *mathscore*) is valid. The result of the proportionality of odds test suggested that the effect of each individual variable might not be consistent across all categories, which in principle might mean that a model as parsimonious as an ordered logit model might not be an appropriate good choice. As shown in Table D.1, the significant covariates are *male*, its interaction with age (*maleage*), *education*, and the presence of children in the household and their number. The coefficient on *male* was significant at the 1% level and positive. This is in line with typical results in these types of studies that males, on average, have slightly better numerical skills. Another possible, if speculative, explanation is the fact that young females were more likely to finish schooling in Newfoundland, as males were often pulled out of school at a young age to go to work. Likely, the effect is due to some combination of these factors. Age has a negative effect on *mathscore* but when included together with *maleage* this effect is not statistically significant. A negative effect of age was expected, since these types of cognitive skills tend to decline with age. The effect of age for males, however, as assessed by the interaction term *maleage* was found negative and weakly statistically significant. This result suggests that older males tend to be worse at math than younger ones and worse at math than females and that the negative effect of age on cognitive skills is only significant for males.

The variable *education* has a positive effect and was significant at the 1%

level, which makes intuitive sense in that more education would mean better numerical skills. It could also be conversely true that individuals with better numerical skills tend to get more education. Finally, the presence of any children and teenagers in the household (*childrenany*) was significant at the 5% and positive. This somewhat surprising result does make sense if we consider the fact that parents who help their children with math homework are routinely exposed to dealing with fractions and decimals, which is exactly what these questions in the survey are assessing. On the other hand, the more children and teenagers in the household, the worse the cognitive skills of the respondent appear to be, as shown by the estimated coefficient on variable *childrennumber*.

Respondents from richer households tend to be significantly more capable of answering the mathscore questions correctly, which likely reveals a two directional effect. Richer individuals are more likely to have become rich by finding employment in sectors where mathematical skills are an asset, so there will be a component of the effect due to the endogeneity of this variable. Additionally, though, those who work in those types of jobs that require frequent use of mathematics and probability will have kept up their skills, so after controlling for education, we will find that richer respondents are more capable of having a higher *mathscore*.

Both *Q13* (self-assessed injury risk from a MVC) and health (self-reported health status) have a weakly significant effect.

When it comes to variable *verysure* (and indicator that tells us who stated that their degree of certainty about their responses to the payment questions,

variable *howsure*, was at least 7 out of 10) we decided to relax[1] the parallel regression restriction in the basic ordered logit model by allowing a different effect between the step going from 0 correct answers to 1, the step going from 1 to 2, and the last step going from 2 to either 3 or 4 (the two categories that we merged). Very sure respondents tend to answer an intermediate number of mathematical questions but are less likely to answer none or many questions than those with a lower value in their numerical certainty scale (*howsure*).

We consider this analysis merely exploratory at this stage and plan to conduct further work to try and identify the effect of additional variables. For example, we have noticed that variables *monthofbirth* and *smoker* would be partly responsible for the movement between categories 3 and 4 in the unaltered *mathscore*.

---

[1] We used Williams' *gologit2* routine for Stata (Williams, 2006; Williams, 2010; Pfarr et al., 2010).

Table D.1: Ordered Logit and Generalised Ordered Logit models of *mathscore*.

	(1) ologit	(2) gologit
age	-0.0069	-0.0071
male	1.8889**	1.8539**
maleage	-0.0242*	-0.0236+
education	0.2896**	0.2892**
childrennumber	-0.3673+	-0.3518+
childrenany	0.6375+	0.6349+
income	0.1478**	0.1431**
verysure	-0.3232*	-0.8407**
Q13	-0.0010*	-0.0010*
health	-0.0061+	-0.0060+
constant		1.4111*
cut1		
constant	-1.0226	
cut2		
constant	0.8858	
cut3		
constant	3.0622**	
Into Category 2		
verysure		0.0524
constant		-1.1193+
Into Categories 3 and 4		
verysure		-0.6191**
constant		-2.8372**
N	904	904
pseudo R^2	0.10	0.11
AIC		
Wald $\chi^2(10)$ test	142.21**	155.47**
Proportionality of Odds Test		

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Appendix E

Survey questionnaire text

Note: The following contains the raw format of the computer program used by the interviewers during the phone interviews. As such, it has been only minimally formatted for inclusion in this report and should not be read separately from Chapter 6. Note also that, since the surveys were delivered by phone, respondents had not actual access to any of this text.

AVALN:

SAMPLE BASED Avalon OR Outside OPEN FOR TESTING

- AVALON   1

- OUTSIDE   2

QI1:   Good Morning/Evening, my name is _ _ _ _ _ _ _ _ _ _ and I am calling on behalf of a team of Economists from/u Memorial University of Newfoundland, we are doing a brief study about the benefits of reducing moose-vehicle collisions. We will

need about 15 minutes of your time to ask you questions mainly about your direct or indirect experience with moose accidents, about your perceived exposure to the risk of a moose vehicle collision, and about your willingness to pay, hypothetically, to reduce that risk. Then your answers will be analysed anonymously, together with those of many other respondents in the province, in order to estimate how worthwhile would be to reduce the risk of collisions to different degrees. May I speak to the youngest adult in the household?

- Yes - CONTINUE   1

- NO - Is there would be a better time to call - SCHEDULE CALLBACK   2

- NO - Thank and discontinue - LOG AS HOUSEHOLD REFUSAL   3

  [DO NOT READ]:

- Don't Know/Refused - LOG AS HOUSEHOLD REFUSAL 9

QI2:   First I need to give you a few formal details about the survey: Your participation, which is greatly appreciated, is completely voluntary and totally anonymous and confidential. All data will be treated confidentially and stored securely for 5 years as required by Memorial University's ethics policy. All your personal information (we do not need names or anything, just your phone number, which was chosen randomly) will be kept confidential and will not be included in data files or other disclosed materials. We do not need to know who said what: we just work with the anonymous data, so after the call, we will not even need to keep your phone number again. A couple

of questions about moose accidents could potentially cause some stress to those who have personally experienced them. However, it is important to collect information from all types of affected people and we have tried to minimise the chances of causing any discomfort. In any case, you may terminate the interview at any point as well as skip any questions you do not feel comfortable with. In that case, the anonymity and confidentiality of your data would remain protected. The proposal for this research has been reviewed by the Interdisciplinary Committee on Ethics in Human Research and found to be in compliance with Memorial University's ethics policy. If you have ethical concerns about the research (such as the way you have been treated or your rights as a participant), you may contact the ICEHR at 864-2561.

QI3: You can contact the ICEHR by e-mail at icehr@mun.ca or by telephone at 864-2561

QI4: The results of this study will be available online at the website of the Economics Department at Memorial University. To learn more about this project or if you would like to contact the researchers with suggestions, questions, concerns, etc. feel free to email Dr. Lyssenko at nlysenko@mun.ca or phone 864-2149.

QS1: Are you 19 or older?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

QS2: Do you agree to go ahead with the interview?

- Yes   1

- No   2

  [DO NOT READ]:

- Don't Know/Refused - LOG AS HOUSEHOLD REFUSAL 9

Q1: [READ LIST] What best describes you?

- I was born in Newfoundland and have lived here for the last 6 months   1

- I moved to Newfoundland from another Canadian province and have lived here
  for the last 6 months   2

- I moved to Newfoundland from another country and have lived here for the last
  6 months   3

- I have lived in NL for less than 6 months   4

- [DO NOT READ]: Other   8

- Don't Know/Refused - LOG AS RESPONDENT REFUSAL   9

Q27. GENDER. [RECORD SEX OF RESPONDENT (ASK ABOUT THEIR
NAME IF UNSURE)].

- Male   1

- Female   2

Q22:   What is the postal code of your main residence?

- [DO NOT READ]:

- Don't Know/Refused - LOG AS HOUSEHOLD REFUSAL 999999

Q22_2: Do you live within or outside the Avalon area?

- Within - AVALON   1

- Outside Avalon - OUTSIDE   2

   [DO NOT READ]:

- Don't Know/Refused   9

KAVON:

RESPONSE BASED AVALON OR OUTSIDE AVALON

- AVALON   1

- OUTSIDE   2

- REFUSED   9

Q2_1: READ LIST IF NECESSARY

How long have you now lived in Newfoundland?

[INTERVIEWER NOTE: If the respondent replies with Don't know; probe with:

Just approximately...]

- I moved to Newfoundland less than 5 years ago   1

- I moved to Newfoundland between 5 and 10 years ago   2

- I moved to Newfoundland more than 10 years ago   3

- I have lived here all my life   4

- I used to live elsewhere and moved back here less than 5 years ago   5

- I used to live elsewhere and moved back here between 5 and 10 years ago   6

- I used to live elsewhere and moved back here more than 10 years ago   7

   [DO NOT READ]:

- Don't Know/Refused - PROBE BEFORE ACCEPTING   9

- Q3:   Do you regularly drive a vehicle?

- No   2

- Yes   1

  [DO NOT READ]:

- Don't Know/Refused   9

Q4:   READ LIST

What type do you drive most often?

- SUV-pickup truck   01

- Small-midsize car   02

- Full size car   03

- Minivan   04

[DO NOT READ]:

- Other - BUONLY/U/B - If vehicle make/year volunteered   77

  [DO NOT READ]:

- Don't Know/Refused   9

Other   88

Q5:   READ LIST IF NECESSARY

Do you know the year of the vehicle?

- 2009-2013   1

- 2005-2008   2

- 2000-2004   3

- 1995-1999   4

- Older than 1995   5

  [DO NOT READ]:

- Don't Know/Refused   9

Q6: About how many Km do you drive a year?

$R.0 0.1 300000

- [DO NOT READ]:

- Don't Know/Refused   999999

Q7: Do you drive more than 30 Km to work?

- YES   1

- NO   2

  [DO NOT READ]:

- Don't Know/Refused   9

Q7B: Does your job involve frequently driving between 12 midnight and 6 am?

- YES   1

- NO   2

  [DO NOT READ]:

- Don't Know/Refused   9

Q8: Have you seen a moose crossing the highway in NL in the last 3 years?

- YES   1

- NO   2

  [DO NOT READ]:

- Don't Know/Refused   9

RCQ9:   Recall for Q9

- your   1

- a   2

Q9: Have you ever hit a moose with <RCQ9> vehicle or had a near-miss (had to swerve/brake suddenly to avoid it)?

- YES   1

- NO   2

- [DO NOT READ]: Don't Know/Refused   9

RCQ10:

Recall for Q10

- else   1

-      2

Q10: Do you know of anyone <RCQ10> you personally know (family/friend) who ever hit a moose while driving?

- YES   1

- NO   2

- [DO NOT READ]: Don't Know/Refused   9

RR: Real Risk Holding Cell

- 60   1

- 80   2

- 100   3

- 120   4

RRX2:   Real Risk times 2

- 120   1

- 160   2

- 200   3

- 240   4

RM: 10-year mortality rate holding cell

- 4   04

- 6   06

- 8   08

- 10   10

- 12   12

RI: Risk of injury RI=30*RM

Q12D:

ENTER NUMBER FROM 0-100,000

In parts of Newfoundland drivers often hit moose. Which, apart from the death of the moose themselves, results in: car damages, injuries, and, even in some occasions, human deaths. The 10-year average traffic mortality risk in Newfoundland is <RR> in 100,000. So in 10 years in a city like St. John's (with its 200,000 people) one would expect about <RRX2> people to die in car accidents. Similarly, the 10-year average risk of dying from hitting a moose in Newfoundland is <RM> in 100,000. Of course, this risk varies from person to person depending on: where one lives, how much one drives, the type of vehicle one drives, whether one drives at night or not, the type of roads used, and how carefully one drives... Q12. Now, considering all this, how high do you think is your own risk of dying in a moose-vehicle collision in the next 10 years? That is, how many times in 100,000 you think your own risk is if the average in Newfoundland is <RM> in 100,000?

$R.0 0 100000

[DO NOT READ]:

- Don't Know/Refused   9999999

Q12ND:

ENTER NUMBER FROM 0-100,000

In parts of Newfoundland drivers often hit moose. Which, apart from the death of the moose themselves, results in: car damages, injuries, and, even in some occasions, human deaths. The 10-year average traffic mortality risk in Newfoundland is <RR> in 100,000. So in 10 years in a city like St. John's (with its 200,000 people) one would expect about <RRX2> people to die in car accidents. Similarly, the 10-year average risk of dying from hitting a moose in Newfoundland is <RM> in 100,000. Of course, this risk varies from person to person depending on: where one lives, how much one drives, the type of vehicle one drives, whether one drives at night or not, the type of roads used, and how carefully one drives... Q12. Now, considering all this, how high do you think is your own risk of dying in a moose-vehicle collision in the next 10 years? That is, how many times in 100,000 you think your own risk is if the average in Newfoundland is <RM> in 100,000? [IF NECESSARY: May add "as a passenger" as needed.]

$R.0 0 100000

[DO NOT READ]:

- Don't Know/Refused   9999999

Q12:

ENTER NUMBER FROM 0-100,000

In parts of Newfoundland drivers often hit moose. Which, apart from the death of the moose themselves, results in: car damages, injuries, and, even in some occasions, human deaths. The 10-year average traffic mortality risk in Newfoundland is <RR> in 100,000. So in 10 years in a city like St. John's (with its 200,000 people) one would expect about <RRX2> people to die in car accidents. Similarly, the 10-year average risk of dying from hitting a moose in Newfoundland is <RM> in 100,000. Of course, this risk varies from person to person depending on: where one lives, how much one drives, the type of vehicle one drives, whether one drives at night or not, the type of roads used, and how carefully one drives... Q12. Now, considering all this, how high do you think is your own risk of dying in a car accident involving a moose in the next 10 years? That is, how many times in 100,000 you think your own risk is if the average in Newfoundland is <RM> in 100,000?

$R.0 0 100000

[DO NOT READ]:

- Don't Know/Refused   9999999

Q13D:

ENTER NUMBER FROM 0-100,000

All these rates are what we call mortality rates but, luckily, most car accidents result in injuries but not deaths. In fact, the average 10-year injury rate from moose-vehicle accidents in NL is <RI> in 100,000. Q13. Again, given your individual circumstances and driving habits. How high do you think is your own risk of getting injured from hitting a moose in the next 10 years?

$R.0 0 100000

[DO NOT READ]:

- Don't Know/Refused   9999999

Q13ND:

ENTER NUMBER FROM 0-100,000

All these rates are what we call umortality/u rates but, luckily, most car accidents result in injuries but not deaths. In fact, the average 10-year injury rate from moose-vehicle accidents in NL is <RI> in 100,000. Q13. Again, given your individual circumstances and travel habits. How high do you think is your own risk of getting injured in a car accident involving a moose in the next 10 years?

$R.0 0 100000

[DO NOT READ]:

- Don't Know/Refused   9999999

Q13:

ENTER NUMBER FROM 0-100,000

All these rates are what we call "mortality/rates" but, luckily, most car accidents result in injuries but not deaths. In fact, the average 10-year injury rate from moose-vehicle accidents in NL is <RI> in 100,000. Q13. Again, given your individual circumstances and driving habits. How high do you think is your own risk of getting injured from hitting a moose in the next 10 years? or Q13. Again, given your individual circumstances and travel habits. How high do you think is your own risk of getting injured in a car accident involving a moose in the next 10 years?

$R.0 0 100000

[DO NOT READ]:

- Don't Know/Refused   9999999

SCTN:

Respondent Section

- VERSION A) PRIVATE GOOD   1

- VERSION B) PRIVATE GOOD   2

- VERSION C) PUBLIC GOOD   3

- VERSION D) PUBLIC GOOD   4

- VERSION E) B PRIVATE GOOD + D PUBLIC GOOD/D PUBLIC GOOD

  + B PRIVATE GOOD   5

HSLEV:

Holding cell which Version of E the respondent should get

VERSION E1)   1

VERSION E2)   2

COST:

The product cost

15.00   1

30.00   2

45.00   3

60.00   4

75.00   5

100.00   6

120.00   7

150.00   8

COST2:

The product cost TIMES 2

30.00   1

60.00   2

90.00   3

120.00   4

150.00   5

200.00   6

240.00   7

300.00   8

COSTH:

The product cost HALFED

7.50   1

15.00   2

22.50   3

30.00   4

37.50   5

50.00   6

60.00   7

75.00   8

MULTI:

The multiplier

2

3

4

ALOWM:

RETURN TO TWO DECIMAL POINTS Section A lower risk of mortal injury

$R.2

ALOWI:

Section A lower risk of injury

$R.2

QA14:   <SCTN> We have now some questions about your willingness to pay for increased traffic safety. Imagine that you are offered a new safety device that is not inconvenient, ugly, or complicated to use. In fact, you would not notice it. It reduces only your own mortality risk from its current level of <Q12> down to <ALOWM> in 100,000 and your injury risk from <Q13> to <ALOWI> in 100,000 should you be involved in a moose vehicle collision. It is only you as a driver who can personally benefit from it by reducing the risk of dying or being injured but only from hitting a moose: it does not help reduce the risks of other types of car accidents; it does not protect your passengers, other drivers, the moose, or the vehicle and, you could not lend it to anyone - even in your household. Assume that this personal device can be used in any of the vehicles you drive. Assume that its effect lasts only one year, so after that, you must make another payment if you want to continue the risk reduction. B Remember that that there would be other ways to improve your safety and that paying for this device would mean having less money for other personal expenditures such as rent, food, gas, and so on./b [INTERVIEWER NOTE: If asked, stress that the level of risk would revert to <Q12> and <Q13> in 100,000 after discontinuing the use of the device.] QA14. Would you be willing to pay $<COST> per year for this device?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

QA15X:

QA15. Would you be willing to pay \$<COST2> per year?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

QA16X:

QA16. Would you be willing to pay \$<COSTH> per year?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

BLOWM:

RETURN TO TWO DECIMAL POINTS Section B lower risk of mortal injury

\$R.2

QB14:   We now have some questions about your willingness to pay for increased

traffic safety. Imagine that you are offered a new safety device that is not inconvenient,

ugly, or complicated to use. In fact, you would not notice it. It reduces only your

own mortality risk from its current level of <Q12> down to <BLOWM> in 100,000

should you be involved in a moose vehicle collision. It is only you as a driver who can

personally benefit from it by reducing the risk of dying but only from hitting a moose:

it does not help reduce the risk of you dying in other types of car accidents or your risk

of getting injured; it does not protect your passengers, or other drivers, or the moose,

or the vehicle and you could not lend it to anyone, even in your household. Assume

that this personal device can be used in any of the vehicles you drive. Assume that

its effect lasts only one year, so after that, you must make another payment if you

want to continue the risk reduction. B Remember that there would be other ways to

improve your safety and that paying for this device would mean having less money for

other personal expenditures such as rent, food, gas, and so on./b [INTERVIEWER

NOTE: If asked, stress that the level of risk would revert to <Q13> in 100,000 after

discontinuing the use of the device.] QB14. Would you be willing to pay $<COST>

per year for this device?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

QB15X:

QB15. Would you be willing to pay \$<COST2> per year?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

QB16X:

QB16. Would you be willing to pay \$<COSTH> per year?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

HLDFP:

Holding Cell Fed/Prov

- Federal   1

- Provincial   2

RCLC6:

Section C Holding Cell 50% of respondents see an extra line

so if a majority of Newfoundlanders supported the program, it would go ahead; otherwise there would be no program and no increased <RCLC4>.

1

2

CLOWM:

RETURN TO TWO DECIMAL POINTS Section C lower risk of mortal injury

$R.2

CLOWI:

Section C lower risk of injury

$R.2

QC14:   We now would like to know if you would support a program aimed at reducing the risk of moose-vehicle collisions in Newfoundland by installing and maintaining fences along the highways, together with under and over passes for the moose to cross the road safely. There are many good reasons why one is willing to pay or not. Before answering the question – we would like to remind you that there may be other causes to support, including programs aimed at promoting health and safety in other ways, and we would also like to remind you that supporting the program would mean having less money for other personal expenditures such as rent, food, gas, and so on. Imagine a [randomized federal/provincial] fencing program that would run for five years. If carried out, the program would reduce the mortality risk from a MVC (Moose-Vehicle Collision) for the general population of the province from <RM> to <CLOWM> in 100,000. Note that the risk of injury to you, your passengers or other drivers would be now also proportionally less, as well as the effects of injuries to the moose. In particular, the 10-year injury risk from moose-vehicle accidents would go down from <RI> to <CLOWI> in 100,000. The funds for this would come from an increase in [annual federal taxes/provincial driver's licence fees]. The extra [taxes/fees] your household would have to pay would be $<COST> per year for the five years of the program if it went ahead. If in trying to decide about the program the [federal/provincial] government had a referendum... <RCLC6> QC14. Would you vote yes in this referendum?

- Yes   1

- No   2   [DO NOT READ]: Don't Know/Refused   9

QC15X:

QC15. Would vote YES if the extra <RCLC5> for you were $<COST2> per year for the five years?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

QC16X:

QC16. Would vote YES if the extra <RCLC5> for you were $<COSTH> per year for the five years?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

HDDFP:

Holding Cell Fed/Prov becomes biv flag for this section

- Federal   1

- Provincial   2

384

RCLD4:

Section D Holding Cell 50% of respondents see an extra line becomes biv flag for

respondents seen

so if a majority of Newfoundlanders supported the program, it would go ahead; otherwise there would be no program and no increased <RCLD6>.   1

    2

DLOWM:

Section D lower risk of mortal injury

$R.2

QD14:   We now would like to know if you would support a public policy aimed at reducing the general mortality risk of car drivers by reducing the risk of moose-vehicle collisions in Newfoundland. There are many good reasons why one is willing to pay or not. Before answering the question, we would like to remind you that there may be other causes to support, including programs aimed at promoting health and safety in other ways, and we would also like to remind you that supporting the program would mean having less money for other expenditures such as rent, food, gas, and so on. Imagine a [federal/provincial] program that would run for five years. If carried out, the program would reduce the mortality risk from a MVC (Moose-Vehicle Collision) for the general population of the province from <RM> to <DLOWM> in 100,000. Although this is a bit unrealistic, please imagine that the risk of injury to you or other drivers, or to any passengers, would remain the same and that the risk of injuries to the moose would also remain unchanged with the program. In other words, the policy would only reduce your mortality risk, nothing else. The funds for this would come from an increase in [annual federal taxes/provincial driver's licence fees]. The extra [taxes/fees] your household would have to pay would be $<COST> per year for the five years of the program if it went ahead. If in trying to decide about the program the [federal/provincial] government had a referendum... <RCLD4> QD14. Would you vote yes in this referendum?

- Yes   1

- No   2   [DO NOT READ]: Don't Know/Refused   9

QD15X:

QD15. Would vote YES if the extra <RCLD3> for you were $<COST2> per year for the five years?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

QD16X:

QD16. Would vote YES if the extra <RCLD3> for you were $<COSTH> per year for the five years?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

E1BLM:

Section E1_B lower risk of mortal injury

$R.2

E1B14:   We now have some questions about your willingness to pay for increased traffic safety. Imagine that you are offered a new safety device that is not inconvenient, ugly, or complicated to use. In fact, you would not notice it. It reduces only your own mortality risk from its current level of <Q12> down to <E1BLM> in 100,000 should you be involved in a moose vehicle collision. It is only you as a driver who can personally benefit from it by reducing the risk of dying but only from hitting a moose: it does not help reduce the risk of you dying in other types of car accidents or your risk of getting injured; it does not protect your passengers, or other drivers, or the moose, or the vehicle and you could not lend it to anyone, even in your household. Assume that this personal device can be used in any of the vehicles you drive. Assume that its effect lasts only one year, so after that, you must make another payment if you want to continue the risk reduction. Remember that there would be other ways to improve your safety and that paying for this device would mean having less money for other personal expenditures such as rent, food, gas, and so on. [INTERVIEWER NOTE: If asked, stress that the level of risk would revert to <Q13> in 100,000 after discontinuing the use of the device.] E1B14. Would you be willing to pay $<COST> per year for this device?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

E1B5X:

E1B15. Would you be willing to pay \$<COST2> per year?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

E1B6X:

E1B16. Would you be willing to pay \$<COSTH> per year?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

HE1FP:

Holding Cell Fed/Prov becomes biv flag for section f/p

- Federal   1

- Provincial   2

RCE14:

Section E1_D Holding Cell 50% of respondents see an extra line becomes biv flag

for seen line

so if a majority of Newfoundlanders supported the program, it would go ahead;

otherwise there would be no program and no increased <RCE16>.

- 1

- 2

E1DLM:

Section E1_D lower risk of mortal injury

$R.2

E1D14:   We now would like to know if you would support a public policy aimed at reducing the general mortality risk of car drivers by reducing the risk of moose-vehicle collisions in Newfoundland. There are many good reasons why one is willing to pay or not. Before answering the question, we would like to remind you that there may be other causes to support, including programs aimed at promoting health and safety in other ways, and we would also like to remind you that supporting the program would mean having less money for other expenditures such as rent, food, gas, and so on. Imagine a [federal/provincial] program that would run for five years. If carried out, the program would reduce the mortality risk from a MVC (Moose-Vehicle Collision) for the general population of the province from <RM> to <E1DLM> in 100,000. Although this is a bit unrealistic, please imagine that the risk of injury to you or other drivers, or to any passengers, would remain the same and that the risk of injuries to the moose would also remain unchanged with the program. In other words, the policy would only reduce your mortality risk, nothing else. The funds for this would come from an increase in [annual federal taxes/provincial driver's licence fees]. The extra [taxes/fees] your household would have to pay would be $<COST> per year for the five years of the program if it went ahead. If in trying to decide about the program the [federal/provincial] government had a referendum... <RCE14> E1D14. Would you vote yes in this referendum?

- Yes   1

- No   2

- [DO NOT READ]: Don't Know/Refused   9

391

E1D5X:

E1D15. Would vote YES if the extra <RCE13> for you were $<COST2> per year for the five years?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

E1D6X:

E1D16. Would vote YES if the extra <RCE13> for you were $<COSTH> per year for the five years?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

E2DLM:

Section E2_D lower risk of mortal injury

$R.2

HE2FP:

Holding Cell Fed/Prov becomes biv flag for section f/p

- Federal   1

- Provincial   2

RCE24:

Section E2_D Holding Cell 50% of respondents see an extra line becomes biv flag

for seen line

so if a majority of Newfoundlanders supported the program, it would go ahead;

otherwise there would be no program and no increased <RCE26>.

- 1

- 2

E2D14:   We now would like to know if you would support a public policy aimed at reducing the general mortality risk of car drivers by reducing the risk of moose-vehicle collisions in Newfoundland. There are many good reasons why one is willing to pay or not. Before answering the question, we would like to remind you that there may be other causes to support, including programs aimed at promoting health and safety in other ways, and we would also like to remind you that supporting the program would mean having less money for other expenditures such as rent, food, gas, and so on. Imagine a [federal/provincial] program that would run for five years. If carried out, the program would reduce the mortality risk from a MVC (Moose-Vehicle Collision) for the general population of the province from <RM> to <E2DLM> in 100,000. Although this is a bit unrealistic, please imagine that the risk of injury to you or other drivers, or to any passengers, would remain the same and that the risk of injuries to the moose would also remain unchanged with the program. In other words, the policy would only reduce your mortality risk, nothing else. The funds for this would come from an increase in [annual federal taxes/provincial driver's licence fees]. The extra [taxes/fees] your household would have to pay would be $<COST> per year for the five years of the program if it went ahead. If in trying to decide about the program the [federal/provincial] government had a referendum?

- Yes   1

- No   2

- [DO NOT READ]: Don't Know/Refused   9

E2D5X:

E2D15. Would vote YES if the extra <RCE23> for you were \$<COST2> per year for the five years?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

E2D6X:

E2D16. Would vote YES if the extra <RCE23> for you were \$<COSTH> per year for the five years?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

E2BLM:

Section E2_B lower risk of mortal injury

\$R.2

E2B14:   We now have some questions about your willingness to pay for increased traffic safety. Imagine that you are offered a new safety device that is not inconvenient, ugly, or complicated to use. In fact, you would not notice it. It reduces only your own mortality risk from its current level of <Q12> down to <E2BLM> in 100,000 should you be involved in a moose vehicle collision. It is only you as a driver who can personally benefit from it by reducing the risk of dying but only from hitting a moose: it does not help reduce the risk of you dying in other types of car accidents or your risk of getting injured; it does not protect your passengers, or other drivers, or the moose, or the vehicle and you could not lend it to anyone, even in your household. Assume that this personal device can be used in any of the vehicles you drive. Assume that its effect lasts only one year, so after that, you must make another payment if you want to continue the risk reduction. Remember that there would be other ways to improve your safety and that paying for this device would mean having less money for other personal expenditures such as rent, food, gas, and so on. [INTERVIEWER NOTE: If asked, stress that the level of risk would revert to <Q13> in 100,000 after discontinuing the use of the device.] E1B14. Would you be willing to pay $<COST> per year for this device?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

E2B5X:

E2B15. Would you be willing to pay \$<COST2> per year?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

E2B6X:

E2B16. Would you be willing to pay \$<COSTH> per year?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

Q17:   On a scale of 1 to 10, where 1 is "not very sure" and 10 is "very sure",

how would you rank your last answers about whether you were willing to pay or not?

- 1 - Not very sure   01

- 2   02

- 3   03

- 4   04

- 5   05

- 6   06

- 7   07

- 8   08

- 9   09

- 10 - Very sure   10

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   99

Q18:   Why would you not be willing to pay any of the previous suggested
amounts? [DO NOT READ]

- I don't believe the money would be spent on that   01

- I would not trust the government to do the job properly   02

- Too expensive/I cannot afford that   03

- I already pay too much tax   04

- It should not be financed through taxes/not everyone should have to pay their
  share to protect drivers   05

- I already contribute to other environmental programs/causes   06

- I should not have to pay individually: the province/government should pay for
  that without raising taxes   07

- I do not care about MVCs   08

- I do not drive   09

- Drivers should just slow down   10

- I do not believe that the program would be effective   11

- The drivers should pay for that themselves   12

- The drivers' insurance should pay for that   13

- The government should fund the program with existing revenues, and not ask
  for additional taxes   14

- The government has other higher priorities for spending taxpayers' money   15

- Other reason   77

[DO NOT READ]:

- Don't Know/Refused   9

- Other   88

- Not necessary/MVC risk is low/waste of money (unspecified)   20

- I am a careful driver/am not worried about hitting moose   21

- Brush should be trimmed from roadsides to enable visibility   22

- Device protects only drivers/not the car/other passengers   23

- I don't/rarely drive at night/these accidents occur at night   24

- I don't/rarely drive on the highway   25

- Moose population should be decreased/culled   26

- MVC prevention should focus on driver safety/awareness   27

- Need more information/proof/evidence of effectiveness   28

- There are too few/no moose in my area to worry about it   29

- I do not drive enough in moose-populated areas   30

- The problem exists because moose are not native to the area   31

- Should not have to pay for the poor habits of other drivers   32

Q19: Have you hunted for moose in the last 5 years?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

Q19B: Have you eaten moose at home in the past 12 months?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

Q20: How many people in your household are under 18, including babies and small children? [IF NECESSARY: Clarify that dependants who do not live at the respondent's home full time also count. Example; children who just come for the weekend but whose custody is shared by the respondent.]

$E 1 20

- None (Zero)   00

  [DO NOT READ]:

- Don't Know/Refused   99

401

Q21: [READ LIST IF NECESSARY] What is your highest level of education?

- Some grade school/high school   01

- High school graduate   02

- Some tech, vocational, or trade school   03

- Graduate of a tech, vocational, or trade school   04

- Professional undergraduate degree (nurse, teacher, community college)   05

- Some university   06

- University graduate   07

- Masters, doctorate, or professional degree   08

- No response   09

[DO NOT READ]:

- Don't Know/Refused   99

Q23:   ENTER EXACT AGE

How old are you?

limits: 19 119

- 18-29   201

- 30-49   202

- 50-64   203

- 65 and over   204

[DO NOT READ]:

- Don't Know/Refused   999

Q23_2: [READ LIST]

Which of the following age ranges would you fall into?

- 18-29   1

- 30-49   2

- 50-64   3

- 65 and over   4

[DO NOT READ]:

- Don't Know/Refused   9

Q24: On which month were you born?

January   01

February   02

March   03

April   04

May   05

June   06

July   07

August   08

September   09

October   10

November   11

December   12

[DO NOT READ]:

- Don't Know/Refused   99

Q25:   In order to group households into general categories, we would like to ask you which of the following intervals best describes your household income ufor 2012/u before taxes. We would like to remind you that these data will be only treated in aggregate statistics so your anonymity will be preserved. Remember that a machine randomly chose and dialed your number and that after this call, the record of the phone number will be kept separate from the rest of the information and not available to those analyzing the rest of the information. Q25. About how much is your household income before taxes?

- Less than 30,000   1

- Between 30,000 and 50,000   2

- Between 50,000 and 70,000   3

- Between 70,000 and 90,000   4

- Between 90,000 and 110,000   5

- Between 110,000 and 130,000   6

- Between 130,000 and 150,000   7

- Over 150,000   8

[DO NOT READ]:

- Don't Know/Refused   9

Q26: How healthy would you say you are in general? Consider a 1 to 100 scale where 1 means very sick and 100 is perfectly healthy. How much do you think you would currently score?

$E 1 100

[DO NOT READ]:

- Don't Know/Refused   999

Q26_2:   Have you smoked more than 10 cigarettes in the past 10 days?

- Yes   1

- No   2

[DO NOT READ]:

- Don't Know/Refused   9

Q28:   [READ LIST]

What do you think is more likely to happen:

- Something that happens 3 times in 10,000 or   1

- Something that happens 6 times in 100,000?   2

DUMMY   3

[DO NOT READ]:

- Don't Know/Refused   9

Q29:

[READ LIST]

Q29. Which number in the following group of numbers represents the smallest amount?

- 3/4   1

- 0.8   2

- 31   3

- 0.33   4

[DO NOT READ]:

- Don't Know/No Answer   9

Q30:   [READ LIST]

A 20% reduction in a 40% risk level results in a new risk level of:

- 32%   1

- 20%   2

- 35%   3

- 80%   4

[DO NOT READ]:

- Don't Know/Refused   9

Q31:   [DO NOT READ LIST]

A baseball bat and a ball together cost $11. The baseball bat costs $10 more than the ball. How much is the ball?

- $1   1

- $0.5   2

- $0   3

- Any other number   4

[DO NOT READ]:

- Don't Know/Refused   9