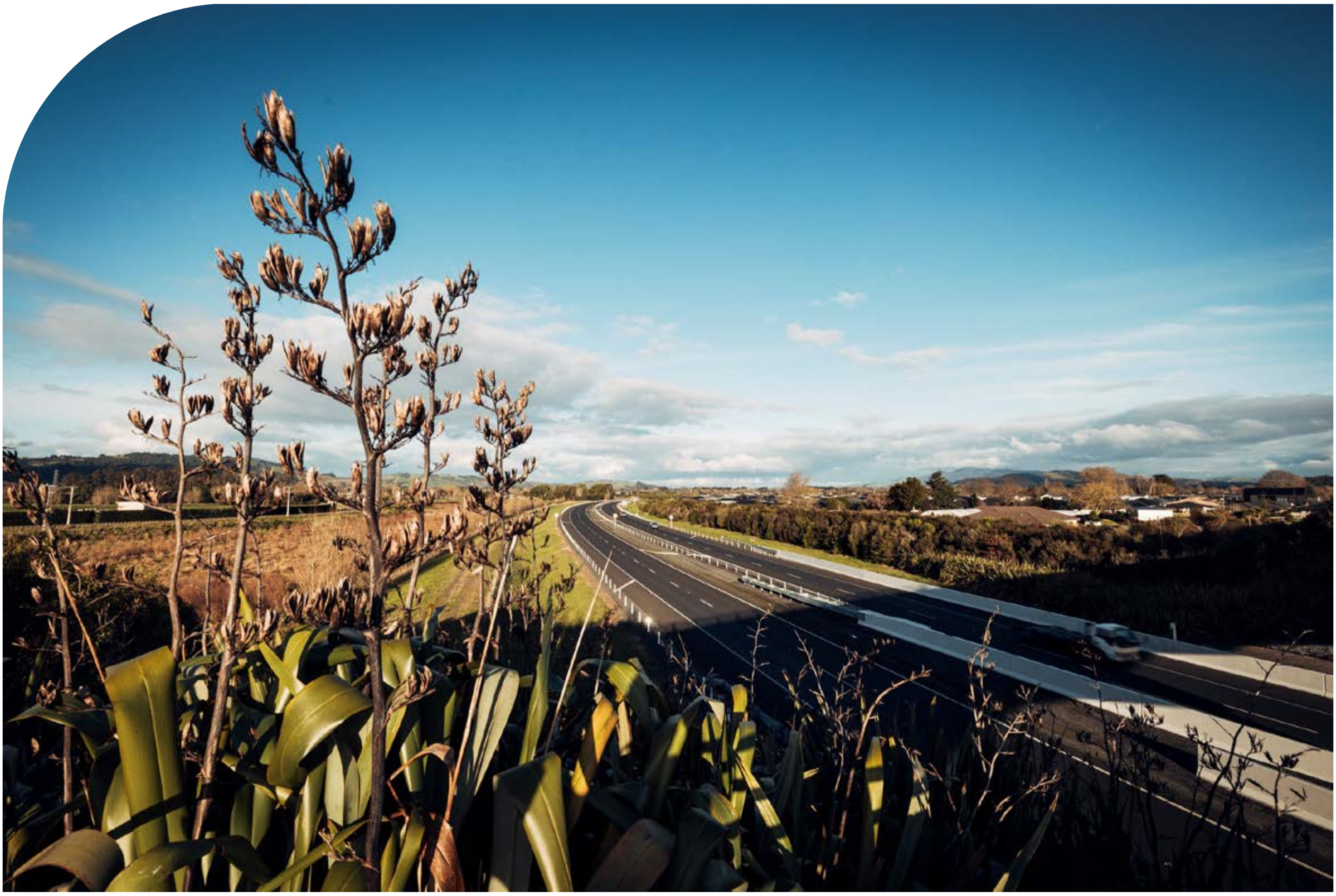


Monetised benefits and costs manual (MBCM) parameter values

Results of a survey to derive values for road safety, travel time and reliability

February 2023

T Denne, Resource Economics, Auckland
G Kerr, Lincoln University, Lincoln
A Stroombergen, Infometrics, Wellington
D Glover, Gravitas Research & Strategy, Auckland
M Winder, Gravitas Research & Strategy, Auckland
B Gribben, CBG Health Research Ltd, Auckland
N Tee, CBG Health Research Ltd, Auckland

Waka Kotahi NZ Transport Agency research report 698
Contracted research organisation – Resource Economics Ltd

WAKA KOTAHI NZ TRANSPORT AGENCY

New Zealand Government

ISBN 978-1-99-004482-3 (electronic)
ISSN 2815-8377 (electronic)

Waka Kotahi NZ Transport Agency
Private Bag 6995, Wellington 6141, New Zealand
Telephone 64 4 894 5400; facsimile 64 4 894 6100
NZTAresearch@nzta.govt.nz
www.nzta.govt.nz

Denne, T., Kerr, G., Stroombergen, A., Glover, D., Winder, M., Gribben, B., & Tee, N. (2023). *Monetised benefits and costs manual (MBCM) parameter values: Results of a survey to derive values for road safety, travel time and reliability* (Waka Kotahi NZ Transport Agency research report TAR 18-04).

Resource Economics was contracted by Waka Kotahi NZ Transport Agency in 2019 to carry out this research.

Keywords: crashes, injuries, fatalities, public transport, reliability, statistical life, time

An important note for the reader

Waka Kotahi NZ Transport Agency is a Crown entity established under the Land Transport Management Act 2003. The objective of Waka Kotahi is to undertake its functions in a way that contributes to an efficient, effective and safe land transport system in the public interest. Each year, Waka Kotahi funds innovative and relevant research that contributes to this objective.

The views expressed in research reports are the outcomes of the independent research and should not be regarded as being the opinion or responsibility of Waka Kotahi. The material contained in the reports should not be construed in any way as policy adopted by Waka Kotahi or indeed any agency of the New Zealand Government. The reports may, however, be used by New Zealand Government agencies as a reference in the development of policy.

While research reports are believed to be correct at the time of their preparation, Waka Kotahi and agents involved in their preparation and publication do not accept any liability for use of the research. People using the research, whether directly or indirectly, should apply and rely on their own skill and judgement. They should not rely on the contents of the research reports in isolation from other sources of advice and information. If necessary, they should seek appropriate legal or other expert advice.

Acknowledgements

We are grateful to the many people who have assisted with this project, particularly our steering group: Ernie Albuquerque, Joanne Leung, Jennifer McSaveney, Sandy Fong and Ian Melsom. We would also like to thank everyone that participated in the survey component of the study. Their generosity in taking the time and effort to answer the questionnaire has been greatly appreciated. Thanks also to the interviewers who worked so diligently to collect the data.

The work has been improved significantly from the valuable inputs from two peer reviewers – Professor Michiel Bliemer (University of Sydney) and Professor Dan Graham (Imperial College, London) – in addition to those from the Institute for Transport Studies, University of Leeds.

Abbreviations and acronyms

AIC	Akaike information criterion
ANA	attribute non-attendance
ASC	alternative-specific constant
ATAP	Australian Transport Assessment and Planning
BIC	Bayesian information criterion
bvkt	billion vehicle kilometres travelled
CAPI	computer-assisted personal interviewing
CASI	computer-assisted self-interviewing
CBA	cost–benefit analysis
CE	choice experiment
CV	contingent valuation
EV	electric vehicle
GMXL	generalised mixed logit
IVT	in-vehicle time
LL	log likelihood
MAIS	Maximum Abbreviated Injury Scale
MAV	multi-attribute valuation
MBCM	*Monetised Benefits and Costs Manual*
MNL	multinomial logit
MRS	marginal rate of substitution
NZTCS	New Zealand Travel Choices Survey
OECD	Organisation for Economic Co-operation and Development
RP	revealed preference
RPL	random parameters logit (model)
RR	reliability ratio
SC	stated choice
SD	standard deviation

SE	standard error
SP	stated preference
SQ	status quo
VFRR	value of fatal risk reduction
VoR	value of reliability
VoSL	value of a statistical life
VPF	value of preventing a fatality
VRMR	value of reduced mortality risk
VRR	value of risk reduction
VTT	value of travel time
VTTS	value of travel time savings
WTP	willingness to pay

Contents

Executive summary

Study objectives

The purpose of this research project is to provide data and analysis to support robust monetary values that could be used to measure the costs or benefits of changes in levels of road safety, travel time, reliability of travel time and other factors for which market prices are not available. The objective is to provide values for inclusion in an updated *Monetised Benefits and Costs Manual* for use in social cost–benefit and other analyses.

The methodology to be used for the research is a national stated preference survey with subsequent analysis of data collected.

Use of a choice experiment

Stated preference analysis is used to identify values when there is no market for an attribute and no suitable related market. It was used to develop the current New Zealand values for risk reduction (fatalities and injuries), time, reliability and public transport quality and has been suggested for use in collecting new attribute data. There are two main stated preference approaches: contingent valuation and choice experiment. Both use surveys to obtain responses, but choice experiment is regarded as introducing fewer biases than contingent valuation and providing survey respondents with more realistic decision options. It is now widely used for transport value analyses because it enables several attribute values to be collected simultaneously, it enables direct analysis of trade-offs between these attributes, and it delivers improved statistical efficiency (lower error margins for a given survey size). Choice experiment is widely regarded as current best practice for deriving non-market values for transport analysis and has been used recently to derive values for transport agencies for risk reduction (in Australia) and for time and reliability (in the UK).

Survey development

The most widely used recent approach to deriving attribute values for road transport has been to use a choice experiment in which a survey gives respondents route choice options for a trip. The options are defined using several attributes with different attribute values and may include travel costs, average travel time, reliability or distribution of expected travel time, congestion levels and safety (risk of death or injury). For public transport, attributes may include quality attributes also, including crowding and service frequency.

An example of a route choice question is shown in Figure ES.1. It uses the style and format used in our survey and combines a range of attributes. It asks respondents to choose their favoured route for a hypothetical trip to a single destination. The attributes include those that will directly affect the respondent (eg, the travel time, proportion of the trip in heavy traffic and cost) and safety attributes that require the respondent to interpret the risks to them from information on the number of deaths per annum on each route. This uses a presentation of risks that builds on approaches developed and used elsewhere.

Figure ES.1 Example route choice question

Which route would you prefer to take?
Please select one only

	Route one	Route two
Average travel time	3 hours 15 minutes	3 hours 0 minutes
Lateness	10% of trips delayed by 40 minutes	10% of trips delayed by 45 minutes
Heavy traffic	20%	10%
Trip cost	$15	$20
Fatalities (per year)	2 deaths	1 death
Serious injuries (per year)	10	2
Minor injuries (per year)	20	5

The route choice approach was tested in a pilot study, particularly to investigate how many attributes might be included in a single survey. In this new study we further tested the route choice approach, in addition to variants of the presentation and format of the questions. Three initial rounds of testing with 60 face-to-face interviews were conducted in 2020, followed by a pilot survey with 100 participants to further check the response to the survey. The final survey was launched in 2021, with over 7,000 participants throughout New Zealand, and was completed in October 2021.

This testing phase involved detailed questioning of respondents during and after they had completed a survey to understand their reactions to and interpretations of the questions. This identified difficulties in interpretation that led to some significant changes and to the adoption of a separate community or citizen value-based approach to questions relating to safety. Rather than asking survey respondents to choose between journey route options that differed in levels of risk that might affect them individually, it asked respondents to choose between national investment programme options that differed in the total resulting number of deaths and injuries for New Zealand as a whole. An example of this approach is shown in Figure ES.2.

Figure ES.2 Investment programme choice task – lower deaths and injuries

QS1 Which of the following would you prefer?

	Current (no changes made)	Investment option one	Investment option two
Deaths (per year)	250	200	250
Serious injuries (per year)	1750	1000	1500
Minor injuries (per year)	6000	5000	6000
Increase in your personal costs (per year)	$0	$200 more per year	$100 more per year
	Prefer this (no change)	Prefer option 1	Prefer option 2

The reason for changing to an investment programme format was partly because the feedback to testing suggested respondents were not able to interpret the risk to them or they discounted risks to them because they believed that they were at lower risk than others based on their driving behaviour or vehicle. In addition, we are uncertain on theoretical grounds that an approach as shown in Figure ES.1 can ever be used to provide an accurate quantification of how people are trading off money and risk.

The final survey used a route choice approach to provide attribute values for:

- value of travel time using private and public transport[1]
- value of reliability in travel time
- road congestion
- crowding on public transport.

An investment choice approach was used to provide attribute values for risk reduction for fatalities, serious injuries and minor injuries.

The final survey is included in Appendix A.

Final survey

The final survey was conducted between February and October 2021 with face-to-face interviews throughout New Zealand (Figure ES.3). The survey roll-out was affected significantly by COVID-19, which delayed the start and, because of lockdowns, meant surveying in parts of the country (Auckland and Northland) had to finish earlier than in others.

Figure ES.3 Number of interviews by date and location

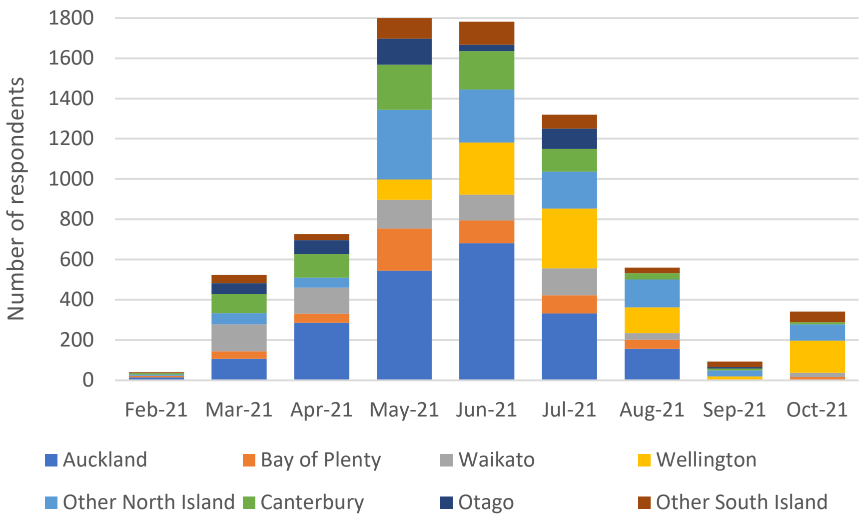


Table ES.1 shows the composition of the achieved sample by ethnicity (total response). The number of responses is greater than the number of respondents, as the selection of multiple ethnic groups was possible.

[1] In this report we use a 'value of travel time', with reductions in travel time treating this value as a benefit. This is consistent with the approach used in the MBCM. We note that other studies refer to a 'value of travel time savings'.

Table ES.1 Sample composition

	Male	Female	Gender diverse	Not reported	Total	%	2018 Census
European	2,356	2,763	5	0	5,124	71.3	70.2
Māori	345	486	1	0	832	11.6	16.5
Pacific peoples	160	226	0	0	386	5.4	8.1
Asian	385	411	1	0	797	11.1	15.1
Other (not specified)	237	264	0	0	501	7.0	1.5
Not reported	17	29	0	131	177	2.5	0
Total	**3,500**	**4,179**	**7**	**131**	**7,817**	**108.8**	**111.4**

The number of Māori and Pasifika respondents included in the survey was lower than expected, due to the COVID-related disruptions to fieldwork. Representativeness was obtained *ex post* by CBG's assignment of weights to individuals. The achieved sample was post-stratified to the New Zealand population using Census 2018 data, by age (16–24, 25–64, 65+), gender, and ethnicity (Māori, Pasifika, other). Weighted responses were used in all statistical analyses.

Results

The survey results are summarised below using mean values, with the greater detail on ranges in the main report (chapter 7). Separate values are provided for time and reliability for public and private transport. Safety attribute values are independent of travel mode; they are 'willingness to pay' values applicable to all people, whether transport users or not. We provide minimum and maximum values below, rather than simply as means.

In Table ES.2, the first row shows that when commuting by public transport for any length of trip, respondents would value:

- a change in certain travel time (where they would be certain of how much the trip would be faster or slower) at $8.16 per hour if they were sitting or $11.88 per hour if standing
- a change in the time between scheduled public transport departures (this is called headway) at $8.28 per hour
- a change in the variability or reliability of travel time, measured as a change in the standard deviation of travel time, at $14.64 per hour.

Table ES.2 Public transport – mean value of certain travel time and reliability ($/hour)

Trip purpose	Trip length	Sitting	Standing	Headway	Reliability (standard deviation of travel time)
Commuting	All	8.16	11.88	8.28	14.64
	Short	10.02	12.18	7.86	15.18
	Long	5.64	12.96	9.96	16.32
Other	All	6.61	10.33	6.13	16.89
	Short	6.30	10.05	5.84	13.36
	Long	6.99	12.98	5.41	14.74

Table ES.3 gives us averaged results for questions about private transport. For example, the first row shows that when commuting, respondents would value:

- a change in certain travel times at $30.90 per hour if the traffic was free-flowing and at $57.24 per hour if they were in heavy traffic, a difference of $26.34 per hour, which is the congestion increment
- a change in the reliability of travel time, measured as a change in the standard deviation, at $26.52 per hour.

Table ES.3 Private transport – mean value of travel time and reliability ($/hour)

Trip purpose	Trip length	Mean	Free-flowing	Heavy traffic	Congestion increment	Reliability (standard deviation of travel time)
Commuting	All	38.40	30.90	57.24	26.34	26.52
	Short	37.56	38.70	73.74	35.04	24.72
	Medium	48.60	37.38	59.64	22.26	31.62
	Long	42.54	30.60	55.80	25.20	20.58
Other	All	37.13	31.97	57.07	25.10	24.96
	Short	37.03	36.32	69.78	33.46	22.05
	Medium	33.12	42.90	61.61	18.71	18.62
	Long	35.31	34.04	60.48	26.44	6.21

Safety attribute values are independent of how people travel. They are 'willingness to pay' values that apply to everyone, whether transport users or not. These are shown in Table ES.4 as minimum and maximum values of what respondents would be willing for the government to pay (for example, as an increase in tax) for a reduction in the aggregate number of each of these types of incidents. For example, on average, respondents are willing to pay $4.30 per annum for one annual road death fewer and $0.225 for one serious injury fewer. To estimate a total value for a reduction in these events, the individual values are multiplied either by the national adult population (maximum) or by the number of households (assuming respondents are stating their willingness to pay out of a household budget). This results in a value of a reduced fatality of $8.1 million to $16.9 million.

Table ES.4 Mean values of preventing fatalities and injuries

	Willingness to pay per event ($/respondent)	Minimum aggregate national value	Maximum aggregate national value
Death	$4.3	$8.1 million	$16.9 million
Serious injury	$0.225	$429,458	$890,681
Minor injury	$0.023	$44,218	$91,707

Abstract

The *Monetised Benefits and Costs Manual* (MBCM) includes non-market values to be used in cost–benefit analysis of transport projects. This study conducted a national stated preference survey using face-to-face surveys (n = 7,203) and a choice modelling approach to derive new values to update those in the current manual. Survey questions that asked respondents to choose their preferred road route and/or public transport service option were used to derive values for time for different trip lengths and purposes, and whether sitting or standing on public transport. They also were used to derive values for reliability of travel time and time in congestion. A different set of questions was used to derive values for reductions in fatalities and injuries. The choice questions provided respondents with options for government programmes that differed in cost to them and in total annual numbers of road deaths and injuries. The results suggest the benefit of significant changes to some of the values in the current MBCM, including for the base value of travel time and for the value of preventing a fatality.

1 Introduction

1.1 Project objectives and requirements

The purpose of this research project is to provide data and analysis to support robust monetary values that could be used to measure the costs or benefits of changes in levels of road safety, travel time, reliability of travel time and other factors for which market prices are not available. The objective is to provide values for inclusion in an updated *Monetised Benefits and Costs Manual* (MBCM) (Waka Kotahi, 2021) for use in social cost–benefit and other analyses.

The brief for the research states that the methodology to be used is a national stated preference (SP) survey, with subsequent analysis of data collected.

Although the preference is to identify values for as large a number of factors as possible, there is a trade-off between the number of categories (or attributes) included in a survey and the quality of the data obtained. Including many attributes in a single survey or survey question places a high cognitive burden on respondents and reduces the quality of an individual response. In contrast, spreading the attributes over different surveys, without increasing the total number of people surveyed, is expected to result in higher error margins for the aggregate results. The survey design was a balancing act between the desired number of attributes and the robustness of the data collected.

1.2 Approach

Previous research for Waka Kotahi examined whether a single SP survey could be used to derive monetary values for several different attributes (Denne et al., 2018). That 'pilot study' used an SP approach in the form of a choice experiment (CE) in which a survey provided respondents with route choice options for a trip, where the options differed with respect to their travel costs, travel time, reliability of travel time, safety (risk of death or injury) and other attributes. This approach has been widely applied elsewhere to obtain values, particularly for reductions in the risks of road deaths and injuries.

In this new study we have further tested the 'route choice' approach, in addition to variants of the presentation and format of the survey questions used in the pilot study. This testing phase involved detailed questioning of respondents during and after they had completed a survey to understand their reactions to and interpretations of the questions. This identified difficulties in interpretation that led to some significant changes and to the adoption of a separate community or citizen value-based approach to questions relating to safety. Rather than asking survey respondents to choose between journey route options that differed in levels of personal risk, it asked them to choose between national investment programme options that differed in the total resulting number of deaths and injuries for New Zealand as a whole.

Because the survey was undertaken face-to-face, it was significantly affected by the COVID-19 outbreak, including national and local lockdowns during which surveying was not possible, and delays while survey staff achieved full vaccination status. The final survey had 7,203 participants throughout New Zealand during 2020 and 2021. It includes questions relating to the following value attributes:

- value of travel time (VTT) using private and public transport[2]
- value of reliability (VoR) in travel time

[2] In this report we use a value of travel time (VTT), with reductions in travel time treating this value as a benefit. This is consistent with the approach used in the MBCM. We note that other studies refer to a value of travel time savings (VTTS).

- road congestion
- crowding on public transport
- value of risk reduction (VRR) for fatalities, serious injuries and minor injuries.

1.3 Report structure

This report describes the process to get to the final survey, the survey approach and the results.

Chapter 2 provides an explanation of why a CE approach was used and discusses the differences between route choice and community value approaches. Chapter 3 provides a detailed explanation of the reasons for using a community value approach for safety questions, and chapter 4 provides a brief summary of literature examined during the survey development phase and how this was used in developing initial options for testing for the non-safety questions.

Chapter 5 describes the survey testing phases, with the final survey and its implementation described in chapter 6. The results and the statistical analysis are presented in chapter 7, with a discussion provided in chapter 8.

2 Background

2.1 Approaches to non-market valuation

2.1.1 Stated or revealed preference

Because there is no market in which safety, time and reliability values can be observed directly, alternative approaches to valuation are required. Techniques to derive non-market values are broadly classified into revealed preference (RP) and stated preference (SP) methodologies. Because of its wide use in the literature, in this report we also refer to the term stated choice (SC) as a category synonymous with SP.

RP studies can be used for non-market valuation only when a market for a similar good exists (Haab et al., 2013). Examples might include:

1. value of travel time (VTT) derived from comparing use levels of a conventional road and a quicker route that involves a toll (with different prices – eg, as used by Small et al., 2005); however, the potential for such a study is very limited because there are very few toll roads in New Zealand[3] and a very limited set of prices
2. safety values derived from a regression analysis on sales and prices of vehicles with different safety ratings; however, even if data were available, it would not enable the separate components of safety to be valued (ie, the value of reducing both injuries and fatalities).

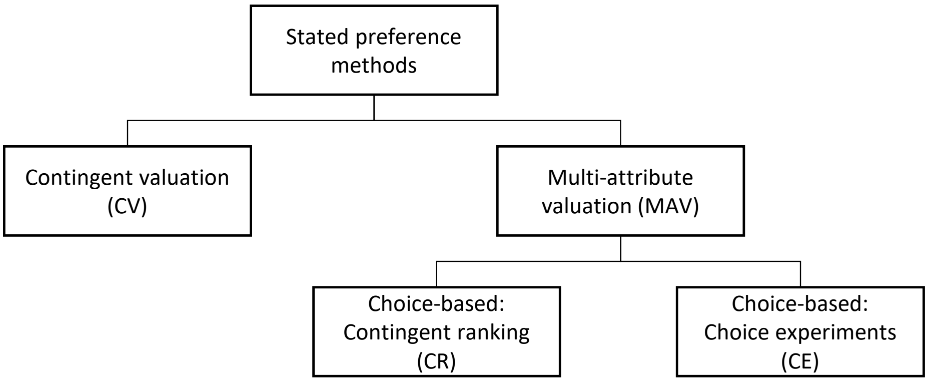
RP approaches only enable a limited number of attribute values to be obtained, whereas SP techniques can be used for all possible values. SP methodologies are not limited by the existence of current markets as they use surveys in which respondents make choices in experimentally controlled hypothetical settings (Hanley & Czajkowski, 2019).

However, despite the benefits of SP, questions remain about whether survey responses to the hypothetical settings are consistent with what actual behaviour would be. Small et al. (2005) suggest this might be the reason for lower VTT in SP than RP studies, although as Fayyaz et al. (2021) note, the Small et al. (2005) SP and RP results used cannot be directly compared because the alternatives and levels are different across data sources. Valuation approaches have changed over time to address this issue, particularly in the shift from the use of simple surveys based on contingent valuation (CV) towards greater use of choice experiments (CEs), particularly because of the opportunity for greater realism in question design.

SP methods include dichotomous choice CV and multi-attribute valuation (MAV) methods (Figure 2.1). CV can value a given outcome, which must be defined prior to the valuation process, but is unable to value attributes of the outcome. MAV methods entail a statement of preferences about two or more alternatives, and systematically vary attribute levels, enabling valuation of the attributes. Those values may then be used to value specific outcomes, which do not need to be defined *a priori*. MAV methods are therefore far more flexible, giving them a significant advantage for policy analysis where new interventions that have different outcomes arise regularly. Some MAV methods ask for the best and worst alternatives, best and second-best alternatives, or a complete ranking of alternatives (contingent ranking). CEs, which ask participants to identify their single most-preferred alternative, are the most common MAV method.

[3] Northern Gateway, Takitimu Drive and the Tauranga Eastern Link.

Figure 2.1 Stated preference methods



We briefly outline the theory behind CV and CE below and the reason for the use of CE in this new survey.

2.1.2 Contingent valuation

CV has been the traditional SP tool for non-market valuation. Examples include a study for Waka Kotahi of the value of enhanced public transport services (frequency, reliability, travel time), in which survey respondents were asked their willingness to pay (WTP) on an ongoing basis (through additional rates or equivalent increase in rents) to obtain these enhancements (Wallis & Wignall, 2012). And in Singapore, a road safety survey asked respondents their WTP for a national programme that would reduce the risk of road deaths by 20% from 40 per million road users to 32 per million (Le et al., 2011).

CV responses differ by question format, all aiming to elicit true preferences,[4] with options including referendum style dichotomous choices (Alberini, 2005; Freeman et al., 2014; Johnston et al., 2017), payment cards (which present respondents with a set of values to choose from – see, for example, Mon et al., 2018), and open-ended questions (Haddak, 2016; Svensson & Vredin Johansson, 2010). Best practice requires strong contextual realism, often with a provision point mechanism[5] (Freeman et al., 2014; Hanley & Czajkowski, 2019; Johnston et al., 2017; Poe et al., 2002).

Despite its wide application, CV has been subject to significant criticisms (Bishop et al., 2017; Carson, 2012; Desvousges et al., 2016; Haab et al., 2013, 2016; Hausman, 2012; Kling et al., 2012; McFadden & Train, 2017). Limitations include that it can value only a single scenario change and is unable to value individual attributes that change between scenarios, although multiple CV studies or questions can overcome these limitations (eg, Jones-Lee et al., 1995) with additional survey participants or greater survey length. Depending on assumptions about how the results will be used, CV also introduces the risk of participants gaming the system. These limitations are amongst the main reasons for development of attribute-based valuation methods such as CE (Bennett & Blamey, 2001; Johnston et al., 2017).

Some studies have attempted to overcome the inability of CV to estimate WTP for multiple attributes, and to decrease cognitive burden by separating the valuation and risk components of the process. Estimating WTP for a single attribute (eg, the value of preventing a fatality) and then estimating the relative values of other attributes (eg, the value of preventing a serious injury) enables estimation of WTP for the other attributes by

[4] Mechanisms where truthful preference revelation is the dominant strategy are referred to as incentive-compatible (Carson et al., 2014).

[5] In a provision point mechanism individuals are asked to donate money to pay for a public good but the donated funding will not be used for the public good unless some lower bound threshold (the provision point) is met (Groothuis & Whitehead, 2009). If the threshold is not met, the donations will be refunded to the individuals. The provision point gives individuals an incentive to reveal their true WTP because of the all or nothing construction.

inference. This approach, sometimes referred to as chaining, typically relies on risk-risk or standard gamble methods to estimate relative values (Balmford et al., 2019; Carthy et al., 1998; Olofsson et al., 2016).

- The risk-risk approach (examples include van Houtven et al., 2008; McDonald et al., 2016; and Nielsen et al., 2019) is a relative value identification method. In one example, Nielsen et al. (2019) used a neighbourhood choice application in which participants are informed of risks of serious injuries and fatalities in their neighbourhood (eg, risk of fatal injury = 50 in 100,000 per decade). Then, assuming they must move, each participant states their preference for moving to one of two alternative neighbourhoods that have different pre-specified risk profiles. Incomes, housing, and all other attributes are identical in each neighbourhood. The risk profile describes the risks of fatal and non-fatal injuries in each location. Based on this initial response, the analyst then varies the risk profiles to identify points of indifference and hence relative values of the two outcomes.
- Jones-Lee et al. (1995, p. 685) provide the following example of a standard gamble to identify the relative magnitude of costs from death or an injury of certain severity.

  Suppose you were in a road accident and you were taken to hospital. The doctors tell you that if you were treated in the usual way, you will certainly experience the consequences shown below [R]. However, they also tell you that there is a different treatment available, but its outcome is not certain. If it succeeds, it will restore you to your normal state of health [J]. But if it fails, you will die [K].

- Jones-Lee et al. (1995) asked participants to identify the probability of failure (π) for the alternative treatment at which 'they would find the "accept/reject treatment" decision to be most finely balanced'. Drawing on expected utility theory, the marginal rate of substitution between state *R* and death is: $mR/mK = \pi$.

Guria et al. (2003) chained values using both the standard gamble and an alternative approach they call matching. The matching approach entailed changing the ratio of differences in outcomes between two public investment projects to identify the point of indifference.

Balmford et al. (2019) argue that concurrently evaluating changes in probabilities and values creates an additional burden on respondents relative to CV applied in non-risky contexts and is likely to be cognitively overwhelming. They assessed the impacts of separating these tasks and estimating the value of risk reduction (VRR) through chaining, first valuing a certain change in health status and then comparing that value to the value of mortality. They also sought to test what they term a 'certainty effect', analogous to lexicographic preferences, in which CV survey participants are unwilling to accept anything but the lowest probabilities of death.

The Balmford et al. (2019) single-chain process is:

1. estimate the value of a specified health change with CV
2. apply the modified standard gamble to assess the relative values of mortality and the specified health change
3. chain those results to estimate VRR.

An alternative 'double chain' process added an additional link to the chain by including two modified standard gambles, one comparing minor and major ill-health, and another comparing major ill-health and death. In the double-chain approach, CV provides estimates of the value of minor ill-health, and the value of a statistical life (VoSL) is derived through a process that is analogous to, but expanded from, the single-chain method.

Whereas policy evaluation occurs *ex ante*, addressing the benefits of avoiding an accident or its impacts, chaining involves *ex post* analysis because an accident has already occurred, which is fundamentally different (Hojman et al., 2005).

Results from CV alone and chaining produce different value estimates (Balmford et al., 2019; Carthy et al., 1998; Jones-Lee et al., 1995). Jones-Lee et al. (1995) identified highly significant differences in estimated marginal rates of substitution derived through chaining. They argued that the reason is likely to rest in problems with the CV estimates, including insensitivity to small changes in risk, insensitivity to injury severity, and failure to isolate risks to other people. The large differences reported in Carthy et al. (1998), which informed the UK Government's VoSL, generated a debate that confirmed the unreliability of chaining, because it resulted in significant inflation of results compared to CV alone (Chilton et al., 2015; Jones-Lee & Loomes, 2015; Thomas & Vaughan, 2015a, 2015b). More recently, Balmford et al. (2019) found strong evidence that single chaining is susceptible to inflation of values, resulting from double-counting utility from own-health. Double chaining exacerbates value over-estimation. The unknown upward bias introduced by chaining signals that this method is unreliable and should not be used.

Whether chaining occurs or not, the question arises of whether CV without chaining is suitable for estimating VRR. The issues of applying CV to VRR can be categorised as issues relating to (i) risk, (ii) scenario framing, (iii) strategic responses, or (iv) constructed responses. Bishop and Boyle (2019) concluded that generally CV can be a reliable tool in decision making, but its accuracy will be specific to the application. The evidence of insensitivity to scope and scale, and inability of participants to process risk-related data, strongly suggest that CV may be an unreliable method for transport-related VRR.

2.1.3 Choice experiments

CE differs from CV in the use of multiple attributes in individual survey questions and the less direct way in which values are derived. In contrast to a CV survey, which might ask a respondent to choose between the status quo and a scenario with a faster journey and a higher cost, a CE survey will ask respondents to choose between options that differ across several attributes simultaneously, with subsequent analysis of the results used to estimate parameter values.

The reason for using a CE, rather than, say, simpler CV or RP analysis, is that it:

- mimics real decisions that people face
- enables collection of data on a large number of attributes of interest to transport policymakers and can collect these simultaneously so that trade-offs are made explicitly
- is the most statistically efficient method.

Many recent transport-related VRR studies have adopted a CE approach in the form of a vehicle route choice. These ask participants to choose between two or more route alternatives that vary systematically over several attributes, including mortality risk and/or risk of injuries of varying severity, as well as other attributes such as travel time (Antoniou, 2014; de Blaeij et al., 2002; Flügel et al., 2015; Flügel et al., 2019; González et al., 2018; Hensher et al., 2009, 2011; Hojman et al., 2005; Iragüen & Ortúzar, 2004; Niroomand & Jenkins, 2016, 2017; Parumog et al., 2006; Rizzi & Ortúzar, 2003, 2006a, 2006b; Rouwendahl et al., 2010; Veisten et al., 2013). It is suggested this approach closely resembles the real decisions that people make (Hensher et al., 2009; Hojman et al., 2005), in contrast to approaches that ask more directly for WTP (Mouter et al., 2017). An example is shown in Figure 2.2 in which the choice game requires respondents to choose between options that differ in terms of the mode (car or public transport), cost (in Swiss francs for fuel, ticket or toll), convenience (number of transfers for the public transport option) and travel time.

Figure 2.2 Route choice task including mode, travel time and cost

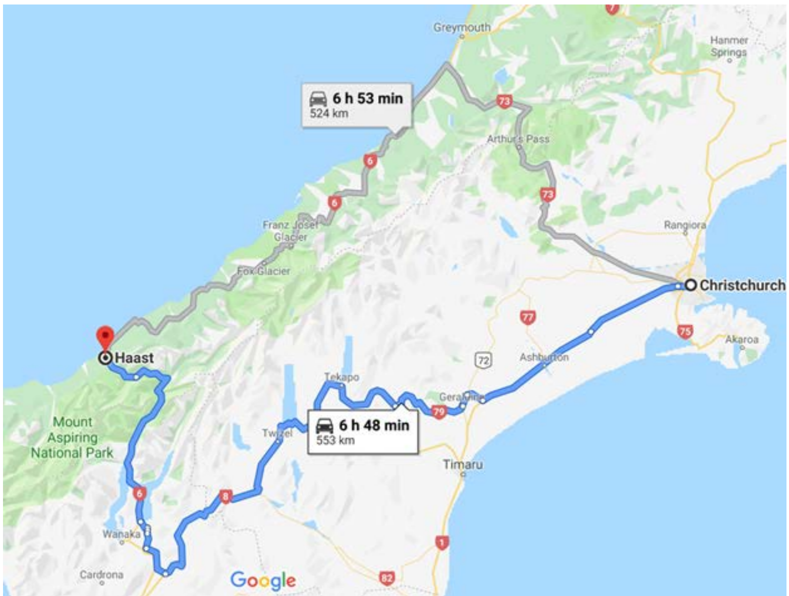
	Car	Rail, Bus Tram
Travel time (min)	10 min	25 min
Fuel / ticket cost	0,40 CHF	1,60 CHF
Number of Transfers		2
Access time public transport		10 min
Interval		10 min
Toll	0,40 CHF	
Parking fee	2,00 CHF	
	❑ ← Your Choice →	❑

Source: Reprinted from Vrtic et al. (2010, p. 112).
Note: CHF = Swiss franc.

In addition to realism, another significant benefit of these studies is that attributes can be valued concurrently – for example, travel time, congestion, and reliability, in addition to values of fatality and injury risk changes (González et al., 2018; Hensher et al., 2009, 2011; Hojman et al., 2005; Niroomand & Jenkins, 2016, 2017).

Recent growth in access to online planning tools will have made route choice based on journey attributes more familiar to the public. Figure 2.3 illustrates a route choice for a journey from Christchurch to Haast from Google Maps. There are two quantitative attributes: distance and expected journey time. The eastern route is 29 km further, but travel time is 5 minutes less. This is a real route choice, so there are other attribute differences that some travellers will be aware of but others will not (such as scenery and road conditions) that will affect route choice decisions. Because of this, hypothetical route choices are often used as these allow the analyst to control these other attributes.

Figure 2.3 Attribute-based route choice



CEs are claimed to reduce the opportunity for strategic responses relative to CV because the purpose of the survey is not obvious to participants (although they may speculate on what it is), the strategy required to strategically influence outcomes is not obvious to participants, and protest ‘zero’ or extreme responses are not possible (Alberini, 2019; Bateman et al., 2002; Bennett & Blamey, 2001; Birol & Koundouri, 2008; Hoyos, 2010; Parumog et al., 2006).

2.2 Private or community values

In CEs the framing of choices is important, and there are two broad approaches. One frames the questions for respondents as private choices with the impacts borne by them personally. Route choices are an example, as discussed above.[6] An alternative approach frames questions as public or social choices in which the costs and benefits might be borne by the community. For example, a Singapore study described by Hess et al. (2017) provided survey respondents with choice games that we understand were similar to that shown in Table 2.1.

Table 2.1 Accident SP game for car users (Singapore)

	Programme A	Programme B
Change in annual tax burden ($)	$25	$50
Number of minor car injuries per year	2,000	1,650
Number of major car injuries per year	50	75
Number of car deaths per year	12	10
Given this I would choose	A	B

A similar approach, with fewer attributes, has been trialled in a series of studies by Mouter and others (eg, Mouter et al., 2017).

Under this approach, people are being asked to consider what is best for the community and not just themselves. This raises the question of whether this is the appropriate approach for collecting VRR data. People often appear to have different preferences when acting as a citizen than when acting as a consumer (Ackerman & Heinzerling, 2002; Arrow, 1950; Marglin, 1963), but as Svensson and Vredin Johansson (2010) suggest, in cases when the intended use is for public policy purposes, the WTP eliciting scenario should be public (see also Abelson, 2008; Andersson et al., 2019; Sagoff, 1988).

Framing questions as social choices has been used in several studies, particularly relating to safety and the VoSL (Mouter & Chorus, 2016; Mouter et al., 2017, 2018; Mouter, Cabral et al., 2019; Hultkrantz & Svensson, 2012; Hess et al., 2017). The social choice approach avoids the requirement for participants to appropriately process small probabilities. It also avoids the imposition of ‘home grown’ probabilities because participants evaluate their own risks as different from (usually less than) the norm, and hence create their own risk attribute levels, which are unknown to the analyst.

One of the significant differences between the approaches is the treatment of safety and risk. The route CEs present respondents with options that are interpreted as different personal risks to them of an injury or fatality; their choice of route is based on how they balance differences in death and injury risks from differences in other attributes (eg, time and costs). In contrast, the community-based CEs, such as in Table

[6] For example, Rizzi and Ortúzar (2003) note that, in a route choice survey aimed at drivers, safety has the dimension of a private good and there is very little room for an altruistic choice; the survey respondent chooses between two existing road options.

2.1 above, do not include a risk element. The choice is between monetary costs to them and total numbers of deaths and injuries for the whole community, where if numbers reduce it is assumed their risks of death and injury fall along with the risk for all other members of the community. The choice task is equivalent to those used frequently in valuing changes in collective environmental outcomes such as cleaner water (eg, Tait et al., 2016) that the respondents and others might enjoy. The values, measured as WTP/person, are multiplied by the relevant population size to generate total (social) value.

In chapter 3 we set out the reasons for adopting a community-based survey in more detail.

2.3 Attributes for inclusion

2.3.1 Suggested inclusions

The attributes suggested for inclusion and which were tested in this project are:

- time-related attributes
  – travel time
  – reliability of travel time
  – time in road congestion
- safety attributes
  – risk reduction for fatalities and injuries
- public transport service quality
- cost.

2.3.2 Limitations

The number of attributes to include in the survey(s) is determined by the trade-off between maximising the number of values that can be collected for updating the MBCM and ensuring data quality and confidence.

A disadvantage of increasing numbers of attributes in CEs is the cognitive burden imposed on participants, and the potential irrelevance of some attributes to some participants. This can result in attribute non-attendance (ANA), which occurs when participants ignore some attributes when making their choices (Hanley & Czajkowski, 2019), particularly in online surveys (Sandorf, 2019). The form of ANA may vary across participants and is not directly observable by the analyst. That is problematic when choice analysis assumes participants attended to the offered levels of all attributes. Methods for addressing ANA during data analysis include latent class models (used to detect unobserved heterogeneity) and models that include only those attributes attended to by the individual. The second approach requires recognition of ANA in the data collection process, questioning participants about any attributes they did not attend to when making their choices.

In testing of the survey for this new study, ANA issues have proven to be critical for safety risk questions. This is because of the very low levels of risk of fatality or injury for any individual trip (and the extent to which drivers believe they are in control of these risks – see fuller discussion in chapter 3).

In addition to the complexity of the choice tasks, the number of choice tasks determines the time taken to complete the survey. Ideally surveys should be long enough so respondents can learn and exhibit their actual preferences consistently, but not so long they tire and begin satisficing (providing ‘good enough’ answers rather than optimising).

Swait and Adamowicz (2001) observed that when respondents were asked to complete a similar task repeatedly, levels of random variation and inconsistency initially decreased before increasing again. These findings suggest levels of repetition and survey length are important design factors. Galesic and Bosnjak (2009) found respondents were faster at answering questions asked later in a survey; this may be partially attributed to respondents' satisficing towards the end of an extended survey, or respondents becoming more skilled at answering the survey in terms of both preference and institutional learning.

Carlsson and Matinsson (2006) found respondents can generally handle a significant number of questions (or choice sets) (up to 12) before there is a noticeable reduction in response rates.

The pilot survey tested different numbers of choice tasks, finding that 10 choice sets with seven attributes per route definition appeared to be a good combination for the general population audience, but that increasing complexity or length beyond this is likely to reduce the quality of engagement and hence increase the random error in responses.

The optimum number of attribute levels is identified during the experimental design process.

2.4 Choice of approach

In this study we have adopted two survey approaches to the collection of attribute values.

1. A community-based CE is adopted for safety questions that estimates the WTP for changes in road safety (numbers of deaths and injuries).
2. A route choice CE is adopted for all other attributes.

In the next two chapters we provide further detail on the reasons for these decisions.

3 Safety benefits in choice experiments

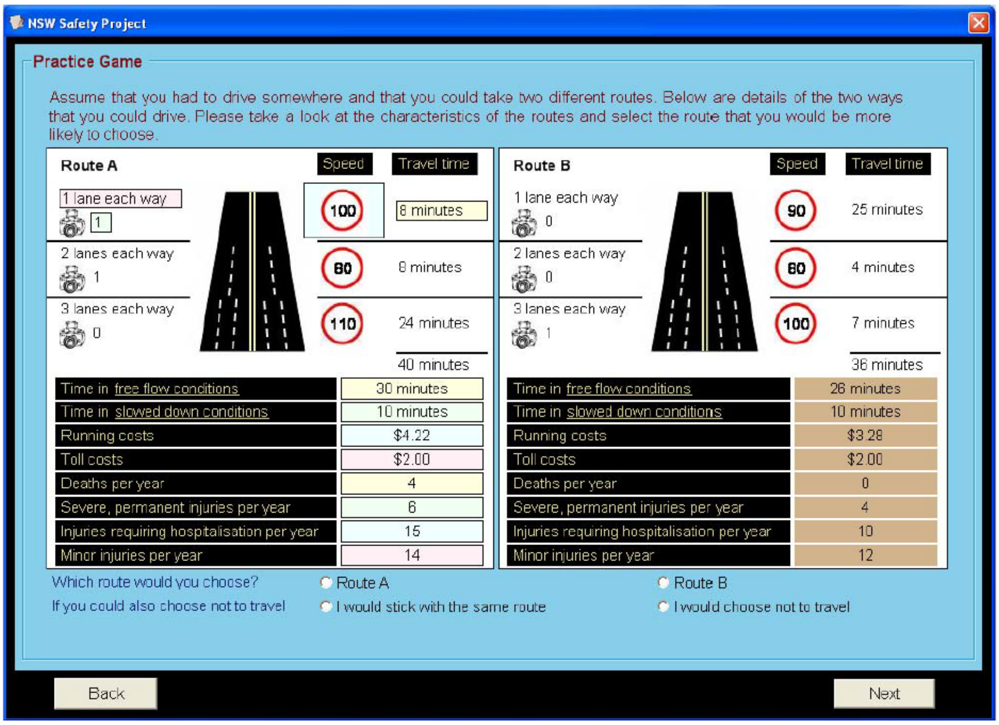
3.1 Route choice studies

In this section we provide additional detail used in the decision to use a community-based approach rather than a route choice for safety risk questions.

3.1.1 Safety attributes in route choice experiments

Recent studies to identify values of risk reduction (VRRs) have tended to use choice experiments (CEs) based on route choices (Bahamonde-Birke et al., 2015; Obermeyer & Hirte, 2019). For example, Hensher et al. (2009) include the number of deaths and injuries per annum as attributes for an individual route, alongside costs, time and other attributes (Figure 3.1).

Figure 3.1 Example stated choice screen from Hensher et al. (2009, p. 10)



Safety issues have also been included in studies that have extended to pedestrians (Hensher et al., 2011; González et al., 2018), where the issues include the differences between the driver's willingness to pay (WTP) to avoid hitting a pedestrian (including some unspecified form of altruism if it exists – see Jones-Lee 1991, 1992) and the pedestrian's WTP to avoid being hit.[7]

3.1.2 Risk levels and risk perception

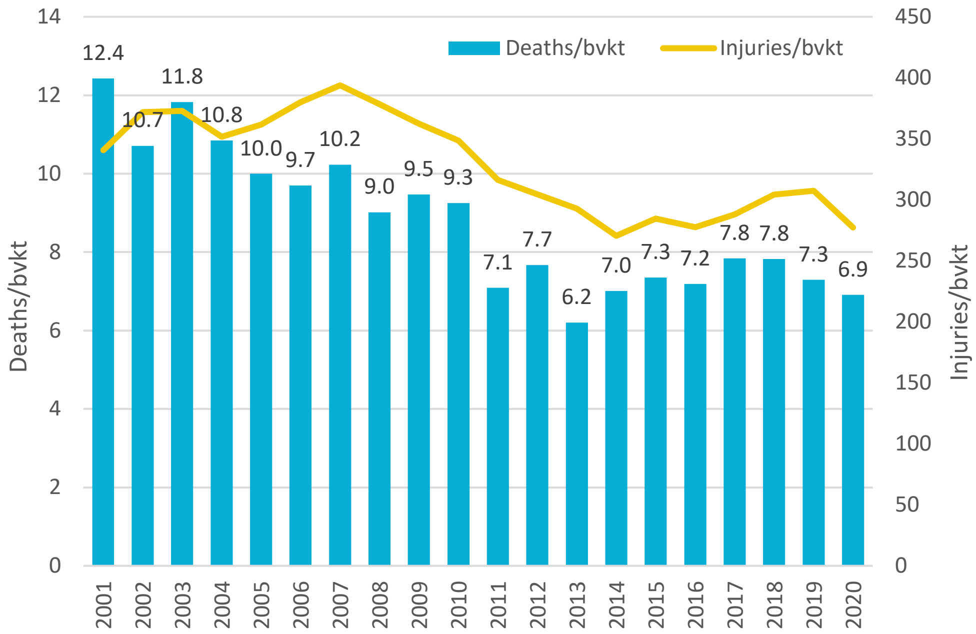
It is important to obtain values close to realistic risk levels so the values can be applied to real-world investments or other decisions. Realistic risk levels are used to address concerns that WTP may not be linear with risk level (as noted by Krupnik et al., 1997). At the same time, risk levels need to be sufficiently

[7] González et al. (2018) consider that the values assessed should be only the private values of the driver, excluding any altruistic concerns for the pedestrian.

high for survey respondents to respond to. Hess et al. (2017) note that values used by Hensher et al. (2011), with deaths per annum on a road varying from 0 to 5, were unrealistically high for a single road in Singapore. The same issue applies in New Zealand.

Current risks of injuries and fatalities are extremely low for any individual, let alone for any single trip. In New Zealand, the average risk of a fatality in recent years is approximately 7–8 deaths per billion vehicle kilometres travelled (bvkt) (Figure 3.2); driver-only risks are lower still, at close to 5 deaths per bvkt.[8] Based on these rates, someone travelling 10,000 kilometres per year would have an annual fatality risk of approximately 1 in 13,000 (or a death every 13,000 years); their annual risk of injury is approximately 1 in 333. Their risks of death and injury on an individual trip of a few kilometres will be extremely small, and their perceived risk may be lower still, depending on their perception of their risk profile.

Figure 3.2 Deaths and injuries per billion vehicle kilometres travelled in New Zealand



Source: Ministry of Transport: Time series of historical casualties and crashes

Questions have been raised over whether the levels of fatality and injury risk are so small in route choices that they will be given little weight in decisions relative to time, reliability and cost attributes – this was observed and noted as a concern in the pilot study[9] and again in testing for this study. In addition, concerns have been raised about whether drivers think crash risks are in their control and that their personal risk is lower than average (Mouter et al., 2018). Again, these attitudes were observed in survey testing.

In addition, unlike cost and speed attributes in a CE, risk is more difficult to present to respondents because individuals differ in their perception of the risk to them in a given situation and in a way that is less easily controlled for (Veisten et al., 2013). To the extent possible, the task of survey design includes ensuring that all individuals have the same perception of the risk.

[8] Joanne Leung, Ministry of Transport, pers comm.

[9] We provided data on exposure to risk in both absolute and relative measures. Other studies have only provided relative measures and inferred exposure exogenously. Hence, other studies rely totally on perceived risk, without knowing what that perceived risk is.

Jones-Lee and Loomes (1995) noted how context made a significant difference both to perceptions and values of risk. They note that psychologists have provided extensive evidence indicating that the public's preferences, perceptions, and attitudes towards risk may vary widely over different hazards, with factors such as familiarity with the risks leading to lower perceived risk and lower WTP to reduce risks.

Hojman et al. (2011) note that the risk literature (eg, Bronfman & Cifuentes, 2003) suggests controllability and knowledge are of particular importance in risk perception. Risk is perceived to be greater when a person considers they do not have control over an activity, such as a public transport trip. Risk is perceived to be less if the activity is more familiar.

3.1.3 Risk presentation in stated choice surveys

There are three main approaches to presenting risk of mortality or injury in choice models for estimation of the value of a statistical life (VoSL) or the VRR.

1. The **Risk** approach: the mortality attribute is the probability of mortality for each journey on each route. This probability is typically extremely small.
2. The **Quantum** approach: the annual number of mortalities on each route, *without* exposure information. Annual mortalities per route is typically a small integer.
3. The **Exposure** approach: the annual number of mortalities on each route, *with* exposure information. Average annual daily traffic counts are the most common form of exposure information.

Table 3.1 lists recent published studies by the approach employed in each and the nature of exposure information provided to participants. Many of these studies have common authors, with Rizzi and/or Ortúzar involved in 10 of the 16 studies. Because method and authorship are related, the prominence of various methods does not indicate the breadth of support for them. Below we explore the advantages and limitations of these approaches.

Table 3.1 Risk presentation approach in selected stated choice studies

Approach	Study (authors include Rizzi and/or Ortúzar)	Study (other)	Exposure information
Risk		• Alberini & Ščasný (2013)	Baseline risk
Quantum	• Flügel et al. (2019) • González et al. (2018) • Hensher et al. (2009) • Hensher et al. (2011) • Iragüen & Ortúzar (2004) • Rizzi & Ortúzar (2003) • Rizzi & Ortúzar (2006a) • Rizzi & Ortúzar (2006b)	• Niroomand & Jenkins (2016) [See also Niroomand & Jenkins (2018)] • Niroomand & Jenkins (2017) • Parumog et al. (2006)	None
Exposure	• Veisten et al. (2013) [Data from this study also in Flügel et al. (2015)]	• Antoniou (2014)	Average annual daily traffic counts
	• Hojman et al. (2005)	• Rouwendahl et al. (2010)	Annual flow

Reliable stated choice (SC) studies require that participants understand and attend to all attributes in the choice alternatives. Correct interpretation of responses requires that the analyst and the respondents treat choice attributes the same way. Where choice study participants ignore, confuse or scale attributes, subsequent analysis that fails to incorporate those types of response will draw false conclusions.

There is substantial evidence that people have difficulty understanding very small risks and interpreting differences in them (Jones-Lee et al., 1993; Krupnick et al., 1997). This is the main reason that authors reject the Risk approach – used only by Alberini and Ščasný (2013) – in Table 3.1.

Recent research has used total annual number of deaths and/or injuries as the risk attribute. In most cases, there is no exposure information, instead relying on subjective risk evaluations (11 of 16 studies in Table 3.1). Sometimes accompanying information, such as average annual daily traffic counts on the route or annual vehicle kilometres travelled on the route, allows participants to evaluate their risk exposure (Antoniou, 2014; Hojman et al., 2005; Rouwendahl et al., 2010; Veisten et al., 2013). It is unknown whether participants do that evaluation or not.

3.1.4 Rizzi and Ortúzar approach

Most recent studies have adopted the Rizzi and Ortúzar perceived-risk framework (Hensher et al., 2009, 2011; Iragüen & Ortúzar, 2004; Rizzi & Ortúzar, 2003, 2006a, 2006b). Hojman et al. (2005) and Veisten et al. (2013) adopted the same framework, but also provided exposure information. Rizzi and Ortúzar argue against use of objective risk information (eg, percentage risk of a fatality per trip) because SC survey participants do not understand this information, and because individuals sometimes perceive their own risks to be different to objective risks. Instead, they advocate for describing risk in terms of total number of events, with participants free to construct their own perceived risks. For example, Figure 3.1 shows two routes, with different trip times and costs for the respondent but fatalities and injuries presented as the number on each road each year, with no presentation of the risk for an individual or a trip.

Rizzi and Ortúzar (2003) note that individuals' personal risk evaluations are context-dependent because of differences in the degree of individual control in different contexts. This is an argument for context-dependent risk valuation, which controls for this effect. Hence, modal choice studies are somewhat problematic. However, even within a specific context, individuals develop personal subjective risk evaluations based on matters such as whether they consider themselves a safe driver, or drive a safe car. Rizzi and Ortúzar (2003) argue that, within-context, these effects will cancel out; relative risk will be constant, although the individual may consider their personal risk to be different to objective risk. They defined the risk attribute as the 'number of accidents during the year in which at least one car occupant dies'. They informed participants of the actual number of fatal crashes on the existing route (eg, 27 fatal crashes per year). Their utility function was a second-order polynomial in risk, but the polynomial term was not statistically significant.

Rizzi and Ortúzar (2006a) make the case for using total fatalities on the route, rather than risk. In Footnote 9 (p. 483) they acknowledge a difficulty with this approach:

> *One referee noted that the two approaches are mutually consistent only when the respondent, when evaluating the number of crashes, has the correct aggregate flow in mind (ie s/he would value an extra fatal crash per year differently if s/he were to make the only trip on that road that year, than when millions of trips would be made on that road). In this sense, although a formulation in terms of number of crashes may sound more natural and easy to understand than a formulation in terms of probabilities to most respondents, the cognitive burden may not become any lighter. Unfortunately, whether the yearly flow indeed affects the valuation of an additional crash cannot be tested with the data available for this paper.*

Rizzi and Ortúzar (2006a, pp. 476–477) suggest that people understand risk in terms of numbers such as those in the media relating to 'the number of fatalities, the number of seriously injured victims, etc., and the frequency with which a certain road is involved in crashes'. Given this, they suggest rational individuals can derive subjective probabilities for events without thinking in terms of objective probabilities:

> *Respondents can combine their personal knowledge of how dangerous a regularly used route is, with objective information related to the number of victims, in order to arrive to a subjective probability of their own likelihood of being involved in a fatal crash. … It is sufficient that people have well-defined preferences in terms of fatal crash reductions to arrive at subjective probabilities.*

They go on to suggest:

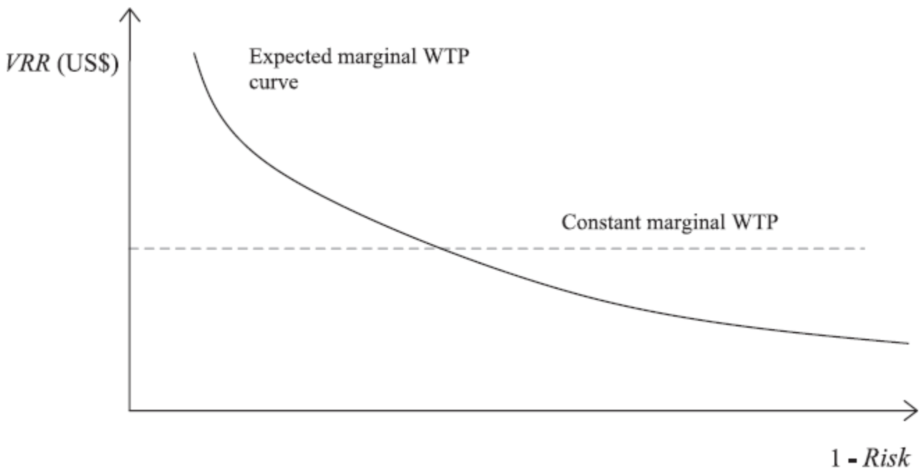
> *In an SC study people have to choose from pairs of alternative routes, the risks of which can be only marginally different from the baseline risk (ie marginally different from the baseline number of crashes). Special care has to be taken to ensure that the alternative routes in the hypothetical choice scenarios are of a similar nature to the route people had driven through. This way one can be confident that respondents are projecting the sample-selection route baseline risk (whatever their risk conceptions are) onto the routes in the choice pairs. Hence, the modelling results should yield plausible monetary values for small changes in a neighbourhood of the baseline risk level of each route but not at all for major changes in road safety.*

Thus, they decided to use the number of fatal crashes as a risk-proxy variable.[10]

Rizzi and Ortúzar (2006b, pp. 72–73) add some additional insights, noting that in analysing the marginal rate of substitution (MRS) between income and the risk of death:

> *one would expect that the WTP for risk reductions should increase with the baseline risk of death and/or with the risk reduction offered, as shown in [Figure 3.3] by the full line curve representing marginal WTP. In this figure the x-axis represents one minus risk, or safety; that is, the abscissa of a point corresponding to a safer road on the curve would be located to the right of the abscissa for a less safe route. We choose this rather 'bizarre' convention (namely, having baseline risk decreasing in a rightward direction) to obtain the usual downward sloping demand curve for our good (that is, safety or risk reduction).*

Figure 3.3 Expected VRR pattern as a function of baseline risk



Source: Reprinted from Rizzi and Ortúzar (2006b, p. 73).

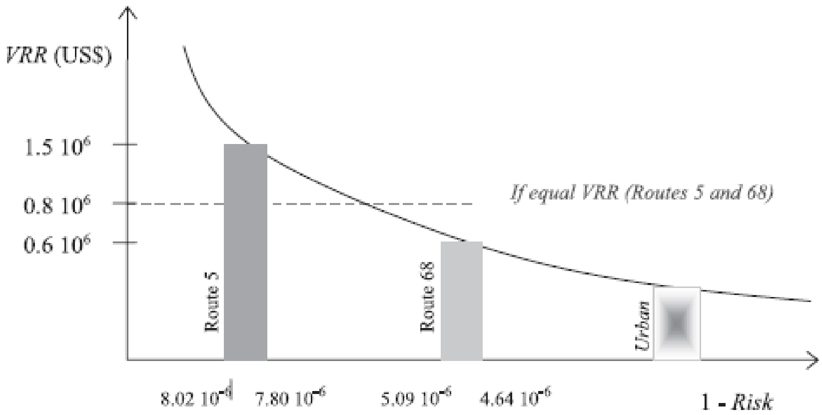
This conceptualisation (risk reducing from left to right on the chart) is consistent with a non-linear specification of risk in the utility function, in which marginal WTP is an increasing function of risk. This figure

[10] Also see discussion in Rizzi and Ortúzar (2006b, pp. 80–81).

clearly identifies that Rizzi and Ortúzar view the absolute level of risk to be an important determinant of WTP, but this is not a feature of their estimated utility function. This view causes little harm if risk changes are small (and therefore can be approximated linearly),[11] and if all participants correctly believe actual risk is in the same small region of the *x*-axis in Figure 3.3. Where different individuals (even those with identical preferences) have significantly different expectations about risk exposure, they will value a similar change in risk somewhat differently. Hence, it is important to either (i) locate participants on the same, correct portion of the risk-exposure (*x*) axis, or (ii) identify the individual's beliefs about risk-exposure and analyse responses using a utility function that does not impose constant marginal utility to account for significant differences in perceived exposure.

Rizzi and Ortúzar's (2006b) field results are supportive of non-constant marginal WTP. In their Figure 3 (included as Figure 3.4 below) they report estimated WTP for routes with different risk exposure. As anticipated, marginal WTP increases with exposure.

Figure 3.4 Implied VRR curve from our three datasets (un-scaled values)



Source: Reprinted from Rizzi and Ortúzar (2006b, p. 83).

This result is not surprising in situations where travellers are familiar with and perceive differences in relative risks of the three routes in the different studies, as was the case here. However, the veracity of WTP estimates still depends on study participants locating themselves at the appropriate point on the *x*-axis. There is no evidence that they did that. Hypothetical routes, about which study participants have no prior knowledge of riskiness, add even more complexity and uncertainty.

Hensher et al. (2009, 2011) adopted the Rizzi and Ortúzar approach, observing that 'MRS depends on personal risk perceptions'. They advocate for describing attributes as total number of events, rather than risks, in order to estimate the 'subjective value of fatal injury (by class) reductions' (Hensher et al., 2011, p. 73). They do this on the basis that doing so 'embodies the definition of community WTP for road safety as the sum of individual marginal rates of substitution between income and number of fatalities and injuries'. Further justification for this approach is avoidance of the assumption of zero covariance between the MRS and change in risk (Rizzi & Ortúzar, 2006a).

Hensher et al. (2009, p. 694) state 'rather than asking people to place a value on risk reductions, they should be asked to value a reduction in fatal or injury class crashes; we believe this task is far easier from the

[11] Within an order of magnitude, the correlation between risk and ln(risk) is very high. For example, when risk takes five equally spaced levels between $\alpha.10^n$ and $3\alpha.10^n$ the correlation coefficient is .9744, irrespective of the values of α or *n*.

respondents' standpoint'.[12] However, stated route choices do not ask participants to 'value a reduction in crashes'; they simply ask people which route they would choose to travel, given the stated number of crashes on each route. Participants' responses will have no effect on the number of crashes. Their expectations about their own exposure to risk, as well as the value of risk to them, condition their route choices. Participant exposure for a single journey is not the number of annual accidents; it is the number of annual accidents divided by the annual traffic flow on the route.

The following hypothetical, exaggerated examples illustrate the importance of exposure. Suppose the following (Table 3.2).

Table 3.2 Different numbers of deaths, traffic flows and risks

Highway	Annual deaths	Annual traffic (trips)	Risk (per trip)
A	3	10^6	3×10^{-6}
B	1	10	1×10^{-1}

Assuming no other differences in trip attributes and despite the greater annual frequency of mortalities on Highway A, the best choice is to select Highway A, where there is a much lower chance of dying. Note, however, that this is not the scenario envisaged by Rizzi and Ortúzar, whose case studies used alternative routes with similar traffic flows.

Continuing with the theme of differences in exposure, suppose now the choice is between two highways with equal annual deaths, but unequal flows (Table 3.3).

Table 3.3 Same numbers of deaths; different traffic flows and risks

Highway	Annual deaths	Annual traffic (trips)	Risk (per trip)
A	3	10^6	3×10^{-6}
B	3	10^1	3×10^{-1}

Motorists aware of exposure would not be indifferent between the equal annual number of deaths on these two highways. Again, the risk is much lower on Highway A. Suppose now the choice is between two highways with equal flows, but unequal numbers of deaths (Table 3.4). This is the Rizzi and Ortúzar scenario.

Table 3.4 Different numbers of deaths and risks; same traffic flows

Highway	Annual deaths	Annual traffic (trips)	Risk (per trip)
A	3	10^6	3×10^{-6}
B	1	10^6	1×10^{-6}

Now it is clear that the risk of death is lower on Highway B; in fact, it is one-third the risk of death on Highway A. The change in risk by choosing Highway B is -2×10^{-6}. All else being equal, participants should prefer Highway B. Relative risk is sufficient to make the right choice, but is contingent on the *ceteris paribus* assumption. Suppose now the choice is again between two highways with equal flows (although smaller this time) and unequal numbers of deaths (Table 3.5).

[12] We note that this is what we have done in this new research.

Table 3.5 Different numbers of deaths and risks; same (but lower) traffic flows

Highway	Annual deaths	Annual traffic (trips)	Risk (per trip)
A	3	10^3	3×10^{-3}
B	1	10^3	1×10^{-3}

As before, Highway B is safer, with only one-third the mortality rate of Highway A. However, the change in risk by choosing Highway B is now -2×10^{-3}, which is 10^3 times the previous example. Consistent with Rizzi and Ortúzar (2006b), one might expect this difference in exposure to affect WTP to avoid Highway A, with WTP substantially higher in the low traffic case.

In short, it is not reasonable simply to use relative mortality risk to estimate VoSL unless the value of risk is a function of the logarithm of risk, rather than of the level *per se* (see Box 3.1 below).

Box 3.1 Logarithmic utility of risk

Consider a utility function in which utility is a linear function of the logarithm of risk.

$$U = \alpha + \beta \ln(Risk) + \gamma Cost$$

$$dU = \beta/Risk.\delta Risk + \gamma\delta Cost = 0 \quad \Rightarrow \quad \delta Cost/\delta Risk = -\gamma Risk/\beta$$

WTP is a function of the absolute level of risk. WTP is more to reduce a near certain risk than for a reduction of the same magnitude of a very low probability risk.

$$\begin{aligned}
\text{Let } Risk_1 &= \phi Risk_2 \\
dU &= U_1 - U_2 = 0 \text{ for constant utility} \\
&= \gamma(Cost_1 - Cost_2) + \beta(\ln(Risk_1) - \ln(Risk_2)) \\
&= WTP + \beta/\gamma(\ln(Risk_1) - \ln(Risk_2)) \\
WTP &= -\beta/\gamma(\ln(\phi Risk_2) - \ln(Risk_2)) \\
&= -\beta.\ln(\phi)/\gamma
\end{aligned}$$

Under this specification, WTP is a function of relative risk, independent of the absolute level of risk. This utility specification supports presentation of relative accident risk measures in SC surveys (unanchored by exposure). This is the Rizzi and Ortúzar approach. However, Rizzi and Ortúzar do not estimate utility functions that are logarithmic in risk, yet they derive VoSL by scaling by objective risk, so this approach is inconsistent. While Rizzi and Ortúzar may well be correct that drivers assess relative risk, their statistical models do not reflect that. Their results can only be correct if (i) they estimate utility functions logarithmic in risk (they don't) and (ii) participants do address relative risk independent of exposure to risk. Rizzi and Ortúzar argue strongly for the latter.

Arguably, participants know about exposure and are able to identify risks from that information. However, our own work during the pilot study showed that it takes some effort to find traffic flow and/or mortality risks, even when one knows about the concept. As the examples above show, if participants get that wrong, then they are valuing something different from the analyst, but the change that participants are valuing is unknown to the analyst.

3.2 Community values

3.2.1 Examples from road safety studies

Despite the simplicity and attraction of the approach adopted by Rizzi and Ortúzar and others, there seems to be significant non-attendance to risk attributes in testing, in addition to the theoretical objections discussed above. We describe the testing experience in chapter 5.

Risk evaluations have been posited as social choice questions in several recent studies with choices focused on investments or public policies that result in changes to an aggregate number of annual deaths and injuries for a road, region or nation.

The social choice approach has the important advantage that it avoids the requirement for participants to appropriately process small probabilities because the numbers of deaths spread across a region or a nation can be both relatively high and realistic. A social choice question also avoids the imposition of ‘home grown’ probabilities that arise from participants evaluating their own risks as different to (usually less than) the average, and hence create their own risk attribute levels that are unknown to the analyst.

The approach used in the survey was that shown in Figure 3.5, which asks the respondent to choose between options with different total national death, injury and cost outcomes. Choosing an outcome in which there are 50 fewer road deaths per annum is a reduction in the risk of death for the respondent and for everyone else in the community.

Figure 3.5 Social choice question for safety attributes

The following questions are about a national highway investment programme

QS1 Which of the following would you prefer?

	Current (no changes made)	Investment option 1	Investment option 2
Deaths (per year)	250	200	250
Serious injuries (per year)	1750	1000	1500
Minor injuries (per year)	6000	5000	6000
Increase in your personal costs (per year)	$0	$200 more per year	$100 more per year

3.2.2 Differences in values

Several studies have tested the differences in values between individual and community-based analyses.

Johannesson et al. (1996) undertook a study in which one group of respondents was offered the opportunity to purchase a safety device to be installed in their cars, while another group was offered a public safety programme (improved road quality) that resulted in the same size risk reduction. The WTP for the private safety device was higher than the WTP for the public safety measure.

A very similar study was undertaken by Andersson et al. (2019), who also found, contrary to their initial expectations, that WTP for a private road safety device was more than twice the WTP for an equivalent (same risk reduction) public safety measure, with no overlap of confidence intervals. The authors discuss, but largely reject, possible explanations such as lack of trust in public provision of the programme in question (or even of public provision in general), strategic behaviour (fearing a rise in taxes) or the possibility of free-riding by others. They conclude:

> *Often WTP is framed as a private safety measure even if the findings are going to be used for public safety measures. In this study we have shown that this may be problematic and that thorough validity tests should be conducted.* (p. 174)

An alternative approach by Mouter et al. (2017) used CEs to test whether people had the same relative value for time and safety if they were being asked as individuals (making route choices) or as citizens (deciding the road or type of roads that the government should spend money on). One example of an experiment in which people were asked to decide as citizens is shown in Figure 3.6. The authors found the share of respondents choosing the safest route is considerably higher in the citizen experiments than in the consumer experiment.[13] This is consistent with our pre-test finding that individuals believe their own risk to be lower than average risks presented to them as attribute levels.

Figure 3.6 Citizen experiment 4

The government decided to build a new road.
The government still needs to decide about the route of the new road.
The government asks you whether you would recommend Route A or Route B for the new road that the government will build. Below you will find the characteristics of both routes.

Assume the following:

- Both routes are 2x2-lane motorways
- Both routes will carry 80,000 trips per day, which means around 29 million trips per year
- 80,000 trips per day corresponds with an average 2x2-lane motorway in the Netherlands
- The routes only differ in terms of travel times and number of fatalities on the road per year
- The routes do not differ in costs, environmental effects and non-fatal accidents, amongst other things
- The government is interested in general preferences of Dutch citizens. Hence, it is not made clear whether or not you would experience any effects (positive and negative) from either of the two routes

Please select the Route which you would recommend to the government.

	Route A	Route B
Travel time	40 minutes	30 minutes
Number of traffic deaths on the road	1 per year	5 per year

Source: Reprinted from Mouter et al. (2017, p. 338).

This example did not permit monetisation but does reveal the MRS. Travel time saving is denominated in minutes per journey, so is not directly comparable to the deaths attribute. To put them both in the same (annual) units requires survey participants to multiply time by exposure (eg, the time difference in this scenario of 10 minutes per trip results in aggregate time saving of 4.83 million hours per year). In the equivalent private evaluation choice game, the mortality risk changes from 0.000000034 per trip to 0.000000172 per trip. Hence, both cases require appropriate calibration (either of time or risk) to enable meaningful comparison and interpretation. They are equally problematic.

This is clearly a very simple approach. Svensson and Vredin Johansson (2010) note the implausibility of changing the number of fatalities without any impact on the number of injuries. If that is true, then participants are likely to construct ‘home grown’ changes for non-included but related attributes.

Another game described by Mouter et al. (2017) uses a private toll in a personal route choice scenario to enable monetisation, with other attributes for travel time and number of deaths. The limitations of excluded attributes and calibration requirements remain. In a social choice valuation scenario, the cost attribute was a ‘one time tax increase for Dutch households’. A summary of the results is presented in Table 3.6 for five different experimental frames.

- Monetisation did not have a significant effect on estimated MRS for 1 & 2, but it did for 3 & 4.
- Framing as total time or time reduction did not have a significant effect (4 & 5 are not significantly different).

[13] A student of Mouter’s conducted an experiment in Indonesia and produced similar results – that is, that consumers tend to choose the faster option and citizens tend to prefer the safer option (Nisa, 2018).

- The experimental design is confounded, so it is not possible to draw inferences about the effects of participant road use (eg, 1 & 3 also differ on proposal and cost attribute, and 2 & 4 differ on proposal). Similarly, it is not possible to isolate effects of specific cost attributes or proposal framing.

Table 3.6 Marginal rates of substitution

Frame	Value	Proposal	Time attribute	Cost attribute	Number of trips/year	Participant uses the road	MRS (death/ time)	SE* (MRS)
1	Private	Personal route choice	Travel time	Toll	29 million	Yes	2.71	0.23
2	Private	Personal route choice	Travel time	None	29 million	Yes	2.53	0.17
3	Public	Recommended new route	Travel time	One time tax increase for Dutch households	29 million	Not stated – anonymous road	5.43	0.53
4	Public	Recommended new route	Travel time	None	29 million	Not stated – anonymous road	16.31	4.35
5	Public	Recommend road project	Travel time reduction	None	29 million	Not stated – anonymous road	10.73	2.49

Source: Adapted from Mouter et al. (2017, p. 343).
* SE = standard error

Reflecting on the results of their analysis, Mouter et al. (2018) suggest car drivers will assign a relatively low value to mitigating accident risk because they believe:

- such risks are trivial on an individual level
- their personal risk is lower than the average risk
- their personal risk is controllable
- they would not be able to distinguish relative safety levels in real life
- their choices for others are more risk-averse than choices for themselves (which, given the above discussion on the work by Jones-Lee and others, is probably invalid) and unlike citizens, they are not explicitly evaluating risky choices for others.

They also suggest individuals believe the government should assign more value to safety because:

- as a citizen they are more prone to base their choices on social norms that prescribe risk-averse behaviour in this context
- governments have a duty of care concerning the safety of the transportation network
- drivers have a relatively high degree of responsibility to reduce their own travel times
- governments should account for drivers' tendencies to choose faster routes by building safer ones
- governments should ensure the safety of the road network because this allows drivers to choose the fastest route without being concerned about the impact of their route choice on accident risk.

In the earlier study, Mouter et al. (2017) commented on possible strategic behaviour. They noted the relative percentage of respondents who always chose the safest option, which might reflect (1) strong (lexicographic) preferences, (2) simplification of the decision task because of boredom or fatigue or (3) political or strategic behaviour – for example, belief that by expressing their preferences in this way they can influence policy decisions. Substantially more respondents always chose the safest option in the citizen experiments, compared to the consumer experiments.[14]

Mouter, Koster et al. (2019) widened the discussion from transport to comment on cost–benefit analysis (CBA) more widely. They suggest CBA for public policies assumes 'consumer sovereignty' – that is, that total social benefits can be measured by aggregating individuals' WTP. However, they note individuals' WTP might not accurately reflect preferences towards public policies. They promote Participatory Value Evaluation as an evaluation framework in which individuals are asked to choose the best portfolio of projects with corresponding impacts for society and themselves, subject to governmental and private budget constraints.

3.2.3 Community values of travel time

In the survey development process we considered whether time could still be included in a community-based survey. The value of travel time (VTT) varies with trip mode, purpose and length (amongst other things). This means any community-based CE that included time alongside injuries and fatalities – for example, government strategy options with different numbers of total national fatalities and trip times – would need to be specific about the trip type for the time variations. However, specifying a single trip type is problematic because it will be more relevant to some respondents than others, and it is likely that the respondent would be making (unknown) assumptions about the change in times for other journey types. Including expected changes relevant to all possible trip options would be far too much information to process. To further complicate matters, for equivalence with annual numbers of fatalities, information would be required on the assumed number of trips taken annually by trip type.

Including time in both social and private choice studies is also unnecessary. Money is an existing numeraire that is common to both choice scenarios. Provided the same individual responds to a survey used to value risk in money terms and value time in money terms, the MRS between time and risk can be estimated.

3.2.4 Active travel mode studies

Because of the difficulty in identifying a payment vehicle in route CEs, active modes (walking and cycling) have been measured using community values.

Hensher et al. (2011) used a CE that examined the preferences of pedestrians for road safety, travel times and cost, but the payment mechanism used was council rates or housing rent increases, and the choice being made was for infrastructure investments to improve safety and/or travel time. For example, a respondent was asked to assume they had to walk somewhere that required crossing a busy road[15] and that there were two different routes. The options differed in terms of characteristics of the road to cross (lanes, speed limit), the crossing type (zebra, traffic lights, overbridge), walking time, and risk attributes (numbers of deaths and injuries per annum); the cost attribute was specified as an increase in council rates or housing rent.

This study involved a mix of public and private costs and benefits.

[14] As discussed by the New Zealand Productivity Commission (2014), pressure for regulatory intervention to deliver levels of risk-reduction that go beyond the level of risk they would accept as individuals (weighing up their individual costs and benefits) may be simply an attempt to obtain benefits (lower risks) while passing costs on to the wider community. These issues are discussed in a transport safety perspective by Denne and Wright (2017).

[15] The survey was limited to those who had walked along and crossed a main road as a pedestrian in the previous week.

- The costs are the respondent's share of the municipal charges for walkways that are public goods, and others will obviously benefit from their use in addition to the respondent.
- The time attribute is for the pedestrian's own trip.
- The risk presentation is for the total annual number of pedestrian deaths and injuries per annum on the road that must be crossed. The authors note that 'the WTP is the pedestrian's marginal rate of substitution between income and number of annual pedestrian fatalities in the road environment' (p. 77), which they describe as a community value – that is, it is a social benefit. To convert to the same per trip basis as the time attribute, these results were used to calculate an estimate of the VRR per pedestrian trip expressed as a WTP per pedestrian activity per month. To do this, they use the annual total kilometres walked by pedestrians, and risk data in terms of the numbers of fatalities and injuries in each death or injury class per annum for persons walking.

Whereas normally a route CE is choosing a private good with little room for an altruistic choice, here people could choose a particular investment option partly for the benefits for others.

A survey in France (Haddak, 2016; Haddak et al., 2014) asked respondents to envisage contributing financially to the implementation of a local project to improve the safety of all road users in the Rhône Département. Each respondent was given information relating to a project to reduce the risk of minor, serious or fatal injuries. However, although Haddak examined the relationship between WTP and injury level, and between income and WTP, the results were not used to identify a VRR.

3.3 Approach used

The approach adopted for including safety attributes in the final survey was that shown in Figure 3.5 above.

4 Route-based choice tasks

4.1 Route choice experiments

Route choice games have been used widely in choice experiment (CE) studies to develop values for road travel and public transport use attributes. In this survey, a route choice approach was adopted for car, motorcycle and public transport trips, for the following attributes:

- travel time savings, including in congested traffic (car trips only)
- reliability of travel time
- public transport attributes, including frequency and standing vs sitting for travel time.

The questions were used to collect data applying to different trip purposes, in addition to modes. In this section we include a discussion of relevant literature that has been used in designing the survey questions.

4.2 Trip purpose

The *Monetised Benefits and Costs Manual* (MBCM) currently uses three trip purposes for which it provides different values of travel time (VTTs):

1. working – trips carried out in the course of paid employment
2. commuting – trips between home and work
3. other – all other non-work trips (ie, other than commuting).

Defining the trip purpose in a CE provides necessary realism,[16] in addition to enabling separate values to be assessed.

Although included in the MBCM and in some other studies,[17] our perception is that survey questions for working trips are likely to produce unreliable results because the driver, passenger or public transport user will usually not face the financial costs of the trip if these are paid for by the employer. In some circumstances this problem might also apply to company vehicles used for private use (a principal-agent market failure problem – International Energy Agency (IEA), 2007); for owner-operated businesses, work use of vehicles will directly impact the driver, but there will be a tax offset to complicate any WTP estimates, even if they can be identified. Because of the confounding factors associated with work trips, these have not been included in the survey.

Commuting trips are readily defined, but other (non-work, non-commuting) trips might be further differentiated to test whether the VTT differs with whether the trips are time-dependent (eg, dropping children at school) or flexible (eg, shopping). Such differentiation increases the number of options, and the longer the list of purposes included, the wider the confidence intervals will be for any sample split, even more so when there are interactions between journey length and purpose.

[16] The purposes need to be relevant to the respondent – for example, not suggesting a commuting trip for someone who does not commute. Introductory questions can be used to ensure respondents are provided with a relevant purpose for their choice tasks.

[17] Batley et al. (2019) noted that the Department of Transport directed them to examine business travel from the employer's and employee's perspective – for example, an employee was asked to report a WTP representative of his/her employer's interests. They noted that they did not use the Hensher approach, which we understand to be the employee's value of substituting travel time for leisure.

The following options are included:

- commuting trips
- other time-dependent trips
- other flexible trips.

'No trip' options are not included.

4.3 Travel mode

The MBCM includes the VTT and other values for drivers and passengers of cars, motorcycles, commercial vehicles and public transport, and for active modes (walking and cycling).

One approach is to combine all attributes into a multi-mode CE as shown in Figure 2.2. The choice game required respondents to choose between options that differ in terms of the mode (car or public transport), cost (in Swiss francs for fuel, ticket or toll), convenience (number of transfers for the public transport option) and travel time. However, for this study the nature of the attributes, and the limitation on the number that could be included into a single choice game without adding too much to cognitive burden, meant separate questions were used, particularly between private and public transport.

4.3.1 Car drivers and motorcyclists

Car drivers and motorcyclists are easily incorporated in a CE as a standard route choice game covering time and cost attributes. Both make choices about routes to get to a destination, weighing up time and cost attributes, and are affected by related attributes including heavy traffic/congestion and trip time reliability.

Although values might differ with the (assumed) number of vehicle occupants and their relationship to the driver, this is much reduced by excluding safety questions from the route choice approach (where concern with passenger safety might be included in the choice), although there may still be a concern with how other passengers will be affected by trip duration. Because respondents may differ in whether they assume any other vehicle occupants, the simplest approach is to instruct drivers to assume they are driving alone, and this was the approach adopted.

4.3.2 Car passengers

The MBCM includes separate VTTs for drivers and passengers of cars. Ian Wallis Associates (IWA) examined the potential for passenger surveys of the VTT (IWA, 2014). Their review of the international literature suggested there were few studies that had analysed passenger VTT, but for those that did exist, values were generally lower (typically by 25–40%) than driver values, and that these relativities can differ substantially by trip purpose. IWA presented options for how this might be analysed in New Zealand, including via a new car passenger survey, while noting the difficulty of finding a suitable payment mechanism that would be regarded as realistic for survey respondents. In this study the approach used is to ask a passenger to assume they:

- were the only passenger in a vehicle driven by someone else
- contributed a stated amount (50%) to the cost of the trip
- had a say in the route chosen.

In survey testing this was found to be realistic and understood by respondents.

4.3.3 Public transport

Choice experiments have been applied to public transport users with respondents being offered trip options with different journey times, fares and quality factors (eg, rider comfort). The MBCM includes separate values for the VTT for public transport, including values that differ with whether the passenger is standing or sitting. It also includes values for improvements in reliability and increases in service frequency, and for quality parameters. We discuss these in sections 4.4 (Time), 4.5 (Travel time reliability) and 4.6 (Public transport questions) below.

4.3.4 Active modes

The use of route CEs to derive values for active modes is more difficult because there is no suitable payment mechanism: there is no fare paid and no appreciable variable cost of a trip.

We have not identified any road safety stated choice (SC) studies that have interviewed cyclists. Hensher et al. (2011) used a route CE to examine the preferences of pedestrians for road safety, travel times and cost, but the payment mechanism used was council rates or housing rent increases and the choice being made was for infrastructure investments to improve safety. Niroomand and Jenkins (2017) used a similarly designed experiment in North Cyprus. This approach is different from the route choices used with drivers, where the benefits of the route choice are limited to the drivers themselves. The investment choices are public goods where others will benefit from their use.

We have not included CE questions for VTT for active modes, but users of these modes are included in the social choice safety survey.

4.4 Time

4.4.1 Components

The MBCM values include total travel time savings where

Total travel time savings = base travel time benefits for improved flow or shorter trips
+ travel time benefits for reduced traffic congestion (if applicable)
+ travel time benefits for improved trip reliability (if applicable).

In this section we discuss the collection of data for base travel time benefits and traffic congestion. Improved trip reliability is discussed in the next section (4.5).

4.4.2 Base travel time

Savings in travel time are one of the key benefits of transport investments for which values are required. The base travel time is simply the WTP for a change in the average travel time. VTT is expected to vary with:

- total trip duration (Abrantes & Wardman, 2011) reflecting a declining marginal VTT – Wallis et al. (2015), for example, found that unit values for travel time savings for a two-hour trip are in the order of half those for a 20-minute trip
- travel mode (Abrantes & Wardman, 2011), reflecting factors that include the differences in transport users (eg, income) and travel time quality (Batley et al., 2019; Douglas & Wallis, 2013; Wardman et al., 2016), including the opportunity to work while travelling (Wang & Hensher, 2015).

Although initially differentiated by mode, since 2013, the *Economic Evaluation Manual* (NZ Transport Agency, 2018), and more recently the MBCM, has used a single (‘equity’) base VTT across all modes, although it varies with trip purpose. Travel mode is dealt with in the new survey by using separate questions for car use and public transport.

An ‘exploratory survey’ by Wallis et al. (2015) focused on respondent preferences between smaller and larger time savings per trip, and between time savings on longer compared with shorter trips. They found survey respondents had a strong preference for a given time saving on a shorter (*c*.20 minutes) trip than a longer (60–120-minute) trip, confirming the need to differentiate between trip length in the choice tasks.

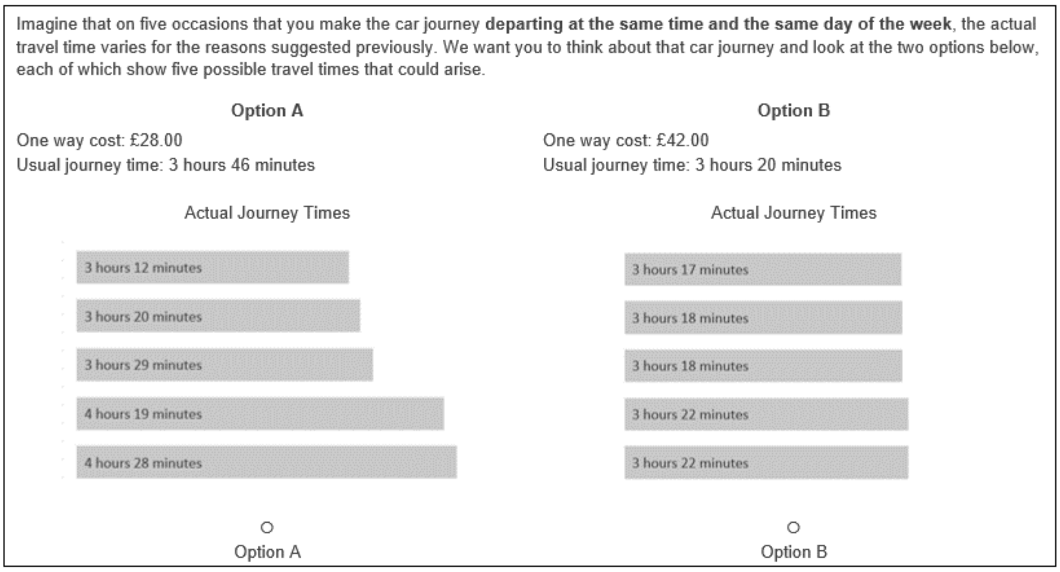
The survey used in this study differentiates by travel mode and travel purpose when collecting data on VTT. Different trip lengths are used in choice tasks to explore this relationship.

4.4.3 Distribution of time

The real experience of people making trips is that journey times are often different each time the trip is made. Presenting trip time as a distribution rather than as a single average time might approximate this. This was trialled in the pilot study, with respondents being shown charts of distributions but with no information on average travel times. This was found to be too complicated, with respondents generally not being able to identify differences in average trips, despite this being vital for estimating VTT.

Figure 4.1 shows an example in which the distribution is used alongside a statement of the usual time. This provides a visual presentation of variability and provides the useful information on average travel times, assuming ‘usual journey time’ is interpreted in the same way as ‘average journey time’.

Figure 4.1 Time vs cost vs reliability experiment (car example)



Source: Reprinted from Arup et al. (2015, p. 32).

This approach was tested and adopted in the new survey, although the complexity of this presentation (and cognitive burden) meant that it could not be included with all other attributes of interest for private transport. We discuss this in the next chapter.

4.4.4 Time in heavy traffic

The MBCM provides increments to the VTT for driving in congested conditions. A substantial congestion level is defined relative to a non-congested/free-flow situation as a situation that would add at least 10% to the typical peak period trips (of typical trip length) travel times.

Surveys have used slightly different terminology to derive estimates of VTT in different conditions. For example:

- Hensher et al. (2009) included trip time in ‘free-flow conditions’ and in ‘slowed-down conditions’, described as:

Freedom to manoeuvre within the traffic stream is noticeably restricted, and lane changes require more care and vigilance on the part of the driver. Minor incidents may still be absorbed, but the local deterioration in service will be substantial. Queues may be expected to form behind any significant blockage. (p. 11)

- Arup et al. (2015) used the terms ‘light traffic’ (you can travel close to the speed limit most of the time, but you have to slow down every so often) and ‘heavy traffic’ (your speed is noticeably restricted and frequent gear changes are required).

A decision was made to use the heavy traffic terminology but to adopt a different definition from Arup et al. because of the greater prevalence of automatic cars in New Zealand compared with the UK. A suggested alternative explanation is ‘Traffic is not free flowing: it is often moving slower than the speed limit and speed changes often.’

We tested different approaches to including congestion in the surveys with the wording using the term ‘heavy traffic’.

4.5 Travel time reliability

4.5.1 Measurement approaches

In addition to VTT, there is a value of reliability (VoR) for road users – that is, the extent to which they can have certainty over trip duration. This applies both to car drivers and passengers, and to users of public transport.

Some studies estimated higher values for reducing travel variability by a given amount (eg, 5 minutes) than for reducing the scheduled journey time or the average travel time. Where measured, studies have shown travel time reliability to be valued anywhere from half the value attributed to time savings to double the VTT (Bhat & Sardesai, 2006; Li et al., 2010). Tilahun and Levinson (2010) found higher income individuals valued reliability less; they reasoned this was because of the greater flexibility provided by higher income jobs.

The most common approach to measuring the VoR is the mean variance model. It assumes people would prefer to leave later and arrive earlier, and that utility (wellbeing) is accumulated at home and at work, but not during travel (see, for example, Fosgerau, 2016). This approach, which is adopted by the MBCM, defines reliability as a measure of the unpredictable variations in travel times that are experienced for a trip undertaken at broadly the same time every day. Under the mean-variance model, the value of improving variability is calculated as a change in the standard deviation (SD) of travel time for a route and the VTT. In section 3.7 of the MBCM, improvements in reliability are valued as 0.9 times the product of the VTT ($/h) and the change in the SD.[18]

In addition to the mean variance model, Li et al. (2010) describe two others:

- The scheduling model considers the consequences of unreliable travel time. Unlike the mean-variance model, which assumes travel time variability leads to the loss of utility by itself, the scheduling model considers disutility is incurred when not arriving at the preferred arrival time, either early or late. The model uses separate parameter values for variability either side of the preferred arrival time.
- The mean lateness model is becoming the ‘standard’ approach for analysing reliability for passenger rail transport in the UK. Travel unreliability is measured by the mean lateness at departure and/or arrival, while the mean earliness (negative lateness) is not considered.

[18] The 0.9 factor is the value of reliability based on a typical urban traffic mix. For projects with a significantly different vehicle mix, it is suggested evaluators use 0.8 for cars and 1.2 for commercial vehicles.

Although trip lengths vary in both directions around the mean (early and late), the nature of costs is quite different. Small (2012) notes that the traveller suffers not only the occasional costs of unexpected delays, but also suffers some routine costs of choosing an earlier schedule to allow a safety margin. Since the precise costs are not known in advance, it is usually assumed that travellers choose a departure time that minimises the expected value of the sum of costs, taking into account the distribution of travel times.

For public transport, the MBCM also uses minute-late ratios. These are the equivalent time values to being a minute late – for example, at a ratio of 2.0, one minute late is treated equally to a trip taking an additional two minutes. In the MBCM they vary from 2.0 to 6.4 (MBCM Table 31).

4.5.2 Survey examples

Surveys have tended to concentrate on lateness rather than variability around the mean (early and late).

Outwater et al. (2010) provided public transport choices with 10% of trips experiencing different levels of lateness – for example, ‘1 in 10 trips experience delay of 2 mins or more’ or ‘1 in 10 trips experience delay of 10 mins or more’. Their analysis suggests a VoR equivalent to 0.5 minutes per work trip and 0.8 minutes per non-work trip.

Consistent with Outwater et al. (2010), the pilot study (Denne et al., 2018) used choice tasks with delays to the normal time: 10% of trips were said to be delayed by 5–20 minutes for short trips (of 20–40 minutes average duration) or by 15–40 minutes for long trips (of 180–240 minutes). This enabled estimates to be made of a value (or cost) of lateness. Bates et al. (2001) express concerns about the use of simple presentations of lateness, which they suggest are often misinterpreted by respondents. For example, a study they conducted of presentations such as ‘one in ten trains is 20 minutes late’ suggested people interpreted that phrase to mean all other trains were on time (or that no other train is more than 20 minutes late). However, the analyst might make the interpretation that 10% of trains are at least 20 minutes late. They note work that presented rail passengers with three alternative patterns of lateness (Table 4.1), each of which had a mean lateness of 5 minutes, but with very different SDs.

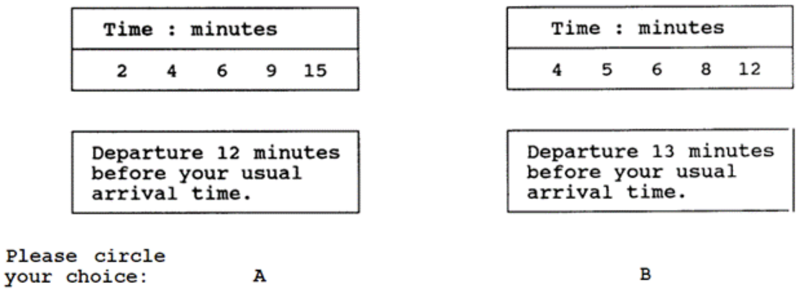
Table 4.1 Alternative patterns of lateness

Series of delay patterns	Standard deviation	Ranked first (%)	Ranked last (%)
0, 0, 5, 6, 8, 7, 6, 4, 5, 9	2.86	38	47
0, 0, 0, 0, 0, 0, 25, 5, 10, 10	7.75	6	29
0, 0, 0, 0, 0, 0, 0, 0, 20, 30	10.25	56	24

Source: Adapted from Benwell and Black (1984) as cited in Bates et al. (2001, p. 216).

As Bates et al. (2001) note, 56% of respondents preferred the pattern of arrivals with the greatest SD but with the lowest probability of being late. Small et al. (1995) developed a survey for commuters in Los Angeles. An example screen is shown in Figure 4.2. For each pair of options, it provides five different travel times and a departure time.

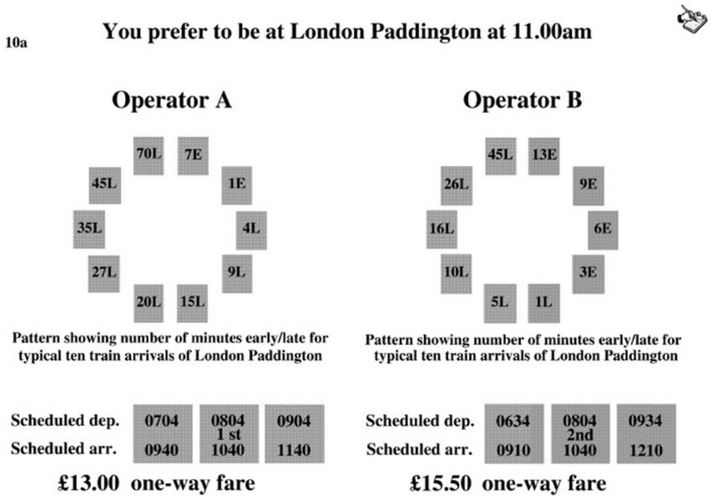
Figure 4.2 Commuter survey with time distributions



Source: Reprinted from Small et al. (1995, p. 128).

Bates et al. (2001) built on these ideas for their work with train times. They used a survey in which respondents were offered a distribution of early and late arrival times (Figure 4.3).

Figure 4.3 Presentation of train use choice task with time distribution



Source: Reprinted from Bates et al. (2001, p. 222).

Bates et al. (2001) suggest respondents had more difficulty in correctly interpreting this presentation than simpler options, and this is a concern over the use of detailed time distributions.

In New Zealand, a survey of train users in 2007 provided survey respondents with choices between two services with different certainties of pick-up times, travel times and fares; a separate set of questions asked about certainties of arrival times (Vincent, 2008). The analysis was used to estimate the VTT and the value of lateness; a valuation of one minute's average lateness was estimated at approximately 3 to 5 times in-vehicle time (IVT). Douglas (2017) undertook an SC survey of train users in Auckland, Christchurch and Wellington in 2012–13 that included choice sets differing in frequency, travel time and fare.[19] It was used to derive a WTP to save time estimated at $7.50/hr compared to a value of IVT of $9.89/hr.

Different approaches were tested for measuring VoR, as discussed in chapter 5 below, but the final approach adopted was similar to that used by Arup et al. (2015) (Figure 4.1).

[19] It also included a train station quality attribute.

4.6 Public transport questions

4.6.1 Components

In addition to average travel time and cost, several other factors were considered for inclusion. Studies have examined the values or passenger preferences for other factors relating to public transport, and often these are factors used by local authorities to monitor the service performance of public transport companies (eg, Litman, 2017). The factors are:

- availability (daily hours of service)
- service frequency (how many trips per hour or day)
- speed (particularly compared with car travel)
- reliability (how well a service follows published schedules)
- comfort (whether passengers have a seat and adequate space)
- security (feelings of safety)
- affordability (user costs relative to their income and other travel options)
- information (ease of obtaining information)
- cleanliness (including minimal mess, dirt, unpleasant smells, and graffiti and vandalism)
- aesthetics (appearance of transit vehicles, stations, waiting areas and documents).

The MBCM includes values for increased service frequency, interchange reduction benefits, other public transport user benefits and infrastructure and vehicle features. It includes a higher VTT for standing onboard buses and trains than when seated for commuting. It also includes values for bus and rail attributes expressed in terms of IVT equivalents. For trains, this includes values for the presence of an attendant, a quiet ride, onboard toilets and air-conditioning; for buses it includes whether a pass needs to be shown, the driver's attitude, and clean windows.

There is a very long list of attributes – too many to be included in this survey because of the potential cognitive burden and attribute non-attendance (ANA). The study therefore includes priority attributes identified by Waka Kotahi.

4.6.2 Survey examples

An SC survey in New South Wales asked passengers to choose between three bus journeys differing in bus fare, travel time, reliability, walking time and nine other service quality attributes relating to the vehicle, bus stop and information (Hensher, 2015; Hensher & Prioni, 2002).

Rose et al. (2013) conducted an SC survey of buses in Australian cities that included the frequency of services and in-vehicle crowding, but no cost attribute. It was addressing whether people valued improved bus quality enough that they would walk further to use a higher quality bus. Their survey included distance to stop, frequency of service, total journey time and crowding level (defined as the percentage of seats occupied and the number of people standing).

AECOM (2009) undertook a series of stated preference (SP) surveys to understand the relative importance of 'soft factors' (Table 4.2) to the travel choice decision. They addressed this using five separate surveys. An additional 'Fare Simplification' SP survey was conducted by Institute for Transport Studies (ITS) Leeds as part of the study.

Table 4.2 Soft impacts/outcomes: definitions

Soft impact	Measures
Quality of In-Vehicle Experience	Vehicle: age, ease of access, seating quality, cleanliness, entertainment, CCTV. Driver: training to achieve politeness and smooth ride.
Increased Awareness of Service Availability	Conventional and unconventional marketing approaches.
Improved Knowledge Whilst Travelling	Real-time information, public service announcements on vehicle.
Ease of Use	Smart cards, travel cards, ticket structure, low floor vehicles.
Quality of Waiting and Walking Experience	Shelters, bus stations, ticket machines, seating, information provision, CCTV, staff presence, lighting.
Safety and Security	CCTV, staff presence, lighting.

Source: Adapted from AECOM (2009, p. 4).

Each respondent received two SP games, which differed with the area, main mode of transport and whether they had a choice of using a quality and/or non-quality bus. The SP experiments were grouped into six questionnaire designs. An example question is shown in Box 4.1.

Box 4.1 Example question: Leeds bus survey

Now I want you to imagine that you have to make a single journey, by bus, from Leeds City Centre to [X] at about 11 o'clock on a cloudy but dry morning.

I want you to imagine that there are two bus services available to you, each going from a different bus stop.

Can you look at the green sheet please.(pause)

Imagine that you have choice between services A and B. (pause)

> You will see that service A has a journey time of 20 minutes, involves 10 minutes walking to and from bus stops and has a fixed fare of £1. (pause)
>
> Service B has a journey time of 30 minutes, involves 15 minutes of walking and has the 'as now' fare – which you have estimated as [XXX]. (pause)

Faced with this choice, which would you use?

Source: Adapted from AECOM (2009, Appendix H).

Douglas (2016) reviews several international studies that address attributes relating to in-vehicle quality and the quality of stations, bus stops and staff. This included many of the examples listed above. In addition, in a study for Tranz Metro Wellington, passengers were asked to rate attributes for the rail station they used, including signage, seating, and staff friendliness, but no valuation was involved (Douglas Economics, 2005, as cited in Douglas, 2016).

Douglas (2017) reports on a New Zealand SC survey that included different levels of station and vehicle quality. Stop/station and bus/train quality were defined using a five-star system, alongside verbal descriptions. The verbal descriptions are not included in the report, so it is difficult to know whether each participant would have understood the ratings and their relativities in the same way.

4.6.3 Approach adopted

The number of attributes was limited by presumed cognitive burden. The final survey for Waka Kotahi was limited to service frequency and standing/sitting time, in addition to time, reliability and cost attributes.

4.7 Cost

Cost is a necessary component of the CE because to obtain attribute values (VTT, VoR etc) it is necessary to understand the MRS between the other attributes and cost.

For route CEs, previous studies have used different cost bases.

- Antoniou (2014) included fuel consumption, maintenance and capital costs for cars.
- Rizzi and Ortúzar (2003) and Hojman et al. (2005) used toll costs alone, rather than including the full costs of trips.
- Hensher et al. (2009) included vehicle running costs and toll costs.

For conventional vehicles, fuel price is a more obvious cost per trip than other running costs, although including a wider set of running costs is both justifiable (they are legitimate components of variable costs) and helps to increase the costs used in the CE, thus making costs more likely to be considered in making a route choice. For electric vehicles (EVs) operating costs per kilometre travelled are extremely low. However, currently EVs are only approximately 1% of the total light vehicle fleet,[20] so they are unlikely to affect the survey results. This may become more of an issue in future surveys, although EVs are also expected to face road user charges, which will be a significant per kilometre cost.

In other countries road tolls have been used in CE surveys, and this enables greater variation in cost. However, there are few toll roads in New Zealand, so they are not familiar to many drivers, and early testing suggested when tolls were included some participants thought the intention of the survey was the introduction of tolls and sought to game the responses accordingly. Fuel and other running costs are used for car drivers and motorcyclists in this survey.

For surveys of car passengers, the assumption given to participants is that they pay some of the fuel costs and have an influence on the route choice.

Ticket prices provide an obvious mechanism for public transport questions.

Active transport modes have no suitable payment mechanism. Usually there is no payment made or appreciable running cost for an individual trip. Surveys that have included costs have usually asked about investment costs for infrastructure (eg, walking bridges), which provides more of a community rather than a private value, as discussed in section 2.2. Given the payment mechanism problems, we have not included active modes in any route CEs used to estimate a VTT, but walkers and cyclists have been asked the same (community-based) safety question as all participants.

4.8 Summary

The attributes and other components, including the differentiation by mode and purpose for the individual attributes, are shown in Table 4.3 for time and other non-safety attributes.

[20] Ministry of Transport Fleet Statistics

Table 4.3 Non-safety attributes and other components

Mode	Time	Purpose	Trip length	Cost
• Car driver • Car passenger	• Average time • Variability • Time in heavy traffic	• Commuting • Time-dependent • Flexible	• Short • Medium • Long	• Fuel and running costs • (50% contribution for passenger)
• Bus passenger • Train passenger	• Average time • Service frequency • Sitting or standing (public transport)	• Commuting • Time-dependent • Flexible	• Short • Medium	• Public transport fare

5 Survey development

5.1 Development process

The survey went through several rounds of testing.

- An initial testing phase comprising 60 in-depth, face-to-face interviews was conducted to refine and test the questionnaire;[21] 22 interviews were completed in early 2020 prior to the initial COVID-19 level 4 lockdown, followed by another 11 interviews conducted post-lockdown in July 2020. A further 27 interviews were conducted in September and October 2020, after final alterations were made to the questionnaire following discussions with the study steering group.
- A survey pilot with 100 participants was conducted in December 2020.
- The final survey was launched in February 2021, with a review of the data and some further adjustments after 758 surveys had been completed in April 2021.
- The survey field work was completed at the end of October 2021.

5.2 Initial testing

5.2.1 Route choice presentation of safety questions

The initial testing in the pilot study and for this new study used a route choice approach for all attributes, including safety. It provided participants with a risk of a fatal, serious or minor injury. This was presented as the number of deaths per year on each route and the number per billion vehicle kilometres travelled (bvkt) on the route.[22] Participants were also informed of how the death rate for each route compared with an average New Zealand highway. Hence, participants received information on (i) actual risk exposure, (ii) risk outcomes and (iii) relative risk.

A second approach tested was a format similar to that used by Rizzi and Ortúzar and others – that is, a simple presentation of the number of deaths and injuries per year (Figure 5.1). Cognitive burden was reduced by removing trip reliability values (fastest/slowest/usual travel time), which were addressed in a separate choice experiment (CE).

[21] Interviews were conducted in Auckland, Hamilton, Cambridge, Wellington and Porirua covering participants who live in the city centre, suburbs, smaller towns and rural areas.

[22] As noted by Douglas (2021), an error in the pilot study meant the wrong exposure information was provided to respondents. Numbers of deaths and injuries were provided for each route and described as numbers per 100 bvkt, whereas more realistic numbers would have been per bvkt. The error confirmed the problem with this approach, that these risk levels were ignored by respondents. A corrected version was used in initial testing.

Figure 5.1 Route choice presentation of safety question

QD1 Which route would you prefer to take?
Please select one only

This trip is for: [insert & colour code trip purpose]

	Route one (2000 vehicles per day)	Route two (2000 vehicles per day)
Usual travel time	3 hours 15 minutes	3 hours 0 minutes
Heavy traffic	30 minutes	20 minutes
Trip cost	$15	$20
Fatalities (per year)	2 deaths	1 death
Serious injuries (per year)	10	2
Minor injuries (per year)	20	5

The feedback received was that, in the context of a relatively long list of choice attributes, people generally paid low or no attention to the numbers of deaths and injuries and justified this approach because:

- a significant majority did not think of safety in terms of the number of deaths per road or the risk of death per road – at most they might think about whether a road feels safe or not
- for some participants, the numbers of deaths and injuries felt too low to be of significance in their route choice decisions, so they felt their chances of being affected personally were too low
- many of those tested thought the deaths and injuries would not apply to them, or they could avoid deaths and injuries because they are good drivers or have a safe car.

In short, most participants indicated they did not include the death and injury values in their choice decision. Even where the safety attributes were considered, we had little confidence they were treated in the same way as other attributes such that the marginal rate of substitution (MRS) could be derived in any meaningful way.

In addition, the overall cognitive burden was found to be high. To simplify the decision task, people used simple heuristics – for example, some ‘eyeballed’ the values and concluded that one column looked roughly worse than the other without fully considering what each attribute represented, or the trade-offs implied by their choices.

5.2.2 Social choice presentation

5.2.2.1 Including time attribute

A further round of testing trialled an investment question for a programme of improvements. This presented the choice between different public good outcomes with differences in total national road deaths and injuries, alongside differences in costs paid via taxation or an equivalent payment mechanism. For this we started with a CE with cost, death and injury attributes, as well as a travel time attribute. The challenge for this approach is that the safety questions were focused on national-level effects, whereas it was not credible to develop a similar time-attribute question (eg, for all trip times to improve). Options were examined that included reduced trip times for typical journeys, or for trips that they do not take such that the benefit was of a public good nature, similar to that for the safety improvements. However, this was problematic because it was not possible to control either the type of trip envisaged by the respondent or the extent to which they would benefit personally, even if told they would never use the road.

In addition, participants who felt the Government has not been able to reduce travel time with past road improvement projects, particularly in Auckland, were sceptical of this attribute. This, in turn, led to scepticism about death and injury attributes. Some felt that if time reduction projects have been ineffective, then safety improvements may also be ineffective. Hence, this approach was not adopted.

5.2.2.2 Safety and cost only

Because of the difficulties of including a time attribute in a social choice question, testing shifted to a presentation as shown in Figure 5.2. The assumption was made that the trade-off with time could still be explored by asking the same individuals questions relating to time (in a route choice format) and safety (in a social choice format). This presentation retains three alternatives, which has significant statistical advantages. In testing, the version with no time attribute worked well, with lower cognitive burden and trade-offs well understood.

Figure 5.2 Social choice question for safety attributes

The following questions are about a national highway investment programme

QS1 Which of the following would you prefer?

	Current (no changes made)	Investment option 1	Investment option 2
Deaths (per year)	250	200	250
Serious injuries (per year)	1750	1000	1500
Minor injuries (per year)	6000	5000	6000
Increase in your personal costs (per year)	$0	**$200** more per year	**$100** more per year

5.2.2.3 Deaths

This was the most important attribute to most participants and was generally well understood; however, some were sceptical that the road toll can be reduced through roading infrastructure investment, feeling that human behaviour is the biggest contributor to accidents. In later application of this question, interviewers stressed that respondents should assume it was possible, with some examples given.

Most participants tend to think of benefit to the community at large, rather than personal safety/safety of own family, where a tendency to think ‘this won’t happen to me’ is prevalent (and the reason the personal risk choice games are ineffective).

5.2.2.4 Injuries

Serious injuries were taken into account more than in individual route choice games, particularly as higher attribute numbers could be presented realistically in a national context. Where the differences in number of deaths between choices are more pronounced, participants tended to pay less attention to injuries.

Minor injuries were regarded as less important and described as ‘just a fact of life’ by some, with similar injuries thought to occur in other settings (home, sports, outdoor activities etc).

5.2.2.5 Cost

The cost attribute was generally well understood. A small number questioned how a tax would or would not apply to them, but accepted this when explained by the interviewer. Differences between willingness and ability to pay were observed. Interviewers reinforced that choices need to be based on affordability as well as on benefit preference, to mitigate overstatement of ability to pay.

5.2.2.6 Other issues

Plausibility/credibility is an issue, with some not believing the values (ie, too small a reduction in deaths/injuries for the cost), some not having faith in the ability of government to effectively improve safety, and others blaming driver behaviour rather than roads. Some indicated no investment in all games despite ability to pay because they could not accept the idea of road travel investment reducing deaths and injuries. This exercise requires a certain level of suspension of disbelief, and interviewers were briefed to instruct participants to respond with how they would feel if a guarantee could be made that the changes would provide the reductions in deaths and injuries specified.

Some participants mentioned other potential impacts of any investment programmes, including increased disruption due to roadworks, particularly in Auckland where some participants felt their travel has been impacted regularly by roadworks already and want a ‘return to normal’ on the routes they use. This group tend to be able to see the long-term benefits, but question how long they will be disrupted in the meantime.

5.2.3 Time

Average travel time was well understood as a realistic and relatable attribute considered regularly in reality by participants. Options were tested that included combining average travel time and time reliability using a distribution. Later versions split these, limiting the attributes in the time-related CEs to average time and time in heavy traffic, as shown in Figure 5.3.

Figure 5.3 Private vehicle (car driver/passenger or motorbike rider)

	Route one	Route two
Average travel time	45 minutes	30 minutes
Heavy traffic	0 minutes Of your travel time is spent in heavy traffic	10 minutes Of your travel time is spent in heavy traffic
Cost	$3	$5

5.2.4 Reliability

The route reliability questions assess willingness to pay (WTP) for travel time reliability. Several options were tested, including presenting (1) both the fastest and slowest possible trip times alongside an average, (2) just the slowest possible time alongside an average, and (3) other versions that included frequency of fast or slow trips. There were various comprehension difficulties, in addition to a difference between the way the values were presented and common parlance. The final iteration for car drivers, passengers and motorcyclists presented the variation in travel times across several trips per route, alongside an average time overall (Figure 5.4).

Figure 5.4 Route reliability (both public transport and private vehicle)

This approach, with accompanying explanation by the interviewer, was the most effective of those tested. But this attribute needs careful explanation by the interviewer as graphs are not easily understood by some. However, generally participants understood the conceptualisation of time variability when it was carefully explained.

The graphs were tested with five, six and ten trip times depicted. Some participants read five trips as equating to days of the week and, if in ascending or descending order, would start to consider days they would travel on or avoid. Randomisation of the five trips would avoid this issue, but it makes the graphs more difficult to read and comprehend correctly. Options presenting either six or ten trips mitigate this assumption; however, the version with ten trips was too visually overwhelming for some. A six-trips option avoids respondents ascribing them to days of the week.

A line showing average time was added to support comprehension and was felt to help.

5.2.5 Public transport

Average travel time was well understood, and reliability was tested using a similar presentation format to that used for cars.

Service frequency was an important attribute for public transport users, but interpretation varied, including interpretation as a proxy for reliability (eg, as the longest they might have to wait when buses were not running to timetable) or as an experienced waiting time at a bus stop rather than a scheduled frequency they could rely on. These misunderstandings were corrected in later testing and the final survey version (Figure 5.5).

Figure 5.5 Public transport (bus/train)

	Route one	Route two
Service frequency	Every 15 minutes	Every 10 minutes
Average Travel time	35 minutes	45 minutes
Cost (fare)	$3	$5
Crowding	You will have to stand for the whole trip	You will be able to sit for the whole trip

5.2.6 Other observations (non-game questions and survey overall)

5.2.6.1 Survey length

Test interview surveys ranged between 25 and 45 minutes with an average length of approximately 35 minutes, excluding wider discussions held to assess comprehension.

The cognitive burden of the final choice tasks is nearing the upper limit for participants. The games had already been successfully streamlined and explanations of the games have been reduced as far as practical while still effectively explaining all necessary information. Some unnecessary demographic data collection was dropped.

5.2.6.2 Number of games

The number of games and attributes presented (five games per section – safety and either car or public transport use (ie, 15 in total) with just one trip type and purpose per person) worked well. A significant amount of interview time is taken up explaining the attributes and games and countering common misconceptions.

5.2.6.3 Other observations

In early rounds of testing, attribute non-attendance (ANA) was an issue, with some participants opting to ignore certain attributes altogether. In general, other than the minor injury attribute for some respondents, no participants were observed to be ignoring any attributes in the final versions.

5.3 Study pilot

5.3.1 Overview

After the testing rounds, a larger survey pilot was launched. The sample consisted of 100 respondents who completed the survey face-to-face, in their own home. One respondent was selected from each dwelling, with dwellings being randomly selected from within primary sampling units.[23]

Three separate survey designs are used, with some variants:

1. a route choice survey for car drivers and passengers (entailing two CEs)
2. a public transport service survey applied to bus and trains users (entailing two CEs) (train users in some parts of New Zealand only)
3. a government programme choice survey focused on safety issues (one CE).

The sample frame for the pilot consisted of ten Stats NZ Statistical Area 1 (SA1) units, which were purposively selected from within Auckland, Christchurch, Lower Hutt, New Plymouth and Tauranga. The blocks were chosen to ensure the sample included households of varying socio-economic status, as well as providing an even split of urban and rural areas. Additionally, half of the SA1s selected had a New Zealand Deprivation Index (NZDep) score of 9 or 10 (most socio-economically deprived). The rationale for including more deprived areas in the pilot was to help to ensure people with lower literacy/education levels (factors which are highly correlated with deprivation) were included. Eighteen dwellings were selected in each SA1.

[23] Primary sampling units refer to sampling units that are selected in the first (primary) stage of a multi-stage sample. In the case of the New Zealand Travel Choices Survey (NZTCS), a primary sampling unit refers to a geographical area containing dwellings. Individual dwellings within the primary sampling unit are selected during the second stage of sampling.

Within the selected SA1s, a total of 180 households were chosen. Households were sent an invitation pack and information pamphlet, which informed the household that an interviewer would be visiting their area and that somebody in the household would be invited to take part in the survey. The invitation pack also included a flyer detailing the measures that CBG was taking to ensure the safety of respondents and interviewers in relation to COVID-19.

Contact attempts were recorded by the interviewer into a tablet computer. Once a respondent was selected and agreement to participate secured, the respondent was required to sign a consent form prior to the survey commencing.

5.3.2 Feedback capture

Interviewers approached each pilot interview with the following questions in mind:

- Do any questions require respondents to think too long or hard before responding?
- Do any questions seem to irritate, embarrass or confuse people?
- Do the selection responses capture respondents' answer choices?

During the survey, interviewers made a note of any seemingly problematic questions. Interviewers were trained to complete this task without disrupting the flow of the interview. At the end of the questionnaire administration, interviewers invited the respondent to identify any problem questions/sections that they had struggled to answer.

Following the exit section of the questionnaire, participants self-completed a respondent burden assessment. The assessment aimed to gauge the respondent's satisfaction with the survey across four metrics:

- length of survey
- number of questions
- complexity of the questions
- intrusiveness of the questions.

Respondents were also asked to identify any questions that confused, frustrated or caused them to regret taking part.

Following each survey, interviewers made notes on their experiences administering the survey tool. They also took part in focus groups during, and at the end of, the fieldwork to explore any issues arising from the pilot and to discuss their experiences.

Additionally, all respondents were re-contacted by a researcher to obtain further feedback and satisfaction ratings.

5.3.3 Interview duration

The mean interview duration in the pilot was 20 minutes and 22 seconds, with a range of 6 to 59 minutes. Eighty-eight per cent of interviews were completed in under 30 minutes.

5.3.4 Survey programming

The survey was programmed in the Askia platform (www.askia.com). No routing or logic errors were reported during the pilot.

Due to time constraints, the pilot survey was administered using an online web link. For the main study, an offline version of the survey was developed to eliminate the need for a live Internet connection in-field.

5.3.5 Conclusion

The findings of the field pilot confirmed that the survey can be completed by the vast majority of respondents, with varying degrees of support being required from the interviewer. No operational issues were encountered, with the computer-assisted personal interviewing (CAPI) software and other systems performing well.

6 The final survey

Based on the study objectives, the review of literature, the pre-testing and the pilot, in this section we summarise the components of the stated choice (SC) survey used in this study. The full survey used is included in Appendix A.

6.1 Components

Two separate survey designs were used (Table 6.1). Examples are shown below. Each respondent undertook three sets of choice experiment (CE) games, each of which involved five choices (ie, a total of 15 choices):

- either two public transport games (QP1-5 (Figure 6.1) and QP6-10 (Figure 6.2)) or two private transport games (QC1-5 (Figure 6.3) and QC6-10 (Figure 6.4))
- a safety game (Figure 6.5).

Table 6.1 Summary of SC games

Games	Modes	Attributes	Tasks
QP1-5	Public transport	Service frequency, average travel time, crowding, cost	5
QP6-10	Public transport	Average travel time, time reliability, cost	5
QC1-5	Car driver, passenger, motorcyclist	Average travel time, time in heavy traffic, cost	5
QC6-10	Car driver, passenger, motorcyclist	Average travel time, reliability, cost	5
QS1-5	All	Deaths, injuries, cost	5

Figure 6.1 Public transport survey design – frequency, time and crowding

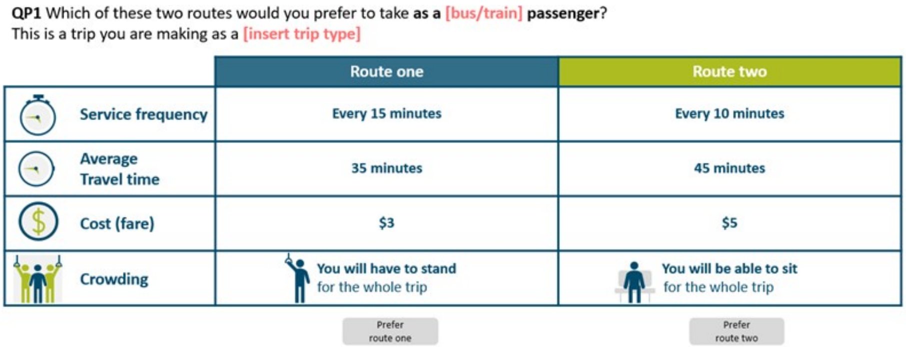


Figure 6.2 Public transport survey design – time reliability

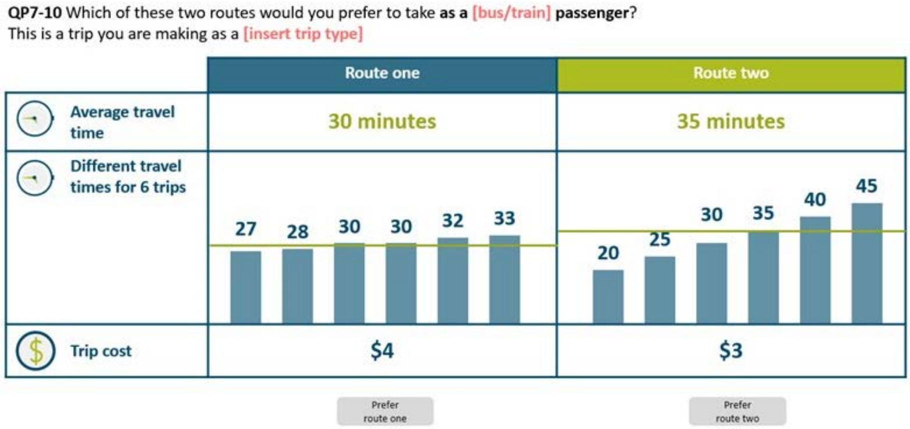


Figure 6.3 Car driver survey design – time/heavy traffic valuation

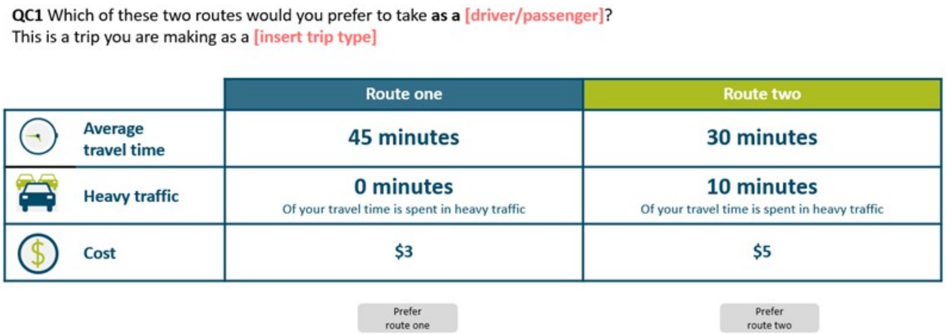


Figure 6.4 Car driver survey design – reliability

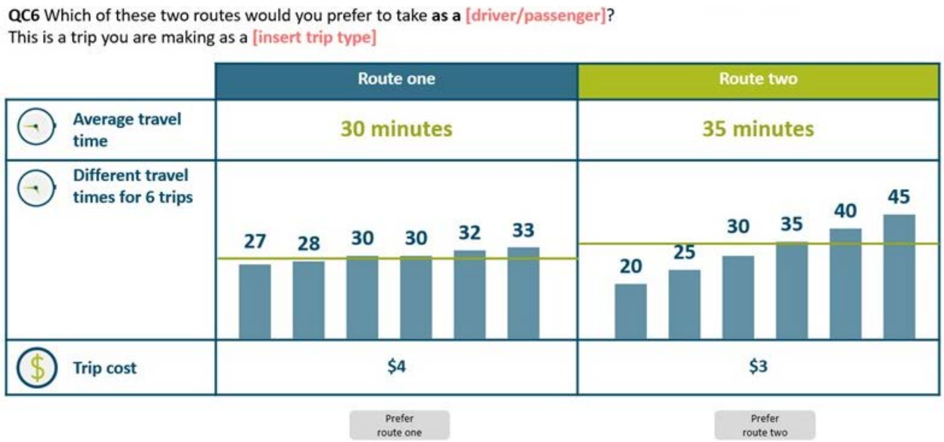
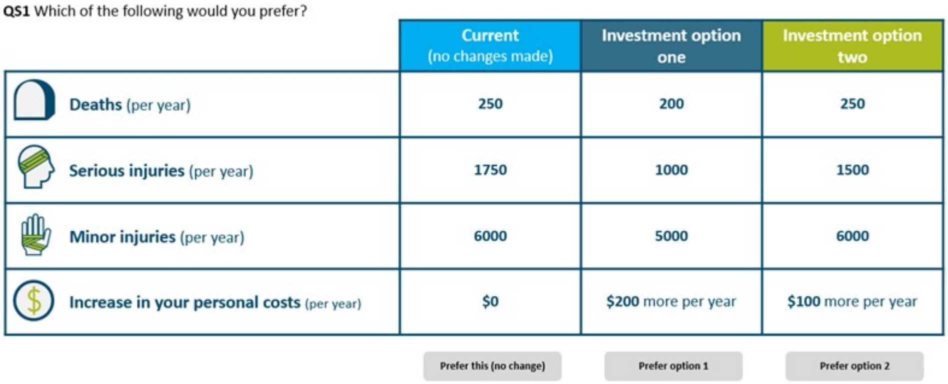


Figure 6.5 Safety survey design – deaths and injuries



QS1 Which of the following would you prefer?

	Current (no changes made)	Investment option one	Investment option two
Deaths (per year)	250	200	250
Serious injuries (per year)	1750	1000	1500
Minor injuries (per year)	6000	5000	6000
Increase in your personal costs (per year)	$0	$200 more per year	$100 more per year
	Prefer this (no change)	Prefer option 1	Prefer option 2

The public transport and private transport route choice surveys are applied to different trip purposes (Table 6.2):

- commute trips
- non-commute trips that are either
  - time-dependent or
  - flexible.

Table 6.2 Route choice survey options

	Private transport	Public transport
Commute trip	✓	✓
Time-dependent trip	✓	✓
Flexible trip	✓	✓

In addition, trips were of different duration, with three durations for private transport journeys and two for public transport.

6.2 Efficient design

There are different approaches to experimental design that estimate the number of choice tasks (and the number of respondents required) so parameters can be estimated with the greatest expected reliability (lower expected standard errors) (Bliemer & Rose, 2006), with the objective being to maximise the D- or A-efficiency (Kim & Haab, 2004; Rose & Bliemer, 2009; Scarpa & Rose, 2008), where:

- A-efficiency minimises the sum of variances of an individual parameter estimate
- D-efficiency minimises the mean variance around multiple estimated parameters.

For a defined sample size, these objectives can be used to develop a plan for:

- the number of choice tasks per respondent
- the number and distribution of attribute levels across the sample.

However, efficient design is not a panacea. For example, Walker et al. (2018) found that D-efficient designs perform very poorly when the priors are mis-specified, and that Bayesian designs' robustness was sensitive to assumed variance of priors. Their mode-choice research found that random designs were more efficient than D-efficient designs for estimating value of travel time (VTT) when assumptions about priors were significantly wrong. Efficient designs can sometimes generate dominated alternatives that do not provide useful information.

We mitigated these potential problems by drawing on the information gained from previously published VTT and value of risk reduction (VRR) studies and the pre-test and pilot studies to establish priors for this study. Designs containing dominated alternatives were identified and removed.

Optimal Bayesian multinomial logit (MNL) model designs were estimated in Ngene software, with all attributes (except the standing dummies in the public transport service scenarios) using 3–5 attribute levels, with the attribute-level ranges designed to accommodate the potential ranges of estimated willingness to pay (WTP), and using the priors identified in analysis of results from initial survey testing. In all cases, five-level attribute ranges were employed. Since Ngene uses what is essentially a trial-and-error approach, it sometimes produces final designs that contain dominated alternatives.

- Fully dominated alternatives contain one alternative that is at least as good as another for all attributes.
- Partial domination arises in the reliability models. While the design software produces designs consistent with model identification for travel time and standard deviation (SD) of travel time, it can produce alternatives for which all travel times (ranked) are shorter for one of the two alternatives. Where that occurs, if the alternative with the shorter-ranked times has the same or smaller cost, that variant will always be chosen, irrespective of larger SD. This is partial domination.

All designs were checked for dominated alternatives, and manual corrections were made to selected attribute levels to remove both forms of domination.

6.3 Pivot designs

In the survey design we aimed to enhance realism for participants by pivoting time-related attributes off participant-nominated levels associated with a relevant, recently completed journey. To enable real-time computer-assisted personal interviewing (CAPI) attribute-level specification, and following recent applications (Batley et al., 2019; González et al., 2018; Hensher, 2008; Hensher & Layton, 2010), we assigned time-related attributes to pre-specified ranges. Predefined attribute-level differences, created through an efficient experimental design process, combined with each participant's reference attribute levels produce an individual-specific experimental design. Safety questions were not pivoted.

In the route choice designs:

- Respondents were asked about the kinds of trips they make, and this was used to limit the choice games they are shown to:
  - a single trip purpose
  - modes – respondents were asked to complete a survey for either public or private transport modes
  - levels for the travel time attribute were pivoted around levels experienced by the respondent on a recent journey for the purpose in the choice scenario.
- Car drivers in the private transport survey were asked to assume they are driving alone.[24]

6.4 Attribute values

The attribute values for the choice scenarios are shown in the tables below.

[24] This is consistent with previous studies.

Table 6.3 Attribute values for public transport service questions

Attribute	Scenario A (short)	Scenario B (medium)
Travel time (minutes)	15, 17, 20, 23, 25	35, 37, 40, 43, 45
Frequency (minutes)	10, 13, 20, 25, 30	10, 13, 20, 25, 30
Standing full journey	0, 1	0, 1
Standing half journey	0, 1	0, 1
Cost ($)	3, 3.5, 4.5, 6, 8	5, 5.5, 6.5, 8, 10

Table 6.4 Attribute values for public transport reliability questions

Attribute	Scenario F (short)	Scenario G (medium)
Travel time (minutes)	15, 17, 20, 24, 25	35, 37, 40, 43, 45
SD (minutes)	1.4, 2.3, 3.9, 6.3, 9.6	2.3, 3.9, 6.3, 9.6, 11.0
Cost ($)	3, 3.5, 3.8, 4.5, 6	4.5, 4.8, 5.25, 6, 7

Table 6.5 Attribute values for private transport service questions

Attribute	Scenario C (short)	Scenario D (medium)	Scenario E (long)
Travel time (minutes)	10, 11, 13, 16, 20	35, 38, 45, 50, 55	100, 105, 115, 130, 160
Heavy traffic (minutes)	0 ,3, 6, 8, 10	0, 3, 8, 15, 20	0, 5, 15, 25, 40
Driver cost ($)	3, 4, 6, 9, 11	6, 8, 10, 12, 14	12, 15, 18, 21, 25
Passenger cost ($)	Driver cost minus $2	Driver cost minus $3	Driver cost minus $5

Table 6.6 Attribute values for private transport reliability questions

Attribute	Scenario H (short)	Scenario I (medium)	Scenario J (long)
Travel time (minutes)	12, 15, 20, 22, 25	35, 38, 45, 50, 55	100, 105, 115, 130, 140
SD (minutes)	0, 1.4, 3.2, 4.9, 6.3	2.3, 3.9, 6.3, 9.6, 15.3	2.3, 3.9, 6.3, 11.0, 20.5
Driver cost ($)	3, 4, 6, 8, 10	6, 8, 10, 12, 14	12, 15, 18, 21, 25
Passenger cost ($)	Driver cost minus $2	Driver cost minus $3	Driver cost minus $5

Table 6.7 Attribute values for safety questions

Attribute	Status quo attribute values	Policy attribute values
Deaths	250	160, 170, 200, 225, 250
Serious injuries	1,750	900, 1,100, 1,400, 1,675, 1,750
Minor injuries	6,000	1,200, 1,500, 2,000, 4,000, 6,000
Cost ($)	Zero	50, 100, 250, 400, 500

A two-stage design process was used. After approximately 800 responses were received, initial data were analysed to assess likely WTP magnitudes, enabling validation of the priors used in the experimental design.

At this point new experimental designs were created to improve efficiency, using the updated priors. (These are the design levels in Tables 6.3 to 6.7. Some initial designs had 3-level attributes).

6.5 Survey implementation

6.5.1 Survey branding and collateral

The survey was undertaken by CBG's public sector surveying team. The study was branded to the general public as the New Zealand Travel Choices Survey (NZTCS). A range of materials were developed to support the survey in-field, including an invitation pack that was mailed to households. The pack included an information brochure and a separate flyer detailing the measures that were being taken to keep respondents and interviewers safe in the context of COVID-19.

6.5.2 Sample selection

The NZTCS employed a multi-stage, stratified, probability-proportional-to-size sampling design. The survey was designed to yield 7,000–8,000 respondents.

The 2018 New Zealand Census Statistical Area 1 (SA1) data were used as the area-based sampling frame and were treated as primary sampling units. SA1s are aggregations of meshblocks, optimised to be of similar population size. As SA1s combine one, two, or more meshblocks, there is less variation in weights than using meshblocks and a minimal reduction in variance of weighted data. SA1s have an ideal size range of 100–200 residents, and a maximum population of about 500. A sample of 750 SA1s was selected from this frame using probability-proportional-to-size sampling.

A list of households was compiled for each selected SA1 using data from the New Zealand Post Postal Address File. A systematic sample of approximately 15 households was then selected from this list by choosing a random start point and selecting every k^{th} household. The skip k was calculated by the 2018 Census occupied-dwellings count divided by 15.

Finally, one adult (aged 18 years or over) was selected at random from each selected household.

6.5.3 Survey approach

6.5.3.1 Interview mode

Interviews were predominantly conducted in respondents' homes, although the interview could be completed at another location at the request of the respondent (eg, their workplace).

Data were collected using a combination of CAPI and computer-assisted self-interviewing (CASI). For the CAPI questions, interviewers recorded responses on behalf of the respondent. For the CASI portions of the interview, respondents entered their own responses. All data were captured directly via a laptop using the Askia survey platform.

'Showcards' with predetermined response options were used to help respondents where appropriate. A separate tablet computer was used to display the showcards, with the applicable card being automatically displayed as the survey progressed.

6.5.3.2 Informed consent

The NZTCS was voluntary. Selected households were mailed an invitation pack prior to the interviewer's first visit. Following successful contact with a household, the interviewer entered all usual residents aged 18 and over into a computer program that randomly selected one occupant to be invited to take part. Respondents selected for the survey were presented with a copy of these documents as part of the informed consent process. Participants were asked to sign an electronic consent form and were given a copy of the consent form to keep. The consent form required the respondent to confirm they had read and understood the information pamphlet, that they could ask questions at any time and that they could contact CBG or Waka Kotahi for more information. The consent form also informed respondents:

- of their right to request an interpreter if required (in a range of 10 different languages)
- that they could stop the interview at any time
- that they did not have to answer every question
- that their participation was confidential, and no identifiable information would be used in any reports
- that their answers were protected by the Privacy Act 2020.

The option was available to match respondents and interviewers by ethnicity and/or gender, although this was rarely requested.

6.5.3.3 Interviewer training

Forty-six interviewers were trained to deliver the survey in-field. Interviewers were trained over a two-week period, which consisted of remote learning and face-to-face in-field assessment. Training covered both sampling procedures and questionnaire administration. Practice interviews were conducted by each interviewer as part of this training. Online training modules were developed, which contained both generic CBG training material as well as material specific to the administration of the NZTCS.

6.5.3.4 COVID-19 protocols

In response to the COVID-19 pandemic, a number of protocols were developed by CBG to ensure the safety of households selected to participate, as well as the field interviewers. A flyer detailing these measures was included in the invitation pack mailed to each selected address. These measures included:

- interviewer training on infection control
- physical distancing
- cleaning and sanitising of equipment and hands
- household and interviewer wellbeing checks
- record keeping.

Once contact had been made with a household, the interviewer ensured that a distance of at least one metre was maintained on the doorstep during the respondent selection process. In those houses where a respondent was selected, an additional COVID-19 screener was administered to identify if anyone in the household was at increased risk of COVID-19. The screener asked the door-opener:

- Is anyone in your household currently unwell and have symptoms similar to COVID-19? This includes fever, coughing, sore throat and sneezing.
- Is anyone in your household self-isolating? For example, because they have travelled back from overseas recently or have been in contact with someone who has had COVID-19?
- Is anyone in your household currently employed in a role where they may come in contact with COVID-19? For example, working at official quarantine facilities, or employed to work on aircrafts that come from overseas?

If the door-opener answered any of these questions affirmatively, then a face-to-face interview was not permitted. In this situation, the respondent had the option of completing the survey at a later date.

A household outcome code was available to interviewers in order to capture COVID-related non-response.

6.5.4 Survey

6.5.4.1 Call pattern

A ‘call’ refers to one visit on one day during a particular time period. Up to 10 calls to each sampled dwelling were made at different times of the day and on different days of the week, before accepting that a dwelling was a non-contact. Calls were recorded as unique events only if they were made at least two hours apart. Calls were spread out over the duration of the fieldwork. Six calls were made in the survey month in which the work allocation was started.[25] If no contact had been achieved by this point, there was a pause with no attempted contact with the dwelling for one to two weeks, before attempting four more calls.

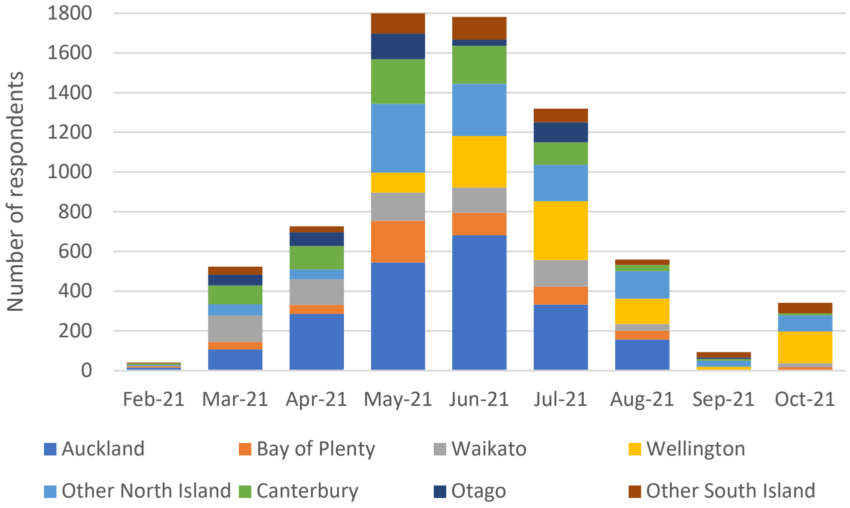
Households where no contact had been established, or where the selected respondent was unable to take part at that time but did not refuse to participate, were revisited during a mop-up phase in an effort to secure participation.

6.5.4.2 Fieldwork

Interviews for the main study were conducted between 24 February and 31 October 2021.

The number of interviews by date and location is shown in Figure 6.6, with more detail in Table 6.8.

Figure 6.6 Number of interviews by date and location



[25] SA1 was used as the primary sampling unit.

Table 6.8 Number of interviews by date and region

Region	Feb-21	Mar-21	Apr-21	May-21	Jun-21	Jul-21	Aug-21	Sep-21	Oct-21	Total
Auckland	15	107	286	545	682	333	156	0	0	**2,124**
Bay of Plenty	8	37	45	208	112	90	46	0	17	**563**
Canterbury	8	94	117	224	191	113	32	8	11	**798**
Gisborne	0	0	0	61	4	0	0	0	22	**87**
Hawke's Bay	0	30	14	33	93	48	4	0	4	**226**
Manawatū-Whanganui	0	0	0	14	76	108	129	27	11	**365**
Marlborough	0	0	0	0	32	31	0	0	9	**72**
Nelson	0	0	0	0	34	1	2	0	2	**39**
Northland	0	0	22	191	53	15	6	4	0	**291**
Otago	0	55	69	129	32	101	0	8	0	**394**
Southland	0	22	25	103	25	27	0	0	23	**225**
Taranaki	5	26	14	47	37	13	0	0	44	**186**
Tasman	0	0	0	0	20	10	25	27	19	**101**
Waikato	0	134	129	144	129	133	32	0	20	**721**
Wellington	0	0	0	101	258	297	128	19	160	**963**
West Coast	4	19	6	0	4	0	0	0	0	**33**
Total	**40**	**524**	**727**	**1,800**	**1,782**	**1,320**	**560**	**93**	**342**	**7,188**

In general, interviewing took place at Alert Levels 1 and 2, with the above COVID-19 precautions in place. However, as a result of the August 2021 Delta outbreak, interviewing was suspended nationwide from 18 August until 16 September. From 17 September, interviewing resumed in areas outside of Auckland and Waikato. From 8 October interviewing was re-suspended in Northland.

6.5.4.3 Sample composition

Table 6.9 shows the composition of the achieved sample by ethnicity (total response). The number of responses is greater than the number of respondents (and total percentages greater than 100%), as the selection of multiple ethnic groups was possible.

Table 6.9 Sample composition

	Male	Female	Gender diverse	Not reported	Total	%	2018 Census
European	2,356	2,763	5	0	5,124	71.3	70.2
Māori	345	486	1	0	832	11.6	16.5
Pacific peoples	160	226	0	0	386	5.4	8.1
Asian	385	411	1	0	797	11.1	15.1
Other (not specified)	237	264	0	0	501	7.0	1.5
Not reported	17	29	0	131	177	2.5	0
Total	**3,500**	**4,179**	**7**	**131**	**7,817**	**108.8**	**111.4**

The number of Māori and Pasifika respondents included in the survey was lower than expected, due to the COVID-related disruptions to fieldwork. In particular, a significant number of cases in Auckland were left unworked or in progress at the time of the August 2021 lockdown, which were never resumed.

6.5.4.4 Interview duration

The mean duration of the survey was 18 minutes. Time taken for the interviewer to engage with the household, to complete the consent process, and to pack away at the end of the survey (an average of 10 minutes) is not included in this timing.

6.5.4.5 Quality control

Interviewers were monitored by CBG management by:

- in-field assessment to ensure survey protocols were being followed correctly
- examination of individual performance metrics and exploration of strategies to improve these if necessary
- checking of a random selection of completed interviews by phoning respondents to confirm the interview was completed according to survey protocols and to collect satisfaction ratings.

Participants were also left with feedback postcards that they could use to send feedback directly to CBG, anonymously if they chose. In addition, CBG operated a toll-free survey helpline that participants could call if they had any questions about the survey or wanted to provide feedback. The results of these quality checks were communicated to the individual interviewers on a regular basis throughout the fieldwork period, with additional training and mentoring provided where required.

6.5.5 Data processing and formatting

Each interviewer was required to upload encrypted survey data to CBG servers every day they were active in the field. The files consisted of all changes that had been made to the Sample Manager and Askia survey databases residing on the interviewer’s laptop since the last upload. For example, this could include new survey data, information on contact attempts or new household outcome coding.

Once received at CBG, the files were decrypted and checked before being processed into a SAS data warehouse. A number of datasets resided within the warehouse pertaining to survey data collected and other survey metrics recorded by the interviewer (eg, respondent information and outcome coding).

The contents of each export file were analysed and directed to the relevant datasets ready for further formatting and cleaning.

Questionnaire responses arrived from the field as raw survey files. Formatting of these raw data was performed to ensure that the supplied datasets were consistent with the questionnaire document. The following tasks were undertaken during the formatting stage.

- Any partial interviews were removed.
- Variables were renamed to match the question numbers used in the questionnaire document.
- Unwanted variables were removed. These were usually ‘dummy’ variables that were included in the survey to achieve desired functionality and behaviour required (eg, complex skip logic and consistency checks).
- Range checks were undertaken to ensure that the values for each variable conformed to the questionnaire document.

At the end of the survey, interviewers were required to code whether the respondent displayed comprehension issues with any of the sections of the survey, and if so, provide detail on the issues encountered. They were also asked to record if the respondent was able to successfully complete the choice games. As part of the dataset creation stage, the responses to these questions were analysed and some interviews were subsequently excluded from the final dataset, where it was evident that the accuracy of the data collected was compromised (see next chapter for details).

7 Results

7.1 Overview

This chapter describes the survey results and the analysis of the results to produce values for the following attributes:

- travel time
- reliability of travel time
- reduction in risk of fatalities and injuries.

The time and reliability values were differentiated by purpose (commuting, time-dependent, flexible), mode (private and public transport), by congestion level (free moving or heavy traffic), and by trip length. All respondents completed a survey on the value of reducing road deaths and injuries.

The survey work was undertaken over the period February to October 2021 and was significantly affected by COVID-19 restrictions, including local lockdowns and the need for surveyors to reach full vaccination status. Analysis of the results started at the end of 2021 and was completed in early 2022.

Below we set out the approach taken to analysing the data, including data cleaning and the statistical models used. The results are presented, with comparisons made with current values and those in other countries.

7.2 Data

7.2.1 Data cleaning

Surveying was stopped prior to (but close to) attainment of a fully representative cross-section because of ongoing COVID-19 restrictions that meant surveying was not possible in some key areas, including parts of Auckland where populations of some ethnic groups were concentrated. Representativeness was obtained *ex post* by survey company assignment of weights to individuals. The achieved sample was post-stratified to the New Zealand population using Census 2018 data, by age (16–24, 25–64, 65+), gender, and ethnicity (Māori, Pasifika, other). Weighted responses were used in all statistical analyses.

The collected data (N = 7,188) were vetted for coding errors, which were corrected where the nature of the error was evident, or the respondent was removed from the sample. Each respondent who used public transport was assigned to the public transport group (N = 902), with the remainder assigned to the private transport group (N = 5,963). All respondents participated in the safety questions (N = 6,865).

Interviewers provided notes on each interview, evaluating the individual participant's level of understanding, engagement, deliberation, and ability to respond to the choice event. Interviewer comments were used to identify problematic cases. For example, in some cases the interviewer noted that the individual did not want to answer the choice questions, so the interviewer chose to record the status quo option for each choice event for that case. While these responses appear complete, they are, in effect, non-responses. Interviewer comments were available for each set of choice events (eg, there were separate comments for the five public transport 'service' choice questions, for the five public transport 'time reliability' choice questions, and for the five safety questions). Interviewer comments enabled choices about exclusion during the modelling phase, where a conservative approach was adopted. The interviewer comments were supported by auxiliary questions that allowed subjective assessment of conformity with the theoretical choice scenario and protest responses (Johnston et al., 2017).

7.2.2 Data summary

The responses rejected during the data cleaning process are summarised in Table 7.1.

Table 7.1 Data summary

	Public transport	Private transport	Safety
Number of respondents	902	5,963	6,865
Number rejected because of non-comprehension or refusal	0	16 Service 8 Reliability	40
Number rejected because of protest responses	0	4 Service 1 Reliability	1,026

Public transport users had no conceptual difficulties with the public transport choices proposed to them, and everyone provided usable responses to the questions. This is unsurprising, given that public transport users routinely make choices over the attributes in these scenarios.

For private transport users, a small number of people either refused to answer or the interviewer evaluated that they had significant comprehension issues that resulted in their exclusion from the analysis.

The safety questions entailed public rather than private choices, had three alternatives rather than the two in the public and private transport service and reliability scenarios, were of a more hypothetical nature, and also posed a moral challenge in requiring people to consider trade-offs between effects on themselves and other members of the community. This set of questions posed significantly more participant burden and introduced reasons that people might reject the stated choice (SC) scenario. As a result, 40 people were excluded from the analysis because they refused to answer, they stated they were unable to comprehend the task, or the interviewer evaluated that they had significant comprehension issues (Table 7.2).

Table 7.2 Rejected safety question responses

Reason for rejection	Number of respondents rejected for this response
Outright rejection (No opinion, don’t understand, don’t want to answer)	40
I already pay for this, taxes are already too high, etc.	270
Paying more taxes won’t overcome driver behaviour, young drivers, alcohol impaired drivers, etc.	245
Paying higher taxes won’t improve safety (without a specific reason)	152
Spend existing transport funding better. Road safety can be improved within existing budgets	106
Offered an alternative solution to address road safety	60
Not my cost/responsibility to pay for improved road safety	54
Needed more information (typically on the type of intervention) to decide	47
Safety is an individual responsibility	35
Don’t trust Waka Kotahi, the government, etc.	31
Funding for safety improvements should come from elsewhere	26

Of the remaining 6,825 safety question respondents, the vast majority (5,799 people; 85%) provided no indications to interviewers or through auxiliary question responses of difficulties with the safety questions or of protest responses. Responses that were interpreted as protests to the safety questions were received from 1,026 participants who were excluded from analysis, including ‘non-traders’ who did not believe the government could do anything to reduce road deaths or injuries, resulting in a total of 1,066 exclusions.

7.3 Statistical modelling

Data analysis requires several key decisions, including:

- whether and how to model respondent heterogeneity – this entails several sub-questions:
  - which statistical model to use (eg, multinomial logit (MNL), random parameters logit (RPL), generalised mixed logit (GMXL), willingness to pay space (WTP-space) models)
  - which variables (if any) to model as random parameters
  - which statistical distributions to use for random parameters
- how monetary value estimates are derived from the statistical model results
- whether it is necessary or desirable to model respondent sub-groups and how that should be done.

7.3.1 Heterogeneity

7.3.1.1 Statistical model selection

Statistical model selection is a trade-off between parsimony and complexity but should account for respondent heterogeneity (Johnston et al., 2017). The multinomial logit (MNL) model is the basic workhorse of discrete choice models.[26] While the MNL model is very simple and quick to apply, it does not exploit the panel data available in this study. The MNL model implicitly assumes every individual has identical preferences (unless modelled as covariate effects) and imposes restrictive substitution assumptions. The contemporary approach to addressing these matters is to use random parameters models. Random parameters logit (RPL) models, also known as mixed logit models, allow for respondent heterogeneity by recognising the patterns within each individual’s responses. Generalised mixed logit (GMXL) models are more advanced forms of RPL models that relax assumptions about scale, which is a measure of the consistency of responses – the varying standard deviation (SD) of the utility function errors across consumers (Hossain et al., 2018). These models are nested – the RPL applies when specific restrictions are applied to the GMXL, and the MNL applies when specific restrictions are applied to the RPL. Hence, using the more flexible models from higher up the hierarchy (ie, GMXL) avoids imposition of assumptions (which may or may not be appropriate) while allowing tests of their implications. A cost of using these more complex models is the number of coefficients that must be estimated, and higher model run times. Model non-convergence is more common with increased complexity.

Discrete choice models can be used to estimate willingness to pay (WTP) in two different ways. In preference-space models, parameters are measured in units of ‘utility’. The coefficients in these models are measures of marginal utility, and WTP is computed by dividing the attribute parameters by the cost parameter. More recently, willingness to pay space (WTP-space) models have been developed in which parameters are units of WTP, rather than utility. Preference-space choice models (eg, MNL, RPL and GMXL) include the cost coefficient in the utility function, whereas WTP-space models do not.

In a preference-space model with two transport-related attributes (X_1 and X_2) and a cost attribute (C), the simplified utility function can be characterised as:

[26] See Appendix B for details of different model types.

$$U = \alpha_0 + \beta_1 X_1 + \beta_2 X_2 + \delta C \qquad \text{(Equation 7.1)}$$

The cost coefficient is δ. The monetary value of a marginal change in X_i is $WTP(X_i) = -\beta_i/\delta$. While statistical models provide point estimates of the means of the utility function coefficients, these estimates are not certain, and there is an associated distribution of each estimated coefficient. Because the distribution of $WTP(X_i)$ is a function of two estimated coefficient distributions, the distribution of $WTP(X_i)$ must be simulated from the appropriately correlated distributions of the two estimated coefficients, or its variance approximated by other methods, such as the Delta method.

WTP-space models estimate (again in simplified terms) the following function by imposing the constraint that the marginal utility of cost is equal to negative one, in effect dividing through by $-\delta$.

$$W = \alpha_0/-\delta + \beta 1/-\delta X1 + \beta_2/-\delta X_2 + \delta/-\delta C = \alpha_1 + WTP_1 X_1 + WTP_2 X_2 - C \qquad \text{(Equation 7.2)}$$

The WTP-space model is extremely convenient because the distributions of the monetary value estimates are produced directly, avoiding the need for simulation or approximation.

7.3.2 Application

The modelling strategy was to invoke the fewest possible restrictions by applying GMXL models wherever possible. MNL models, the results of which we do not report, provided starting values for estimation of more complex models that incorporated aggregate respondent heterogeneity through random parameters, and correlations in individuals' responses to each set of five choice questions. For the service and travel time choices, GMXL models were estimated in WTP-space using NLOGIT software[27] – providing maximum model flexibility.

Safety choice models were applied to the population as a whole. The absence of sub-population models vastly reduced the number of models estimated and provided more opportunity to explore the impacts of model choice. For the safety choices, RPL, preference-space GMXL, and WTP-space GMXL models were all applied, enabling verification of sensitivity of monetary value estimates to model type.

7.3.2.1 Variables modelled as random parameters

Recognised problems with WTP estimation when the cost parameter is random include very wide confidence intervals, positive cost coefficients, and non-defined moments (Bliemer & Rose, 2013, Daly et al., 2012; Hensher et al., 2015). While adoption of a lognormal cost distribution could have resulted in cost coefficient estimates constrained to be negative, 'using lognormals, due to the presence of heavy tails, may produce biased results' (Bliemer & Rose, 2013, p. 200). Daly et al. (2012, p. 19) go further and show 'some popular distributions used for the cost coefficients in random coefficient models, including normal, truncated normal, uniform and triangular distributions, imply infinite moments for the distribution of WTP, even if truncated or bounded at zero'. Consequently, cost parameters in models that were not estimated in WTP-space (which have fixed cost coefficients) were non-random – this applied to only five safety models.

All safety and transport-related coefficients were random, reflecting the expectation that different individuals have different absolute and relative values for these attributes. Alternative-specific constants (ASCs) were also random. However, in a limited number of cases, models that included a random ASC failed to converge, in which case they were estimated with a non-random ASC.

[27] https://www.limdep.com/products/nlogit/

7.3.2.2 Random parameter distributions

Significant ASCs are indicators that respondents are reacting to something other than the attribute levels presented to them in the choice alternatives. For example, a positive ASC can signal ‘stickiness’ around the status quo because people are averse to change. Alternatively, a negative ASC may be an indicator that the individual prefers to be doing something (anything) different to the status quo. This is the case with ‘warm-glow’ effects, where people want to signal the importance of improved outcomes, regardless of the cost (Andreoni, 1990; Haghani et al., 2021). ASCs can also signal order effects – for example, a propensity to always choose (say) the first alternative. Stickiness, warm-glow, and order effects may differ across respondents, so there is no *a priori* expectation of the sign of the ASC. Consequently, normally distributed random ASCs that do not constrain the sign on the ASC, are included, where possible, in all models. In a few cases model convergence required a non-random ASC.

The implausibility of negative WTP (for adverse outcome avoidance) was eliminated in all random parameters models by constraining the random distributions of attribute coefficients to be positive. Previous studies have used various random parameter distributions, including normal (Iragüen & Ortúzar, 2004; Parumog et al., 2006; Veisten et al., 2013), constrained triangular (Hensher, 2008; Hensher & Layton, 2010; Hensher et al., 2009, 2011; Niroomand & Jenkins, 2018), and lognormal (Niroomand & Jenkins, 2016; Parumog et al., 2006, Rouwendal et al., 2010).

7.3.3 Application

Our tests of the normal distribution produced significant numbers of negative WTP estimates, so it was rejected. One-sided distributions[28] evaluated across all the choice experiments (CEs) were lognormal, one-sided triangular, modified Rayleigh, and the maximum distribution, which censors the normal distribution at zero.

Based on convergence criteria (for example, lognormal and maximum models frequently failed to converge) and statistical fit criteria, the one-sided triangular distribution performed best overall and is reported here for all choice tasks. The one-sided triangular distribution models always converged and nearly always had similar or better McFadden’s R^2, Akaike information criterion (AIC) and Bayesian information criterion (BIC) scores than the other distributions. A significant benefit of the one-sided triangular distribution is that, by definition, the spread of the distribution is equal to the mean, reducing the number of estimated coefficients with this distribution relative to other distributions that require estimates of both mean and spread. Each model used 100 Halton draws.

We assessed five different distributions for the safety-related random parameters in the safety models – the one-sided triangular, normal, lognormal, maximum, and Rayleigh distributions. Several of these safety models failed to converge. These included GMXL maximum and lognormal distributions, and WTP-space

[28] Random parameter distributions:

Lognormal	$\beta_i = \exp(\beta + \sigma v_i)$,	$v_i \sim N[0,1]$
One-sided triangular	$\beta_i = \beta + \beta v_i$,	$v_i \sim$ triangular $[-1,1]$
Modified Rayleigh	$\beta_i = \exp(\beta + \sigma v_i)$,	$v_i = 2(-\log u_i)^{\surd .5}$, $u_i \sim N[0,1]$
Maximum	$\beta_i = \mathrm{Max}(0, \beta + \sigma v_i)$,	$v_i \sim N[0,1]$

All of these distributions admit only non-negative values. The lognormal and modified Rayleigh distributions have long right-hand tails, which means they can produce some very large parameter WTP estimates relative to the mean. The maximum value in the one-sided triangular distribution is twice the mean. The maximum distribution is the familiar normal distribution, but with all negative values assigned the value of zero. This can produce a peak of values at zero. The right-hand tail of the maximum distribution is not skewed, as the tails of the lognormal and modified Rayleigh distributions are.

normal, lognormal, and maximum distributions. The GMXL and WTP-space Rayleigh distributions converged, but had poor fit, so were discarded.

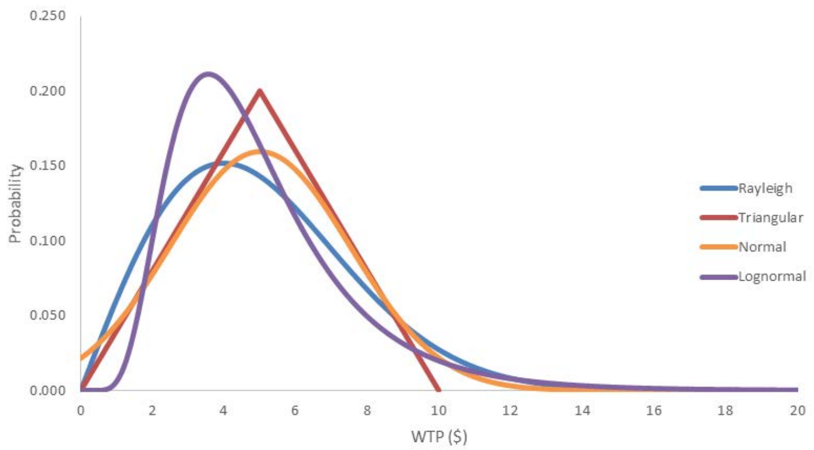
The RPL safety models all converged. Like its GMXL counterpart, the RPL normal distribution resulted in WTP estimates that straddled zero by considerable margins. Hence, the normal distribution was discarded. The maximum distribution RPL model had marginally better fit than the modified Rayleigh distribution model. The lognormal distribution was the worst fitting RPL model.

We report the following safety models:

- RPL: One-sided triangular, lognormal, Rayleigh, maximum
- GMXL: One-sided triangular
- WTP-space: One-sided triangular.

Figure 7.1 illustrates the differences between these distributions for mean = 5 and SD ≈ 2.5. The normal distribution has negative values. The maximum distribution has a spike at zero that accounts for the negative section of the normal distribution. For the same mean the maximum distribution mode is to the left of the normal distribution mode.

Figure 7.1 WTP probability density functions (mean = 5)



7.3.4 Monetary value estimates

WTP-space models without covariate-related heterogeneity greatly simplify estimation of money values. For these models, the mean marginal value of each attribute is estimated directly as a parameter coefficient. Since WTP-space models also estimate confidence intervals for these coefficients, the confidence interval for the mean of the value distribution is also observable in the model results. An important caveat applies here though. Where heterogeneity is modelled with distributions that entail parameter transformations, such as the lognormal distribution, the coefficient estimates are not monetary values, but are relevant transformations of monetary values. These estimates require subsequent calculation to retrieve the monetary value estimates and confidence interval bounds.

Where there is covariate-related heterogeneity in the random parameters of the WTP-space model, estimation of monetary values requires simulation that accounts for the differences in means of the various sub-groups while accounting for correlations in the estimated model parameters. Typically, simulation is undertaken by some form of Monte Carlo analysis that creates a synthetic distribution of monetary values by randomly drawing appropriately correlated coefficients from their distributions and combining those algebraically.

Preference-space models produce estimates of the distributions of marginal utility for the transport attributes and for the cost attribute. Monetary value is the ratio of the transport attribute coefficient (drawn from the appropriate distribution and adjusted appropriately for covariate-related heterogeneity) and the cost attribute (drawn from the appropriate distribution). Hence, estimation of monetary values from preference-space models always requires some form of simulation. Again, this is typically achieved through some form of Monte Carlo estimation.

All transport models have modelled the distribution of transport-related coefficients using the constrained triangular distribution in WTP-space. Hence, mean monetary value estimates and their 95% confidence intervals are reported directly in the model results. We report six safety models, one WTP-space model, one GMXL model and one RPL model – all using the constrained triangular distribution. The remaining models are RPL models using the maximum, Rayleigh and lognormal distributions for the random parameters. The WTP-space model allows direct observation of monetary values. For the GMXL and RPL models, the distribution of monetary value estimates has been derived from 10,000 random draws in the Krinsky and Robb procedure, a Monte Carlo approach incorporating variance and covariance in the estimated model parameters that does not impose the symmetricity of the simpler Delta method (Bliemer & Rose, 2013). The exception is the RPL (maximum) model, for which the simulated distribution of monetary values is the distribution of the 5,799 estimated individual respondent means.

7.3.5 Sub-groups

The *Monetised Benefits and Costs Manual* (MBCM) time-related parameters are differentiated by journey purpose, but not by journey duration. However, one objective of the research was to inform assessment of the potential significance of differentiation by journey duration. Hence, private and public transport service and reliability value estimates are required for each type of journey (Commuting, Time-Dependent, Flexible) as a whole, and for the different journey duration groups (Short, Medium, Long) within each journey type. There was no need to model sub-groups for social value of safety choices because these were independent of journey purpose and duration.

Because of the large number of combinations of journey purposes and durations, it was impractical to address these through dummy attribute interactions in utility functions. Consequently, private and public transport choices were modelled for sub-groups. Two approaches were investigated.

1. For each journey purpose the sample was partitioned by journey duration, and separate models were estimated for each purpose/duration sub-group.
2. For each journey purpose a single model was estimated that modelled heterogeneity in time-related attributes as a function of journey duration.

The advantage of the first approach is simplicity and specificity – results for each group are not influenced by other unknown factors that differ between groups. This approach makes estimation of monetary values for each group straightforward, avoiding the need for complex simulation. The advantage of the second approach is potentially improved ability to test for statistical significance of inter-group differences. The disadvantage is the need for simulation of monetary values. Statistical fit of models using the first approach was better than for the second approach which, coupled with ease of monetary value estimation, resulted in adoption of the purpose/duration sub-group approach. In addition, models were fitted to the whole sample, without differentiation by either journey purpose or journey duration.

Over the two choice events for each transport sector, this resulted in a total of 46 models, 20 for public transport and 26 for private transport. The difference arises because there were only two journey durations in the public transport choices, whereas there were three journey durations in the private transport choices.

7.3.6 Model fit

Preferred models were identified based on convergence (some models would not converge – particularly GMXL models) and statistical measures of fit (Table 7.3). Goodness-of-fit was assessed using McFadden's Pseudo R^2 (which is not analogous to R^2 in a linear regression) and two normalised information criteria that adjust in various measures for sample size and number of parameters estimated in the model. Because the information criteria produce extremely similar results, we report only the more punitive BIC. For the safety models, which all have identical samples, the log-likelihood score provides an alternative, but equivalent, statistic to McFadden's pseudo R^2 since the restricted log likelihood is the same for all these models. Because sample sizes vary, it is not meaningful to directly compare log-likelihood scores for the non-safety models.

Table 7.3 Statistics

Statistic	Description	Interpretation
LL	Log likelihood of the fitted model	This is a negative number. Maximum likelihood estimation fits a model that makes this number as large as possible. In other words – as close as possible to zero.
LL_0	Log likelihood of a constants-only model	This is the log likelihood of a model without any parameters apart from a constant.
N	Sample size	The number of survey responses.
k	Number of parameter	The number of independent parameters estimated in the statistical model.
Pseudo R^2	McFadden's pseudo $R^2 = 1-LL/LL_0$	A measure of model fit. This is a measure of the proportion change in LL of the fitted model compared to the constants-only model. This measure, sometimes labelled Rho^2 to distinguish it, is different from linear R^2. Rho^2 scores are lower than equivalent linear R^2 scores. Approximate equivalencies to linear R^2 (Hensher et al., 2005) are: • pseudo $R^2 = 0.1$ linear $R^2 < 0.3$ • peudo $R^2 = 0.2$ linear $R^2 \sim 0.5$ • pseudo $R^2 = 0.3$ linear $R^2 > 0.6$.
AIC	Akaike information criterion ($-LL$*2 + 2k)	A measure of model fit. A lower value is better. AIC does not compare fit to a base model. AIC is a score based on LL adjusted for the number of estimated parameters.
BIC	Bayesian information criterion ($-LL$*2 + k*ln(N))	A measure of model fit. A lower value is better. BIC does not compare fit to a base model. BIC is a score based on LL adjusted for the number of estimated parameters and the sample size. BIC is more punitive than AIC.

7.4 Results

There are five sets of models (Public Transport Service, Public Transport Reliability, Private Transport Service, Private Transport Reliability, Safety) summarised in separate sections below. The statistical models used are discussed above and summarised in Table 7.4.

Table 7.4 Statistical model used

Model	WTP estimation	Distribution	Application
RPL	Preference-space	One-sided triangular	Safety
	Preference-space	Lognormal	Safety
	Preference-space	Rayleigh	Safety
	Preference-space	Maximum	Safety
GMXL	Preference-space	One-sided triangular	Safety
	WTP-space	One-sided triangular	Time, Service frequency, Congestion, Reliability, Safety

The results are presented in tables below, including estimated parameter values (as WTP) and statistical tests. A guide to their interpretation is given in Table 7.3.

7.5 Public transport

7.5.1 Value of travel time

For each of the three journey purposes there are three models, one of which (All) uses data for both journey durations, imposing the restriction that time values are independent of duration. In addition, the final three models do not distinguish by journey purpose, imposing the restriction that time values are independent of journey purpose. Comparison of estimated money values across these various purpose/duration combinations can help inform decisions about the level of detail required in project evaluations.

The results for the value of travel time (VTT) when those changes are certain are shown in Table 7.5. For non-commuting trips, results are provided for time-dependent trips (eg, dropping children at school) and flexible trips (eg, shopping). In some cases, the sample size is relatively small. For example, flexible-purpose, long-duration journeys had only 345 choice events involving 69 different people. Relatively small sample sizes and substantial respondent heterogeneity result in relatively low pseudo R^2 scores for these models. However, estimated 95% confidence intervals on the means of the WTP distributions are relatively narrow. Models that included journey duration as an independent predictor of heterogeneity in WTP did not perform as well as duration-specific models.

Table 7.5 Public transport service model results for changes in certain time

Purpose	Duration	N	$-LL_0$	$-LL$	k	Pseudo R^2	BIC	WTP[a] headway[b] ($/hour)	WTP seated time ($/hour)	WTP standing time ($/hour)
Commute	***All***	**1,830**	**1,373**	**1,216**	**7**	**0.114**	**1.342**	**8.28 [7.32, 9.24]**	**8.16 [6.36, 9.9]**	**11.88 [9.54, 14.22]**
Commute	Short	1,105	791	669	7	0.155	1.229	7.86 [6.78, 8.94]	10.02 [7.98, 12.06]	12.18 [9.36, 14.94]
Commute	Long	725	580	538	7	0.072	1.513	9.96 [7.98, 11.88]	5.64 [2.82, 8.52]	12.96 [8.58, 17.34]
Time-dependent	***All***	**1,520**	**1,003**	**865**	**7**	**0.137**	**1.153**	**4.62 [3.66, 5.58]**	**5.70 [4.02, 7.38]**	**8.46 [6.48, 10.5]**
Time-dependent	Short	1,060	673	590	7	0.123	1.133	3.78 [2.52, 5.04]	5.10 [3.24, 6.96]	8.22 [6.18, 10.2]
Time-dependent	Long	460	329	268	7	0.187	1.204	4.38 [3, 5.76]	6.96 [4.68, 9.18]	13.08 [9.78, 16.32]

Purpose	Duration	*N*	$-LL_0$	$-LL$	*k*	Pseudo R^2	BIC	WTP[a] headway[b] ($/hour)	WTP seated time ($/hour)	WTP standing time ($/hour)
Flexible	***All***	**1,160**	**750**	**669**	**7**	**0.108**	**1.172**	**8.10 [6.96, 9.3]**	**7.8 [5.94, 9.78]**	**12.78 [10.26, 15.24]**
Flexible	Short	815	528	479	7	0.093	1.200	8.52 [6.66, 10.38]	7.86 [4.56, 11.16]	12.42 [8.16, 16.68]
Flexible	Long	345	221	188	7	0.150	1.143	6.78 [4.74, 8.88]	7.02 [3.54, 10.44]	12.84 [6.6, 19.14]
All	***All***	**4,510**	**3,126**	**2,764**	**7**	**0.116**	**1.231**	**7.44 [6.9, 7.92]**	**7.56 [6.6, 8.58]**	**11.82 [10.5, 13.08]**

[a] WTP values = mean and [95% confidence intervals].
[b] Headway = WTP for a one-hour reduction in time between departure of services.

The results suggest the WTP for more frequent service, and possibly the cost of seated time, are lower for time-dependent journeys than for other journey types. The lower value for time-dependent trips is difficult to explain. One explanation might be that, because time-dependent trips have to be made, respondents are more willing to adjust starting time to ensure they get there, whereas for flexible trips respondents are somewhat indifferent about making the trip and are thus less willing to lengthen the trip. However, this is speculation in the absence of a clear understanding. Given this we suggest that the VTT for all non-commuting trips is estimated from a weighted average of the time-dependent and flexible trip values.[29] These are provided in Table 7.6.

Table 7.6 Public transport service model results – other (non-commuting trips)

	WTP headway ($/hour)	WTP seated time ($/hour)	WTP standing time ($/hour)
All	6.13 [5.09, 7.19]	6.61 [4.85, 8.42]	10.33 [8.12, 12.55]
Short	5.84 [4.32, 7.36]	6.3 [3.81, 8.79]	10.05 [7.04, 13.02]
Long	5.41 [3.75, 7.1]	6.99 [4.19, 9.72]	12.98 [8.42, 17.53]

Note: WTP values = mean and [95% confidence intervals]; Headway = WTP for a one-hour reduction in time between departure of services.

There are differences between journey purposes. In general, commuters place a higher value on headway (frequency of service) than do other travellers. The VTT (sitting and standing) is higher for commuting, apart from on long trips. There are differences between central value estimates for journey durations, but the confidence intervals have significant overlap.

Travellers are willing to pay more for a reduction in standing time than for a reduction in sitting time across all models. The WTP to avoid standing time is reasonably consistent across trip duration, but the difference between standing and sitting time is much greater for long trips. Headway values are dollars per hour of reduced time between services.

[29] We note below that Batley et al. (2019) have similarly averaged results where they were not easily explained.

7.5.2 Value of reliability

The reliability results are shown in Table 7.7. The WTP for changes in travel time is for mean but uncertain travel time, in contrast to the values in Table 7.5. As with the public transport service model results, we provide a weighted average of the time-dependent and flexible trip values in Table 7.8.

Table 7.7 Public transport reliability models

Purpose	Duration	N	$-LL_0$	$-LL$	k	Pseudo R^2	BIC	WTP uncertain travel time ($/hour)	WTP SD of travel time ($/hour)
Commute	***All***	**1,830**	**1,369**	**1,262**	**6**	**0.079**	**1.389**	**16.5 [15.18, 17.82]**	**14.64 [13.08, 16.26]**
Commute	Short	1,105	793	719	6	0.093	1.318	15.72 [13.68, 17.76]	15.18 [12.36, 17.94]
Commute	Long	725	572	540	6	0.057	1.513	13.32 [9.9, 16.74]	16.32 [12.18, 20.46]
Time-dependent	***All***	**1,520**	**1,002**	**913**	**6**	**0.089**	**1.214**	**5.64 [3.24, 8.04]**	**12.24 [9.96, 14.46]**
Time-dependent	Short	1,060	670	619	6	0.077	1.185	7.62 [4.68, 10.5]	11.64 [8.94, 14.34]
Time-dependent	Long	460	329	293	6	0.112	1.307	6.60 [2.16, 10.98]	12.30 [8.4, 16.2]
Flexible	***All***	**1,160**	**750**	**713**	**6**	**0.049**	**1.245**	**25.86 [21.66, 30.06]**	**22.98 [16.5, 29.46]**
Flexible	Short	815	528	503	6	0.046	1.257	31.02 [24.12, 37.86]	15.60 [9.66, 21.6]
Flexible	Long	345	222	200	6	0.099	1.202	17.28 [13.14, 21.42]	18.00 [12.9, 23.16]
All	***All***	**4,510**	**3,122**	**2,914**	**6**	**0.066**	**1.297**	**16.92 [16.14, 17.76]**	**13.56 [12.54, 14.58]**

Table 7.8 Public transport service model results – other (non-commuting trips)

	WTP uncertain travel time ($/hour)	WTP SD of travel time ($/hour)
All	14.39 [11.21, 17.57]	16.89 [12.79, 20.95]
Short	17.79 [13.13, 22.39]	13.36 [9.25, 17.5]
Long	11.18 [6.87, 15.45]	14.74 [10.33, 19.18]

Note: The results are based on the weighted average of time-dependent and flexible trips (Table 7.7).

The results suggest the mean WTP ($/hour) for reduction in average travel time is higher for short trips than for long trips, but the value of reliability (VoR) is greater for long trips.

There is little overall variation in the value of the SD of travel time. Values of average travel time exceed the values for both sitting and standing time in the service models.

7.6 Private transport

7.6.1 Value of travel time

The private transport results are shown in Table 7.9. As with the public transport analyses, average values for time-dependent and flexible journey (non-commuting) purposes are combined in Table 7.10.

Table 7.9 Private transport service models – VTT

Purpose	Duration	*N*	$-LL_0$	$-LL$	*k*	Pseudo R^2	BIC	WTP free-flow travel time ($/hour)	WTP heavy traffic travel time ($/hour)
Commute	***All***	**9,220**	**6,674**	**5,799**	**5**	**0.131**	**1.260**	**30.90 [29.7, 32.16]**	**57.24 [54.9, 59.64]**
Commute	Short	6,930	4,968	4,184	6	0.158	1.211	38.70 [36.54, 40.8]	73.74 [69.6, 78]
Commute	Medium	970	716	675	6	0.057	1.411	37.38 [29.16, 45.54]	59.64 [46.74, 72.6]
Commute	Long	1,320	988	894	6	0.095	1.369	30.6 [26.94, 34.32]	55.80 [49.44, 62.16]
Time-dependent	***All***	**11,505**	**7,948**	**6,785**	**5**	**0.146**	**1.181**	**27.48 [26.52, 28.44]**	**52.08 [50.1, 54.12]**
Time-dependent	Short	9,415	6,481	5,401	6	0.167	1.150	31.02 [29.52, 32.52]	62.28 [58.98, 65.58]
Time-dependent	Medium	815	578	522	6	0.096	1.303	40.74 [31.14, 50.34]	58.62 [46.14, 71.1]
Time-dependent	Long	1,275	889	825	6	0.072	1.308	29.58 [24.96, 34.2]	57.72 [48.9, 66.6]
Flexible	***All***	**8,945**	**5,953**	**5,234**	**6**	**0.121**	**1.173**	**36.00 [34.2, 37.74]**	**60.30 [57.06, 63.48]**
Flexible	Short	7,125	4,647	4,025	6	0.134	1.133	40.98 [38.64, 43.32]	74.82 [70.02, 79.68]
Flexible	Medium	860	603	547	6	0.094	1.292	37.50 [30.06, 44.94]	54.24 [43.98, 64.44]
Flexible	Long	960	702	631	6	0.101	1.333	34.32 [28.68, 39.96]	56.04 [47.94, 64.14]
All	***All***	**29,670**	**20,575**	**17,876**	**5**	**0.131**	**1.206**	**31.20 [30.42, 31.92]**	**57.42 [55.92, 58.86]**

Table 7.10 Private transport service models – other (non-commuting) VTT

	WTP free-flow travel time ($/hour)	WTP heavy traffic travel time ($/hour)
All	31.21 [29.88, 32.51]	55.68 [53.14, 58.21]
Short	35.31 [33.45, 37.17]	67.68 [63.74, 71.65]
Medium	39.08 [30.59, 47.57]	56.37 [45.03, 67.68]
Long	31.62 [26.56, 36.67]	57 [48.49, 65.54]

Note: The results are based on the weighted average of time-dependent and flexible trips (Table 7.9).

The results suggest the cost of time in heavy traffic is considerably higher than the cost of time in free-flow conditions. Free-flow travel time is valued at about $30/hour, whereas time in heavy traffic is valued at about $60/hour.

There are three duration classes for these models. Differences in values between durations are minor, and there does not appear to be a pattern for free-flow time values. For time in heavy traffic, values are higher for short duration journeys.

7.6.2 Value of reliability

Table 7.11 includes the results for the VoR for private transport.

Table 7.11 Private transport reliability models

Purpose	Duration	N	$-LL_0$	$-LL$	k	Pseudo R^2	BIC	WTP travel time ($/hour)	WTP SD of travel time ($/hour)
Commute	***All***	**9,245**	**6,681**	**6,278**	**5**	**0.060**	**1.360**	**38.40 [37.02, 39.72]**	**26.52 [24.72, 28.32]**
Commute	Short	6,950	4,966	4,579	6	0.078	1.321	37.56 [36.3, 38.82]	24.72 [22.68, 26.82]
Commute	Medium	975	721	670	6	0.071	1.392	48.60 [38.58, 58.62]	31.62 [22.74, 40.5]
Commute	Long	1,320	990	927	6	0.063	1.419	42.54 [36.18, 48.9]	20.58 [14.82, 26.34]
Time-dependent	***All***	**11,550**	**7,949**	**7,272**	**5**	**0.085**	**1.261**	**35.04 [34.26, 36]**	**21.06 [19.92, 22.14]**
Time-dependent	Short	9,455	6,470	5,791	5	0.105	1.227	35.22 [34.32, 36.12]	18.90 [17.58, 20.28]
Time-dependent	Medium	815	578	522	6	0.098	1.301	27.90 [25.2, 30.6]	20.04 [14.1, 26.04]
Time-dependent	Long	1,280	892	838	6	0.060	1.324	30.42 [27.3, 33.48]	10.56 [5.52, 15.54]
Flexible	***All***	**8,975**	**5,952**	**5,421**	**5**	**0.089**	**1.210**	**39.82 [39.80, 39.85]**	**29.98 [29.94, 30.02]**
Flexible	Short	7,140	4,639	4,228	6	0.088	1.188	39.42 [37.98, 40.86]	26.22 [24, 28.5]
Flexible	Medium	865	604	547	6	0.095	1.285	38.04 [32.58, 43.44]	17.28 [10.14, 24.48]
Flexible	Long	970	705	651	6	0.077	1.360	41.76 [33.96, 49.56]	0.48 [−12.6, 13.56]
All	***All***	**29,770**	**20,583**	**19,099**	**5**	**0.072**	**1.284**	**41.46 [40.68, 42.24]**	**24.00 [23.16, 24.78]**

Table 7.12 Private transport reliability models – other (non-commuting) VTT

	WTP travel time ($/hour)	WTP SD of travel time ($/hour)
All	37.13 [36.68, 37.68]	24.96 [24.3, 25.59]
Short	37.03 [35.89, 38.16]	22.05 [20.34, 23.82]
Medium	33.12 [29, 37.21]	18.62 [12.06, 25.24]
Long	35.31 [30.17, 40.41]	6.21 [−2.29, 14.69]

Note: The results are based on the weighted average of time-dependent and flexible trips (Table 7.11).

The results suggest that commuters have higher values for time and the SD of time than do travellers with other journey purposes, although the differences are not large.

7.7 Safety

7.7.1 Results

The results for the safety models are shown in Table 7.13 as the WTP for reductions in deaths and injuries and in Table 7.14 as the ratios between the WTP for different severities of death and injury.

Table 7.13 Safety models

Model type	Distribution	*k*	−*LL*	Pseudo R^2	BIC	WTP $/death avoided	WTP $/serious injury avoided	WTP $/minor injury avoided	Participant medians $/death or injury avoided
							Death	Serious	Minor
RPL	Triangular	6	25,059	0.226	1.729	4.46 [4.34, 4.58]	0.236 [0.22, 0.25]	0.029 [0.03, 0.03]	4.26
RPL	Maximum	9	24,719	0.237	1.706	4.13 [4.05, 4.2]	0.223 [0.22, 0.23]	0.027 [0.03, 0.03]	2.98
RPL	Lognormal	9	25,142	0.224	1.736	3.22 [3.03, 3.4]	0.190 [0.17, 0.21]	0.006 [0, 0.01]	3.41
RPL	Rayleigh	9	24,735	0.236	1.708	5.31 [4.18, 6.43]	0.256 [0.13, 0.39]	0.016 [0, 0.03]	3.04
GMXL	Triangular	8	27,371	0.155	1.889	4.29 [4.14, 4.45]	0.252 [0.24, 0.27]	0.041 [0.04, 0.04]	3.45
GMXL (WTP)	Triangular	7	28,970	0.105	1.999	4.17 [4.07, 4.27]	0.193 [0.18, 0.21]	0.02 [0.02, 0.02]	3.99
Constant		1	32,379						

Note: WTP estimates are means for survey participants. The social value of each event (eg, a single death avoided) requires aggregation over the relevant population. Medians are the medians of mean estimated WTP for the 5,799 individual participants.

There are several safety model variants. The RPL models reported in addition to the constrained triangular distribution are the lognormal, Rayleigh and maximum distributions. For the GMXL models, only the triangular and normal models converged, but the normal model is discarded because of negative WTP estimates. The only WTP-space model that converged was the triangular model. The two best-fitting models are the RPL maximum and Rayleigh distribution models. The GMXL and WTP-space models had significantly poorer fit than the RPL models.

Table 7.14 Safety model value estimate relativities of means

Model type	Distribution	Serious injury/ death	Minor injury/ death	Minor/serious injury	Serious/ death (median)	Minor/ death (median)	Minor/ serious (median)
RPL	Triangular	0.053	0.007	0.123	0.054	0.007	0.123
RPL	Maximum	0.039	0.006	0.119	0.059	0.009	0.139
RPL	Lognormal	0.035	0.001	0.012	0.068	0.001	0.017
RPL	Rayleigh	0.041	0.006	0.116	0.058	0.009	0.137
GMXL	Triangular	0.058	0.010	0.160	0.060	0.010	0.163
GMXL (WTP)	Triangular	0.046	0.005	0.101	0.048	0.005	0.102

Note: Reported relativities are estimates of the mean of the distribution of ratios of predicted means of relevant values (eg, mean WTP to avoid a serious injury/mean WTP to avoid a death) for each of the 5,799 individuals in the sample.

Monte Carlo sampling of 10,000 estimates generated from appropriately correlated attribute and cost coefficients generated means and 95% confidence intervals of money values for all except the maximum distribution. The RPL maximum money value estimates are derived from the distribution of mean estimates for the 5,799 individuals who completed these choices. Money value estimates are dollars per event per respondent. Aggregate values require multiplication by the number of relevant people in the population (see discussion below).

The central estimates of mean WTP to avoid a death all fall in the range $3.22 to $5.31. The outliers are the lognormal (low WTP) and Rayleigh (high WTP) RPL models. There are statistically significant differences between some of these estimates. For the other four distributions, estimates of mean WTP to avoid a death are broadly similar, in the relatively narrow range $4.13 to $4.29. Medians of predicted mean WTP for the 5,799 individuals in the sample are somewhat lower than the non-conditional means, ranging from $2.98 to $4.26.

All six models provide broadly consistent estimates of the value of avoiding a serious injury. There is little difference between the means and medians of the value of avoiding a serious injury. All means and medians fall in the range $0.190 to $0.256. Again, the lowest unconditional mean is for the lognormal distribution and the highest is for the Rayleigh distribution.

The lognormal RPL has a particularly low value for avoiding a minor injury, whether considering the unconditional mean or the sample median, whereas the GMXL model is high on both these measures.

The ratio of mean WTP to avoid a serious injury to mean WTP to avoid a death is in the order of 0.05 (with slightly higher medians), whereas the ratio of mean WTP to avoid a minor injury to mean WTP to avoid a death is in the order of 0.006 (again, with slightly higher medians). These ratios highlight the extreme values generated by the lognormal distribution, which is a clear outlier with the lowest value ratio for serious injury/death, and the other two ratios an order of magnitude lower than for the other distributions. The lognormal distribution produces the lowest values for all three attributes, but the difference is particularly marked for the two injury categories, with what appear to be unrealistically low values.

The lognormal model, which is the poorest fitting RPL model, provides the smallest mean WTP estimates across all three attributes. This, combined with the outlier ratio estimates of the lognormal model, suggests it is an inappropriate summary of these data. The Rayleigh model, which has similar fit to the best-fitting maximum model, provides the highest mean WTP values for death and serious injury.

The status-quo ASCs are highly significant and negative in all non-WTP-space models (Table C.1), with associated mean WTP (β_{SQ}/β_{Cost}) of −$459 (Rayleigh: = 1.985/−0.00432) to −$179 (GMXL: = 1.042/−0.00581), depending on the model. In the WTP-space model the status quo ASC WTP is −$175 (standard error (SE) = $6). There is a high-level of heterogeneity in these results, with SDs of the no change or status quo (SQ) parameter in the order of 2–3 times the mean, so the distribution has significant mass on both sides of zero. In aggregate, these results signal an overall preference for choosing an alternative other than the status quo, irrespective of attribute levels. This outcome is consistent with the warm-glow effect having prominence for some participants. However, inclusion of the ASC term in these models means that estimated attribute money values are independent of the warm-glow, stickiness, and order effects embodied in the ASC.

7.8 Aggregation and updating

7.8.1 Aggregation

The safety question has been designed to develop estimates of the WTP of the individual respondent to avoid an additional death or injury in the population. Respondents were asked to choose between options that varied in the total number of deaths and injuries in New Zealand and the change in their ‘personal costs’ via a tax or some other payment platform. This is an expression of a WTP for a public good outcome, with the benefits obtained by the whole population as a reduced number of road deaths and injuries.

In cost–benefit analysis (CBA) the standard approach to valuation of government policies or other programmes with national effects is to aggregate the individual WTP values by multiplying the average WTP by the number of individuals in the relevant population (eg, the total New Zealand adult population of close to four million).[30]

By definition, aggregating by the number of individuals in a population gives a higher total than aggregating by the number of households. A number of researchers have noted that households may be a more appropriate multiplier, depending on household resource allocation (Bateman & Munro, 2009; Delaney & O’Toole, 2006; Lindhjem & Navrud, 2009).[31] Strand (2007) cites different examples of income pooling amongst households, including unitary (or dictatorial) mode (where all household income is managed by one individual), income-pooling and the non-cooperative model. Research has suggested both the unitary model and household pooling means household WTP can be the same as individual WTP (Adamowicz et al., 2013; Bergstrom, 2003; Munro, 2005), while in the non-cooperative model the relationship between individual WTP and household WTP can be complex, and household WTP typically lies in between each individual WTP and the sum of the two (Bateman & Munro, 2009).

[30] For example, this is the same approach as used by Tait et al. (2016) in a study of the benefits of improvements in national water quality. They used a CE to generate individual WTP values, which were aggregated to national benefits by multiplying by the adult population. These values have been used in the measurement of benefits of freshwater policy changes.

[31] Similar differences in interpretation (individual or household) are noted by Comerford et al. (2009) relating to collection of data on expenditure.

This is an issue of household decision-making behaviour rather than income equivalence.[32] Research in New Zealand has suggested that the extent of income sharing within a household, and even between households, will vary with household structure, in addition to other attributes, particularly ethnicity (see Fleming, 1997, as cited in Hodgson & Birks, 2002; and Aziz et al., 2013).

The Ministry of Social Development has used the concept of the income sharing unit as a definition of the unit of income aggregation in its assessment of household poverty (Perry, 2007, 2019). All individuals in the household are assumed to benefit reasonably equally from the combined income of the household, regardless of household structure. Although this is not always the case, the Ministry of Social Development suggests it is defensible as an approximation (Perry, 2019) and is consistent with international practice (United Nations Economic Commission for Europe, 2011). By implication, this would suggest households as the appropriate unit of aggregation for individual WTP estimates.

7.8.1.1 Suggested approach

Although there is widespread acknowledgement that aggregating over individuals will over-estimate total WTP, there is not complete consensus in the literature on whether to aggregate using individual adults or household numbers. In addition, there is a significant practical use of an assumption of household income sharing and of household WTP, particularly in New Zealand. This is a significant assumption for the way in which the data derived from this study could be used, particularly in aggregating to estimate the value of a statistical life (VoSL) or the value of preventing a fatality (VPF).

Because of the ratio of household to adult numbers, aggregating WTP over the number of households produces a lower VPF by approximately 50% compared to aggregation using the number of adults (Table 7.15). We recommend using household aggregation (*c*.$8.1 million) as a minimum estimate of VoSL or VPF. Taking account of the mix of household types might suggest a value in between this and the higher value based on the number of adults.

Table 7.15 Implications of alternative aggregation methods for values of preventing fatalities and injuries

	WTP per event ($/respondent)	Aggregate (individual adults)	Aggregate (households)
Multiplier		3,958,580	1,908,700
Death	$4.13 – $4.46	$16.3 – $17.7 million	$7.9 – $8.5 million
Serious injury	$0.193 – $0.252	$764,006 – $997,562	$368,379 – $480,992
Minor injury	$0.007 – $0.041	$27,710 – $162,302	$13,361 – $78,257

7.8.2 Updating

The approach to updating these values for future years is not immediately obvious. WTP might increase with increased relative income, but it might also be influenced by the starting level of risk events or by the total population.

[32] For example, some research has focused on the development of equivalised household income measures to take account of household structure (see Stats NZ, 2019), which is used to suggest that a household with two adults can have an equivalent level of utility with a lower level of aggregate income than two one-adult households, but that the income equivalence is also affected by the number of children. However, this is addressing a different question from the probability that an adult respondent might be making a decision in a CE on the basis of shared household income or individual income. This might be unaffected by the number of children.

- In a larger population, WTP to reduce a fatality or injury by a given amount (eg, 10 fewer annual road deaths) in the community might be reduced because the probability of any connection to those people would be lower.
- People might value differently a reduction in the number of road deaths per annum from 300 to 290 compared to that from 200 to 190 or from 500 to 490; the number of deaths saved is the same, but the proportional change is quite different. Currently we do not know how marginal WTP changes with a change in the total number of deaths and injuries.

Given these uncertainties, a conservative approach would be to simply update using changes in relative income, as is done currently.

7.9 Comparisons of results

In this section we compare the values derived in this study with those in the MBCM currently and with international values.

The MBCM values are updated to current values using an index using average hourly earnings[33] (ordinary time), which is the same index used by the Ministry of Transport to update the safety values (Ministry of Transport, 2021). International values are first translated to New Zealand dollars at the date of the study using GDP per capita on a purchasing power parity basis (Organisation for Economic Co-operation and Development [OECD], 2022).

The review suggests:

- Many of the values used elsewhere for the attributes in this current study are based on literature reviews conducted some time ago, using mixes of methods and updated to take account of the value of money. Very few are based on primary research and particularly recent primary research.
- The results from this new study are within the range of the results of international studies.

7.9.1 Value of travel time

7.9.1.1 Estimated values compared to MBCM

New estimates

The new results suggest VTTs as shown in Table 7.16 in $/hour using current (2021) NZ$ values. Table 7.17 shows the results as a percentage of the average hourly wage rate ($34.76 at Q2 2021).[34] In the rest of this section we compare these with current values in the MBCM and those from studies in other countries.

[33] This uses StatsNZ average hourly earnings (ordinary time), Infoshare Table Reference: QEM003AA.

[34] Stats NZ Infoshare Table reference QEM003AA (average hourly earnings, ordinary time, all sectors, all sexes).

Table 7.16 Mean VTTs (2021 $/hour)

		Driving – free-flow	Driving – heavy traffic	Congestion increment	Public transport – seated	Public transport – standing
Commute	All	30.90	57.24	26.34	8.16	11.88
	Short	38.70	73.74	35.04	10.02	12.18
	Medium	37.38	59.64	22.26	not included	not included
	Long	30.60	55.80	25.20	5.64	12.96
Other	All	31.21	55.68	24.47	6.61	10.33
	Short	35.31	67.68	32.37	6.30	10.05
	Medium	39.08	56.37	17.29	not included	not included
	Long	31.62	57.00	25.38	6.99	12.98
All	**All**	**31.20**	**57.42**	**26.22**	**7.56**	**11.82**

Note: Other = weighted average of time-dependent and flexible.

Table 7.17 VTTs (% of average hourly wage)

		Driving – free-flow	Driving – heavy traffic	Congestion increment	Public transport – seated	Public transport – standing
Commute	All	89%	165%	76%	23%	34%
	Short	111%	212%	101%	29%	35%
	Medium	108%	172%	64%	not included	not included
	Long	88%	161%	72%	16%	37%
Other	All	90%	160%	70%	19%	30%
	Short	102%	195%	93%	18%	29%
	Medium	112%	162%	50%	not included	not included
	Long	91%	164%	73%	20%	37%
All	**All**	**90%**	**165%**	**75%**	**22%**	**34%**

Note: Other = weighted average of time-dependent and flexible.

Current MBCM values

The MBCM includes behavioural VTTs for transport modelling purposes only. The values are for all mode users by trip purpose in $/hour. Table 7.18 shows the values in 1 July 2002 dollars from Table 14 in the MBCM and updated values in 2021 NZ$ values. Table 7.19 provides VTTs by trip purpose for all modes combined. These values are used for calculating travel time benefits as an input into CBAs. The values are in $/hour for all trip purposes for congested or uncongested traffic conditions.

The revised values for commuting and non-commuting time are approximately double the current MBCM values. The congestion increment is even more significantly increased.

Table 7.18 Behavioural VTTs for mode user for transport modelling purposes ($/h)

Mode user	Commuting to/from work	Other non-work travel purposes	Commuting to/from work	Other non-work travel purposes
	$/h – July 2002	$/h – July 2002	$/h – July 2021	$/h – July 2021
Base VTTs				
Car, motorcycle driver	7.80	6.90	12.40	10.97
Car, motorcycle passenger	5.85	5.20	9.30	8.27
Light commercial driver	7.80	6.90	12.40	10.97
Light commercial passenger	5.85	5.20	9.30	8.27
Medium/heavy commercial driver	7.80	6.90	12.40	10.97
Medium/heavy commercial passenger	5.85	5.20	9.30	8.27
Seated bus and train passenger	4.70	3.05	7.47	4.85
Standing bus and train passenger	6.60	4.25	10.49	6.76
Pedestrian and cyclist	6.60	4.25	10.49	6.76
Congestion increment				
Car, motorcycle driver	–	2.75	–	4.37
Car, motorcycle passenger	–	2.05	–	3.26
Commercial vehicle driver	–	2.75	–	4.37
Commercial vehicle passenger	–	2.05	–	3.26

Source: Adapted from Waka Kotahi (2021, p. 50) (update factor to July 2021 of 1.59 – see MBCM Appendix A12).

Table 7.19 VTTs by trip purpose ($/h/person)

Trip purpose	Base VTT (2002 $/h/person)	Congestion value (2002 $/h/person)	Base VTT (2021 $/h/person)	Congestion value (2021 $/h/person)
Commuting to/from work	7.80	3.15	14.58	5.89
Other non-work travel purpose	6.90	2.75	12.89	5.14

Source: Adapted from Waka Kotahi (2021, p. 50).

7.9.1.2 International comparisons

The Australian Transport Assessment and Planning (ATAP) Guidelines[35] use a value of private travel time equal to 40% of seasonally adjusted full-time average weekly earnings for Australia. This is A$14.99/hr in June 2013 A$ values (ATAP, n.d.), equivalent to current (2021) NZ$18.88/hour. The 40% value is based on an Austroads (1997) study that summarised values from other countries that varied from 21% to 93% of the average wage; it appeared to be particularly influenced by studies in Canada and New Zealand that had suggested multipliers of 40%. The New Zealand studies appear to include that of Symonds Travers Morgan (1997), which produced a value of 40% of the average wage based, in turn, on a review of international

[35] The ATAP Guidelines have replaced the previous *National Guidelines for Transport System Management* and Austroads' *Guide to Project Evaluation* and *Guide to Road Transport Planning*. (https://austroads.com.au/infrastructure/project-delivery/atap-australian-transport-assessment-and-planning).

studies, particularly a UK analysis in the early to mid-1980s. This percentage was used to establish New Zealand values in 1991 dollars.

Two recent Australian studies have found relatively high VTTs. Hensher et al. (2021) found commuting VTT in Sydney was A$25.53/hour, and they identified significantly higher WTP for those who spent more time working from home during the COVID-19 pandemic. For individuals who did not work from home, the mean was $20.39/hour.[36] Fayyazz et al. (2021) found that VTT was different for two classes of travellers identified in a latent class model. Their results also varied according to whether they were estimated in an SC experiment (A$31.30/hour for class 1 and A$12.14/hour for class 2) or from their choices on a driving simulator (A$41.90/hour for class 1 and A$16.20/hour for class 2).

In the UK, a recent study including primary research has provided new values for VTT (Table 7.20), which the authors note have largely been accepted and implemented by the UK Department for Transport. Converted to current New Zealand dollars, the commuting values are similar to the new suggested New Zealand values (Table 7.21), although the non-commute values are lower.

Table 7.20 UK VTTs

Distance	Commute (2014 £/hr)	Non-commute (2014 £/hr)	Commute (2021 NZ$/hr)	Non-commute (2021 NZ$/hr)
All	11.21	5.12	28.50	13.02
< 20 miles	8.27	3.62	21.03	9.20
20–100 miles	12.15	6.49	30.89	16.50
≥ 100 miles	12.15	9.27	30.89	23.57

Source: UK values adapted from Batley et al. (2019, p. 612).

Batley et al. (2019) also estimated ratios of VTT for different modes (Table 7.21) suggesting lower VTTs for public transport commuting relative to commuting by car, a finding consistent with ours, but they found the opposite for non-commuting travel (higher VTTs for public transport), which we did not. They suggest some of this is explained by a ‘comfort’ factor, although they note the values for rail and other public transport[37] are different from their expectations. They note a preference for using mode-free values for other (non-commuting) trips, by averaging the values over the sample of trips for all (motorised) modes, while maintaining the distance weighting.

[36] The direction of causality is not known – that is, whether those who work more from home have a higher VTT or if more working from home results in a higher marginal VTT.

[37] ‘Other public transport’ refers to trams, light rail or the London Underground.

Table 7.21 Ratio of modal VTT by trip purpose for an average person

	Commute	Other non-work	Employee's business
Car	1.00	1.00	1.00
Bus	0.51	2.14	N/A
Other public transport	0.99	3.19	0.69
Rail	0.73	2.29	0.39

Source: Adapted from Batley et al. (2019, p. 606).

In the US, the VTT is based on a percentage of hourly wages, assuming 50% of wages for personal local travel (in a range of 35–60%), 70% for personal intercity travel (by road or rail, and in a range of 60–90%) and 100% (road or rail and range of 80–120%) of the wage rate for business travel (Table 7.22). These percentages were first adopted in 1997 guidance based on literature and the views of experts (US Department of Transportation, 1997).

Table 7.22 US VTTs

	Personal	Business	All purposes	Personal	Business	All purposes
	(2015 US$/hr)			(2021 NZ$/hr)		
Local travel	13.6	25.4	14.1	23.57	44.01	24.43
Intercity	19	25.4	20.4	32.92	44.01	35.35

Source: Adapted from US Department of Transportation (2016, p. 17).

A meta-analysis of European studies suggested a wide range of VTTs as a percentage of wage rates taken from OECD statistics (Table 7.23).

Table 7.23 VTTs in Europe (% of gross wage rate)

Mode	Car commute	Car other	Employer business	Public transport
Urban free-flow	41% (19–85%)	36% (17–75%)	77% (38–155%)	–
Urban congested	58% (27–121%)	51% (23–107%)	110% (53–220%)	–
Intercity free-flow	–	49% (23–104%)	107% (52–214%)	158% (75–321%)
Urban bus commute	–	–	–	31% (14–66%)

Source: Adapted from Wardman et al. (2016, p. 25).

Congestion

UK multipliers for different levels of congestion are shown in Table 7.24. The difference between free-flow and heavy traffic is greater than we find, although this might reflect differences in perceptions of heavy traffic between the two countries.

Table 7.24 UK VTT multipliers

Mode	Multiplier	Commute	Other non-work	Employee's business
Car	Free-flow	0.51	0.47	0.42
	Light traffic	0.72	0.83	0.68
	Heavy traffic	1.37	1.89	1.26
Bus	Value of free-flow	0.99	1.22	N/A
	Value of slow down	1.39	1.36	N/A
	Value of dwell time	0.68	1.57	N/A
	Value of headway	1.68	1.60	N/A

Source: Adapted from Batley et al. (2019, p. 614).

7.9.2 Reliability – Private transport

7.9.2.1 Estimated values compared to MBCM

The current MBCM estimates the benefits of improved reliability in travel time using a formula of the form:

$$0.9 \times \textit{travel time value (\$/h)} \times \textit{(reduction in the network variability (in min)/60)} \times \textit{traffic volume for time period (veh/h)} \times \textit{correction factor} \qquad \text{(Equation 7.3)}$$

Where the reduction in network variability is the difference between the sums of the variability (based on SD of travel time) for all journeys in the modelled area for the do-minimum and project option. The 0.9 factor is the VoR based on a typical urban traffic mix.[38] The correction factor depends on the scale of the analysis (and therefore how much of total variability is accounted for) – for example, it is 100% for analysis using a regional model but only 30% if analysing a single intersection or passing lane.

In this new study we have examined the value of a change in the SD of travel time. This is estimated in \$/hour; in Table 7.25 we show this as a reliability ratio (RR), where:

$$\textit{RR = Value of SD of travel time / Value of travel time} \qquad \text{(Equation 7.4)}$$

Table 7.25 New Zealand RRs – private transport

	Commute	Other	All
All	0.69	0.67	0.58
Short	0.66	0.60	–
Medium	0.65	0.56	–
Long	0.48	0.18	–

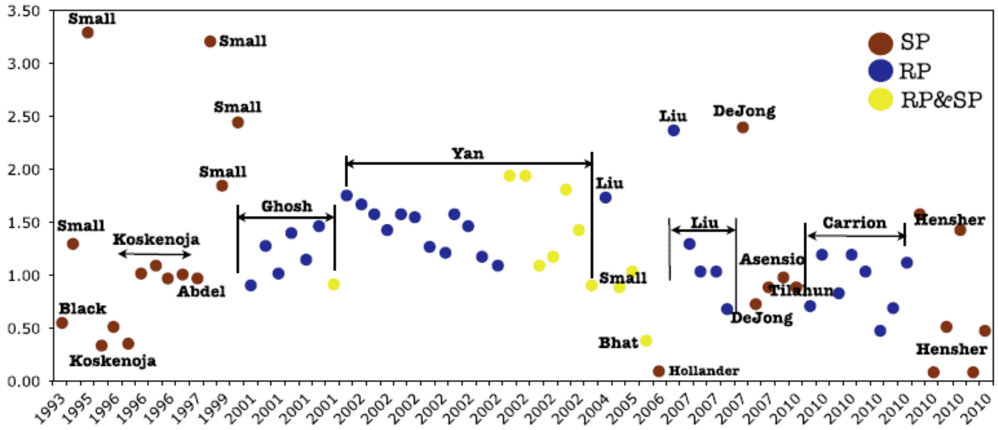
Source: Estimated from data in Table 7.11.

[38] For projects with a significantly different vehicle mix, evaluators should use 0.8 for cars and 1.2 for commercial vehicles.

7.9.2.2 International comparisons

In the UK, an RR of 0.4 is recommended for cars (Department for Transport, 2021b), although other international studies suggest a wide range of ratios (Figure 7.2).

Figure 7.2 RRs in published studies



Source: Reprinted from Carrion and Levison (2012, p. 734).

7.9.3 Reliability – Public transport

7.9.3.1 Estimated values compared to MBCM

Currently the MBCM uses minute-late ratios as the basis for estimating the value of an improvement in reliability, measured as the savings in minutes late. The new study uses the same basis for defining RR as for private transport; the results are shown in Table 7.26.

Table 7.26 New Zealand RRs – public transport

	Commute	Other	All
All	0.89	1.17	0.80
Short	0.97	0.75	–
Long	1.23	1.32	–

Source: Estimated from data in Table 7.7.

7.9.3.2 International comparisons

A recent Australian study has assessed methodologies for valuing reliability but has not identified WTP values to populate it (ATAP Steering Committee, 2021). In the UK, the current assumption is an RR of 1.4 for public transport (Department for Transport, 2021b).

7.9.4 Safety values – Fatalities

7.9.4.1 Estimated values compared to MBCM

Estimates of safety values in New Zealand and elsewhere have focused on the identification of the VoSL, with values for reducing the risk of injuries estimated as a percentage of the VoSL. The VoSL is defined as the value of reducing the risk of a fatality, with analysts stressing that this is not a value of any individual's life or of the benefits of preventing a death of a specific person (Box 7.1). Alternative terminology such as 'value of reduced mortality risk' (VRMR) or 'value of preventing a fatality' (VPF) may be more easily understood.

Box 7.1 Terminology for reduced fatality benefits

The value of a statistical life (VoSL) is the term often used when quantifying the benefits of reductions in the risks of fatalities (Viscusi, 2005) and is widely used in New Zealand. VoSL is not used to estimate the value of the life of any individual but is the value to society of reducing the risk of fatalities. The US Environmental Protection Agency (2022, para. 2) defines VoSL as follows:

> *Suppose each person in a sample of 100,000 people were asked how much he or she would be willing to pay for a reduction in their individual risk of dying of 1 in 100,000, or 0.001%, over the next year. Since this reduction in risk would mean that we would expect one fewer death among the sample of 100,000 people over the next year on average, this is sometimes described as 'one statistical life saved.' Now suppose that the average response to this hypothetical question was $100. Then the total dollar amount that the group would be willing to pay to save one statistical life in a year would be $100 per person × 100,000 people, or $10 million. This is what is meant by the 'value of a statistical life.' Importantly, this is not an estimate of how much money any single individual or group would be willing to pay to prevent the certain death of any particular person.*

Cameron (2010) outlined the misinterpretation and confusion caused by VoSL and WTP. She preferred 'willingness to swap for a microrisk reduction' (WTS). However, Penn and Hu (2018) tested differences in willingness to participate in a VoSL study when invitations used VoSL, WTS and two alternatives, finding no significant differences. Dockins et al. (2018) note that VoSL is well understood by economists (as the marginal rate of substitution (MRS) between mortality risk and money), but to many others, including decision-makers and media professionals, the term resembles 'obfuscated jargon bordering on the immoral'. They examined several alternative terms, suggesting that 'value of reduced mortality risk' (VRMR) was the most effective and readily understood alternative. It is similar to the more generic 'value of risk reductions' (VRR) as used by Rizzi and Ortúzar (2006a) and others, or 'value of a fatal risk reduction' (VFRR) (González et al., 2018).

In the UK the preferred term is the 'value of preventing a fatality' (VPF) (Glover & Henderson, 2010), and this is the term preferred by Clough et al. (2018) also. A US study based on focus group feedback preferred the 'value of improved chance of survival' (VICS) or VRMR (Simon et al., 2019).

Terms such as VRMR (or VRR when referring to mortality and/or injury risks) are more closely related to how the value is used in transport studies – that is, it takes account of the fact that transport investments reduce the risk of fatalities rather than preventing a fatality in any measurable way: *ex post* it could never be shown that a fatality had been prevented, but statistical analysis might be used to show the difference between expected numbers of fatalities and the out-turn number. The term can be extended to injuries also – that is, 'value of reduced injury risk'.

In New Zealand, the current VoSL was derived from a set of questions included in the Ministry of Transport Household Travel Survey of 1989/90. Questions relating to reductions in risks to different people (eg, the respondent, their family, or other people) were provided through different hypothetical goods/services (eg, safer toll road, road safety course, car safety features, safer neighbourhood and extra taxes for road safety) (Miller & Guria, 1991). The survey results were used to recommend a VoSL of $2 million in 1991 NZ$ values, and this was used by the Government to set the VoSL for all transport sector evaluations (Leung, 2009), with annual adjustments being made for changes in the value of a dollar.

A repeated survey in 1997/98 suggested a new VoSL of $4 million (1998 NZ$), significantly higher than the wage inflation-adjusted value of $2.5 million (Guria et al., 2003). Due to 'some unresolved policy issues', the Government did not adopt the revised VoSL estimate, and the VoSL continues to be based on the value established in 1991 (Leung, 2009). One of the arguments against adoption of the higher value was it would result in a shift of resources away from investments in travel time savings and decongestion, and towards road safety features (Clough et al., 2015). If the 1997/98 value had been adopted, it would convert to a current value of approximately $8.3 million.

The different values are shown in Table 7.27 alongside a summary value from this current study.

Table 7.27 New Zealand estimates of VoSL ($ million)

	Value	Study year	2021 NZ$
This study (household aggregation)	7.9–8.5	2021	7.9–8.5
MBCM value[a]	2.0	1991	4.6
1997/98 survey	4.0	1998	8.3

[a] Based on Ministry of Transport (2021) value of $4.42 million in 2020 NZ$.

7.9.4.2 International comparisons

There are numerous studies of VoSL in the literature and also many meta-analyses that bring the results of several studies together. Recently, Ananthapavan et al. (2021) have undertaken a meta-analysis of meta-analyses. They included studies published from 2007 onwards and extracted the minimum, maximum and median value in each review, removing studies that were included in more than one paper; means were used when medians were not available. Their results (in 2007 A$ converted into 2021 NZ$) are shown in Table 7.28. There is a wide range, with means ranging from $3 million to $34 million and a mean of median values of $10 million.

Table 7.28 Summary of meta-analyses of VoSL (2021 NZ$ million)

Study	Minimum VoSL estimate	Maximum VoSL estimate	Median VoSL estimate
Bahamonde-Birke et al. (2015) – SC studies	4.1	129.5	8.4
Dekker et al. (2011)	1.9	3.7	2.1
Hultkrantz & Svensson (2012)	2.8	27.3	8.2
Lindhjem et al. (2011)	0.3	11.4	5.2
Robinson & Hammitt (2016)	10.5	17.5	17.1
Robinson et al. (2019)	0.6	31.3	12.8
Milligan et al. (2014)	1.6	19.9	16.7
Median stated preference VoSL	1.6	19.9	8.4
Mean stated preference VoSL	3.1	34.4	10.1

Source: Adapted from Ananthapavan et al. (2021).

Values used by government agencies for official purposes and CBAs are somewhat different.

Australia

The ATAP Guidelines use a VoSL estimated by the New South Wales Roads & Traffic Authority in 2008,[39] updated to current dollars until a national WTP study is undertaken (ATAP, n.d.). The latest values (Table 7.29) convert to a current New Zealand equivalent of approximately $9.3 million.

[39] The latest updated New South Wales values are provided in Transport for New South Wales (2020).

Table 7.29 ATAP values for VoSL in Australia

	VoSL (2013 A$ million)	VoSL (2021 NZ$ million)
Urban	7.43	9.35
Non-urban	7.34	9.25
Average	7.38	9.30

Source: Australian values from ATAP (n.d., Table 14).

An alternative set of Australian values is derived by studies by Hensher and others. These estimate VoSLs for car occupants and pedestrians that range between approximately NZ$7 million and NZ$10 million in current values (Table 7.30).

Table 7.30 Alternative estimates for VoSL in Australia

	Car occupants (2007 A$ million)	Car occupants (2021 NZ$ million)	Pedestrians (2007 A$ million)	Pedestrians (2021 NZ$ million)
Urban	6.37	10.26	5.35	8.62
Non-urban	6.30	10.14	4.24	6.84
Average	6.33	10.20	4.80	7.73

Source: 2007 A$ values from Hensher et al. (2009, p. 18) and Hensher et al. (2011, p. 90).

UK

Following HM Treasury (2003) guidance, road fatalities have been valued in the UK using a WTP approach with values based on an original 1997 review, updated annually using GDP per capita. The cost per fatality is £2.06 million in 2021£ (Department for Transport, 2021a),[40] equivalent to a current (2021) NZ$ value of $4.63 million.

USA

The US Department of Transportation suggests a 2021 VoSL of US$11.8 million.[41] This is based on hedonic wage studies that were used in a 2013 analysis to yield a value of US$9.1 million in 2012 US$. Values are updated annually using the following formula:[42]

$$VoSL_T = VoSL_0 \times (P_T / P_0) \times (I_T / I_0)^{\varepsilon} \qquad \text{(Equation 7.5)}$$

Where 0 = original base year
T = current base year
P_t = price index in year T (Consumer Price Index for All Urban Consumers)
I_t = real incomes in year T (median usual weekly earnings)
ε = income elasticity of VoSL (assumed to be = 1.0)

[40] See ‘Table A 4.1.1: Average value of prevention per casualty by severity and element of cost’.

[41] US Department of Transportation (2021a).

[42] US Department of Transportation (2021b).

7.9.5 Injuries

7.9.5.1 Estimated values compared to MBCM

The new estimated values and the ratios of the value of fatal to serious and minor injuries are shown in Table 7.31.

Table 7.31 WTP and value of reduced injury by injury type (2021 $NZ)

	Fatal	Serious injury	Minor injury
Cost of injury ($/person WTP)	$4.13 – $4.46	$0.190 – $0.256	$0.006 – $0.041
Value of reduced injury ($/incident)	$7.9 – $8.5 million	$368,379 – $480,992	$13,361 – $78,257
Percentage of fatality	100%	4.6% – 5.9%	0.2% – 1.0%

In the MBCM the value of reduced injury risk has been estimated using a ratio of VoSL. The current approach used in New Zealand is to estimate the costs of injuries using 10% and 0.4% of the VoSL for serious and minor injuries respectively (Table 7.32); this was based on the results of the 1997/98 survey (Leung, 2009).

Table 7.32 New Zealand social costs by injury type (2020 NZ$/event)

	Fatal	Serious injury	Minor injury
Cost of injury	4,423,800	442,400	17,700
Percentage of fatality	100%	10%	0.4%

Source: Adapted from Ministry of Transport (2021, p. 12).

7.9.5.2 International comparisons

Table 7.33 shows values for Australia.[43] The ratio of serious injury to fatality is smaller than used in New Zealand and in the values suggested in this study. The value for minor injuries is within the range suggested by the new estimates.

Table 7.33 Australia social costs by injury type (2013 A$)

	Fatal	Serious injury	Hospitalised injury	Minor injury
Urban	7,425,629	361,733	87,988	19,296
Non-urban	7,342,167	226,025	65,210	23,678
Average	7,383,898	293,879	76,599	21,487
Ratio to VoSL	100%	4.0%	1.0%	0.3%

Source: Adapted from ATAP (n.d., Table 14).

UK value ratios are shown in Table 7.34 using current (2021) value estimates in the webTAG.

[43] These ratios are the same as found by Hensher et al. (2009), although for a study of VRR for pedestrians, higher injury ratios were identified.

Table 7.34 UK social costs by injury type (2021 £)

Injury type	2021 £	Ratio to fatal
Fatal	2,063,940	100%
Serious	202,046	9.8%
Slight	14,790	0.7%

Source: Adapted from Department for Transport (2021a, Table A 4.1.1).

In the US, the Maximum Abbreviated Injury Scale (MAIS) is used to assign ratios of impact relative to VoSL (Table 7.35). Those for serious injuries[44] are assumed to have a value equal to 10.5% of the VoSL. These fractions have been estimated using analysis of changes in quality-adjusted life years for different types of injury.

Table 7.35 Value of reduced injuries as a fraction of VoSL for the US

MAIS level	Severity[a]	Fraction of VoSL
1	Minor	0.003
2	Moderate	0.047
3	Serious	0.105
4	Severe	0.266
5	Critical	0.593
6	Unsurvivable	1.000

[a] More detail on severity descriptions is provided in Federal Aviation Authority (n.d.).
Source: Adapted from US Department of Transportation (2021b, p. 10).

Table 7.36 shows estimates of the ratios used in a selection of other countries as summarised in a recent meta-analysis.[45]

Table 7.36 Fatality and injury benefit reduction values (2012 US$000) in other countries and value ratios

Country	Fatal	Serious injury	Minor injury	Serious (% of fatal)	Minor (% of fatal)
Austria	3,649.2	461.5	32.5	13%	0.9%
Germany	1,391.7	144.2	5.9	10%	0.4%
Netherlands	3,181.5	341.8	7.7	11%	0.2%
Singapore	1,840.8	165.6	15.9	9%	0.9%
Switzerland	2,893.8	523.6	25.6	18%	0.9%

Source: Adapted from Wijnen and Stipdonk (2016, p. 102).

[44] Defined to include major nerve laceration; multiple rib fracture (but without flail chest); abdominal organ contusion; and hand, foot, or arm crush/amputation (Federal Aviation Authority, n.d.).

[45] This study summarises values and ratios from non-English texts. However, we note we estimate different values from Wijnen and Stipdonk (2016) for Australia, New Zealand, the UK and the US.

8 Conclusions and discussions

This study has set out to provide data and analysis to support robust monetary values for a range of attributes. This has been a challenging task and has required us to develop approaches that are different from those adopted in other countries. Comparisons with values in other countries have not always been informative because there has been relatively little primary research developing values adopted by transport agencies in many countries. The robustness of the values developed in this study are therefore attested by the statistical analysis.

The overall results are summarised below using mean values, with the greater detail on ranges in the tables in chapter 7 and as noted underneath Table 8.1 and Table 8.2 below. Separate values are provided for time and reliability for public and private transport. Safety attribute values are independent of travel mode; they are WTP values applicable to all people, whether transport users or not. We provide minimum and maximum values below, rather than simply as means.

Table 8.1 Public transport – mean value of travel time and reliability ($/hour)

Trip purpose	Trip length	Mean	Sitting	Standing	Reliability (SD of travel time)
Commuting	All	16.5	8.16	11.88	14.64
	Short	15.72	10.02	12.18	15.18
	Long	13.32	5.64	12.96	16.32
Other	All	14.39	6.61	10.33	16.89
	Short	17.79	6.30	10.05	13.36
	Long	11.18	6.99	12.98	14.74

Source: Tables 7.5–7.8 in chapter 7.

Table 8.2 Private transport – mean value of travel time and reliability ($/hour)

Trip purpose	Trip length	Mean	Free-flowing	Heavy traffic	Congestion increment	Reliability (SD of travel time)
Commuting	All	38.40	30.90	57.24	26.34	26.52
	Short	37.56	38.70	73.74	35.04	24.72
	Medium	48.60	37.38	59.64	22.26	31.62
	Long	42.54	30.60	55.80	25.20	20.58
Other	All	37.13	31.97	57.07	25.10	24.96
	Short	37.03	36.32	69.78	33.46	22.05
	Medium	33.12	42.90	61.61	18.71	18.62
	Long	35.31	34.04	60.48	26.44	6.21

Source: Tables 7.9–7.12 in chapter 7.

Table 8.3 Value of preventing fatalities and injuries

	WTP per event ($/respondent)	Minimum aggregate national value	Maximum aggregate national value
Death	$4.13 – $4.46	$7.9 – $8.5 million	$16.3 – $17.7 million
Serious injury	$0.193 – $0.252	$368,379 – $480,992	$764,006 – $997,562
Minor injury	$0.007 – $0.041	$13,361 – $78,257	$27,710 – $162,302

Source: Table 7.15 in chapter 7.

References

Abelson, P. (2008). *Establishing a monetary value for lives saved: Issues and controversies*. Working papers in cost-benefit analysis 2008-02, Department of Finance and Deregulation, Australia.

Abrantes, P. A. L., & Wardman, M. R. (2011). Meta-analysis of UK values of travel time: An update. *Transportation Research Part A: Policy and Practice, 45*(1), 1–17.

Ackerman, F., & Heinzerling, L. (2002). Pricing the priceless: Cost-benefit analysis of environmental protection. *University of Pennsylvania Law Review*, 150, 1553–1584.

Adamowicz, W., Dickie, M., Gerking, S., Veronesi, M., & Zinner, D. (2013). *Household decision-making and valuation of environmental health risks to parents and their children* (US EPA National Center for Environmental Economics Working Paper #13-06).

AECOM. (2009). *The role of soft measures in influencing patronage growth and modal split in the bus market in England*. Final Report to Department for Transport, UK.

Alberini, A. (2005). What is a life worth? Robustness of VSL values from contingent valuation surveys. *Risk Analysis, 25*(4), 783–800.

Alberini, A. (2019). Revealed versus stated preferences: What have we learned about valuation and behavior? *Review of Environmental Economics and Policy, 13*(2), 283–298.

Alberini, A., & Ščasný, M. (2013). Exploring heterogeneity in the value of a statistical life: Cause of death v. risk perceptions. *Ecological Economics, 94*, 143–155.

Ananthapavan, J., Moodie, M., Milat, A. J., & Carter, R. (2021). Systematic review to update 'Value of a Statistical Life' estimates for Australia. *International Journal of Environmental Research and Public Health*, *18*(11), 6168. https://doi.org/10.3390/ijerph18116168

Andersson, H., Levivier, E., & Lindberg, G. (2019). Private and public willingness to pay for safety: A validity test. *Accident Analysis & Prevention, 123*, 170–175.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, *100*(401), 464–477.

Antoniou, C. (2014). A stated-preference study of the willingness-to-pay to reduce traffic risk in urban vs. rural roads. *European Transport Research Review*, *6*, 31–42

Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of Political Economy*, *58*(4), 328–346.

Arup, Institute for Transport Studies (ITS), & Accent. (2015). *Provision of market research for value of travel time savings and reliability*. Phase 2 Report.

Australian Transport Assessment and Planning (ATAP). (n.d.). *Crash costs*. https://www.atap.gov.au/parameter-values/road-transport/4-crash-costs

Australian Transport Assessment and Planning (ATAP) Steering Committee. (2021). *Australian transport assessment and planning guidelines, road reliability measurement – research report*.

Austroads. (1997). *Value of travel time savings*.

Aziz, O., Gemmell, N., & Laws, A. (2013). *The distribution of income and fiscal incidence by age and gender: Some evidence from New Zealand*. Victoria University Working Papers in Public Finance, Working Paper 10/2013.

Bahamonde-Birke, F. J., Kunert, U., & Link, H. (2015). The value of statistical life in a road safety context – a review of the current literature. *Transport* Reviews, *35*, 488–511.

Balmford, B., Bateman, I. J., Bolt, K., Day, B., & Ferrini, S. (2019). The value of statistical life for adults and children: Comparisons of the contingent valuation and chained approaches. *Resource and Energy Economics*, *57*, 68–84.

Bateman, I., 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., & Swanson, J. (2002). *Economic valuation with stated preference techniques: A manual*. Edward Elgar.

Bateman, I. J., & Munro, A. (2009). Household versus individual valuation: What's the difference? *Environmental and Resource Economics, 43*, 119–135.

Bates, J., Polak, J., Jones, P., & Cook, A. (2001). The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review*, *37*, 191–229.

Batley, R., Bates, J., Bliemer, M., Börjesson, M., Bourdon, J., Cabral, M. O., Chintakayala, P. K., Choudhury, C., Daly, A., Dekker, T., Drivyla, E., Fowkes, T., Hess, S., Heywood, C., Johnson, D., Laird, J., Mackie, P., Parkin, J., Sanders, S., Sheldon, R., Wardman, M., & Worsley, T. (2019). New appraisal values of travel time saving and reliability in Great Britain. *Transportation*, *46*, 583–621.

Bennett, J., & Blamey, R. (2001). *The choice modelling approach to environmental valuation*. Edward Elgar.

Bergstrom, T. C. (2003). *Benefit cost analysis and the entanglements of love.* https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.331.1056&rep=rep1&type=pdf

Bhat, C. R., & Sardesai, R. (2006). The impact of stop-making and travel time reliability on commute mode choice. *Transportation Research Part B: Methodological*, *40*(9), 709–730.

Birol, E., & Khoundouri, P. (2008). *Choice experiments informing public policy*. Edward Elgar.

Bishop, R., & Boyle, K. J. (2019). Reliability and validity in nonmarket valuation. *Environmental and Resource Economics, 72*(2), 559–582.

Bishop, R., Boyle, K. J., Carson, R. T., Chapman, D., Hanemann, W. M., Kanninen, B., Kopp, R. J., Krosnick, J. A., List, J., Meade, N., Paterson, R., Presser, S., Smith, V. K., Tourangeau, R., Welsh, M., Wooldridge, J. M., deBell, M., Donovan, C., Konopka, M., & Sherer, N. (2017). Putting a value on injuries to natural assets: The BP oil spill. *Science*, 356, 253–254.

Bliemer, M. C. J., & Rose, J. M. (2006). *Designing stated choice experiments: state-of-the-art*. Institute of Transport Studies, University of Sydney.

Bliemer, M. C. J., & Rose, J. M. (2013). Confidence intervals of willingness-to-pay for random coefficient logit models. *Transportation Research Part B: Methodological*, *58*, 199–214.

Bronfman, N. C., & Cifuentes, L. A. (2003). Risk perception in a developing country: The case of Chile. *Risk Analysis*, 23, 1271–1285.

Cameron, T. A. (2010). Euthanizing the value of a statistical life. *Review of Environmental Economics and Policy, 4*(2), 161–178.

Carlsson, F., & Martinsson, P. (2006). *How much is too much? An investigation of the effect of the number of choice sets, starting point and the choice of bid vectors in choice experiments.* http://hdl.handle.net/2077/2731

Carrion, C., & Levinson, D. (2012). Value of travel time reliability: A review of current evidence. *Transportation Research Part A: Policy and Practice*, *46*, 720–741.

Carson, R. T. (2012). Contingent valuation: A practical alternative when prices aren't available. *Journal of Economic Perspectives*, *26*(4), 27–42.

Carson, R. T., Groves, T., & List, J. A. (2014). Consequentiality: A theoretical and experimental exploration of a single binary choice. *Journal of the Association of Environmental and Resource Economists*, *1*, 171–207.

Carthy, T., Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N., & Spencer, A. (1998). On the contingent valuation of safety and the safety of contingent valuation: Part 2 – The CV/SG "Chained" approach. *Journal of Risk and Uncertainty*, *17*(3), 187–213.

Chilton, S., Covey, J., Jones-Lee, M., Loomes, G., Pidgeon, N., & Spencer, A. (2015). Response to 'Testing the validity of the "value of a prevented fatality" (VPF) used to assess UK safety measures'. *Process Safety and Environmental Protection*, *93*, 293–298.

Clough, P., Chilton, S. M., Jones-Lee, M. W., & Metcalf, H. R. T. (2018). A new approach to environmental valuation for New Zealand. *Policy Quarterly*, *14*(2), 50–57.

Clough, P., Guria, J., & Bealing, M. (2015). *Approaches to valuing injury and mortality risk in transport assessments* (NZ Transport Agency research report 571). https://www.nzta.govt.nz/resources/research/reports/571

Comerford, D., Delaney, L., & Harmon, C. (2009). *Experimental tests of survey responses to expenditure questions* (Institute for the Study of Labor (IZA) Discussion Paper No. 4389).

Daly, A., Hess, S., & Train, K. (2012). Assuring finite moments for willingness to pay in random coefficient models. *Transportation*, *39*, 19–31.

De Blaeij, A, Rietveld, P., Verhoef, E., & Nijkamp, P. (2002). The valuation of a statistical life in road safety: A stated preference approach. In *Proceedings 30th European Transport Forum*.

Dekker, T., Brouwer, R., Hofkes, M., & Moeltner, K. (2011). The effect of risk context on the value of a statistical life: A Bayesian meta-model. *Environmental and Resource Economics*, 49, 597–624.

Delaney, L., & O'Toole, F. (2006). Willingness to pay: Individual or household? *Journal of Cultural Economics*, 30(4), 305–309.

Denne, T., Kerr, G., Glover, D., Winder, M., & Wright, L. (2018). *A pilot study to determine the relative value of non-market transport impacts of investment* (NZ Transport Agency research report 648). https://www.nzta.govt.nz/resources/research/reports/648/

Denne, T., & Wright, L. (2017). *A review of methods for analysis of regulatory effectiveness* (NZ Transport Agency research report 604). https://www.nzta.govt.nz/resources/research/reports/604

Department for Transport. (2021a). *TAG data book November 2021 v1.17*. https://www.gov.uk/guidance/transport-analysis-guidance-tag#webtag-data-book

Department for Transport. (2021b). *TAG unit A1.3: User and provider impacts. July 2021.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1079012/tag-unit-A1.3-user-and-provider-impacts.pdf

Desvouges, W., Mathews, K., & Train, K. (2016). From curious to pragmatically curious: Comment on "From hopeless to curious? Thoughts on Hausman's 'Dubious to Hopeless' critique of contingent valuation". *Applied Economic Perspectives and Policy*, *38*(1), 174–182.

Dockins, C., Maguire, K. B., Newbold, S., Simon, N. B., Krupnick, A., & Taylor, L. O. (2018). *What's in a name? A systematic search for alternatives to "VSL"* (No. 2168-2018-8178).

Douglas, N. J. (2016). *Pricing strategies for public transport Part 2: Literature review* (NZ Transport Agency research report 565). https://www.nzta.govt.nz/resources/research/reports/565

Douglas, N. J. (2017). *Valuing public transport quality using passenger ratings & willingness to pay surveys*. Douglas Economics.

Douglas, N. J. (2021). Valuing life when roads are increasingly safe. In *Australasian Transport Research Forum 2021 proceedings*. https://australasiantransportresearchforum.org.au/wp-content/uploads/2022/05/ATRF2021_Resubmission_126-1.pdf

Douglas, N. J., Wallis, I. P. (2013). Predicting the value of public transport in-vehicle time. In *Australasian Transport Research Forum 2013 proceedings*. https://www.researchgate.net/profile/Neil-Douglas/publication/309238209_Predicting_the_Value_of_Public_Transport_In-Vehicle_Time/links/580685dc08ae5ad18816654d/Predicting-the-Value-of-Public-Transport-In-Vehicle-Time.pdf

Fayyaz, M., Bliemer, M. C. J., Beck, M. J., Hess, S., & van Lint, W. C. (2021). Stated choices and simulated experiences: Differences in the value of travel time and reliability. *Transportation Research Part C: Emerging Technologies*, 128, 103145.

Federal Aviation Authority. (n.d.). *Treatment of the values of life and injury in economic analysis.* https://www.faa.gov/regulations_policies/policy_guidance/benefit_cost/media/econ-value-section-2-tx-values.pdf

Fiebig, D. G., Keane, M. P., Louviere, J., & Wasi, N. (2010). The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity. *Marketing Science*, *29*(3), 393–421.

Flügel, S., Elvik, R., Veisten, K., Rizzi, L. I., Meyer, S. F., Ramjerdi, F. & Ortúzar, J. de D. (2015). Asymmetric preferences for road safety: Evidence from a stated choice experiment among car drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 31*, 112–123.

Flügel, S., Veisten, K., Rizzi, L. I., Ortúzar, J. de D., & Elvik, R. (2019). A comparison of bus passengers' and car drivers' valuation of casualty risk reductions in their routes. *Accident Analysis and Prevention*, *122*, 63–75.

Fosgerau, M. (2016). *The valuation of travel time variability* (International Transport Forum discussion paper 2016-04).

Freeman, A. M., Herriges, J. A., & Kling, C. L. (2014). *The measurement of environmental and resource values* (3rd ed.). RFF Press.

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly, 73*(2), 349–360.

Glover, D., & Henderson, J. (2010). *Quantifying health impacts of government policies. A how-to guide to quantifying the health impacts of government policies*. Department of Health.

González, R. M., Amador, F. J., Rizzi, L. I., Ortúzar, J. de D., Román, C., Espino, R., Martín, J. C., & Cherchi, E. (2018). Estimating the value of risk reductions for car drivers when pedestrians are involved: A case study in Spain. *Transportation*, *45*(2), 499–521.

Groothuis, P. A., & Whitehead, J. C. (2009). The provision point mechanism and scenario rejection in contingent valuation. *Agricultural and Resource Economics Review*, *38*(2), 271–280.

Guria, J., Jones, W., Jones-Lee, M., Keall, M., Leung, J., & Loomes, G. (2003). *The New Zealand values of statistical life and of the prevention of injuries: Draft*. Land Transport Safety Authority.

Haab, T. C., Interis, M.G., Petrolia, D.R., & Whitehead, J. C. (2013). From hopeless to curious? Thoughts on Hausman's "Dubious to Hopeless" critique of contingent valuation. *Applied Economic Perspectives and Policy*, *35*(4), 593–612.

Haab, T. C., Interis, M. G., Petrolia, D. R., & Whitehead, J. C. (2016). Interesting questions worthy of further study: Our reply to Desvousges, Mathews, and Train's (2015) Comment on our thoughts (2013) on Hausman's (2012) Update of Diamond and Hausman's (1994) Critique of Contingent Valuation. *Applied Economic Perspectives and Policy*, *38*(1), 183–189.

Haddak, M. M. (2016). Estimating the willingness-to-pay for road safety improvement. *Transportation Research Procedia*, *14*, 293–302.

Haddak, M. M., Havet, N., & Lefèvre, M. (2014). *Willingness-to-pay for road safety improvement* (GATE Groupe d'Analyse et de Théorie Économique Lyon-St Étienne Working Paper WP1406).

Haghani, M., Bliemer, M. C. J., Rose, J. M., Oppewal, H., & Lancsar, E. (2021). Hypothetical bias in stated choice experiments: Part II. Conceptualisation of external validity, sources and explanations of bias and effectiveness of mitigation methods. *Journal of Choice Modelling, 41*, 100322.

Hanley, N., & Czajkowski, M. (2019). The role of stated preference valuation methods in understanding choices and informing policy. *Review of Environmental Economics and Policy, 13*(2), 248–266.

Hausman J. (2012). Contingent valuation: From dubious to hopeless. *Journal of Economic Perspectives*, *26*(4), 43–56.

Hensher, D. A. (2008). Joint estimation of process and outcome in choice experiments and implications for willingness to pay. *Journal of Transport Economics and Policy*, *42*, 297–322.

Hensher, D. A. (2015). Customer service quality and benchmarking in public transport contracts. *International Journal of Quality Innovation*, *1*(4), 1–17. http://dx.doi.org/10.1186/s40887-015-0003-9

Hensher, D. A., Beck, M. J., & Balbontin, C. (2021). What does the quantum of working from home do to the value of commuting time used in transport appraisal? *Transportation Research Part A: Policy and Practice*, *153*, 35–51.

Hensher, D. A., & Layton, D. (2010). Parameter transfer of common metric attributes in choice analysis: Implications for willingness to pay. *Transportation*, *37*, 473–490.

Hensher, D. A., & Prioni, P. (2002). A service quality index for area-wide contract performance assessment. *Journal of Transport Economics and Policy*, *36*(1) 93–113.

Hensher, D. A., Rose, J. M., & Greene, W. H. (2005). *Applied choice analysis: A primer*. Cambridge University Press.

Hensher, D. A., Rose, J. M., & Greene, W. H. (2015). *Applied choice analysis* (2nd ed.). Cambridge University Press.

Hensher, D. A., Rose, J. M., Ortúzar, J. de D., & Rizzi, L. I. (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.

Hensher, D. A., Rose, J. M., Ortúzar, J. de D., & Rizzi. L. I. (2011). Estimating the value of risk reduction for pedestrians in the road environment: An exploratory analysis. *Journal of Choice Modelling*, *4*(2), 70–94.

Hess, S., Murphy, P., Le, H., & Leong, W. Y. (2017). Estimation of new monetary valuations of travel time, quality of travel, and safety for Singapore. *Transportation Research Record*, *2664*, 79–90.

HM Treasury. (2003). *The green book: appraisal and evaluation in central government*.

Hodgson, R., & Birks, S. (2002). *Statistics New Zealand's definition of family, its implications for the accuracy of data and effectiveness of policy targeting.* Massey University Centre for Public Policy Evaluation, Student Paper no. 4.

Hojman, P., Ortúzar, J. de D., & Rizzi, L. (2005). On the joint valuation of averting fatal and severe injuries in highway accidents. *Journal of Safety Research*, *36*, 377–386.

Hojman, P., Ortúzar, J. de D., & Rizzi, L. (2011). *Internet-based surveys to elicit the value of risk reductions* [Paper presentation]. 9th International Conference on Transport Survey Methods, 14–18 November 2011, Termas de Puyehue, Chile.

Hossain, I., Saqib, N. U., & Haq, M. M. (2018). Scale heterogeneity in discrete choice experiment: An application of generalized mixed logit model in air travel choice. *Economics Letters*, *172*, 85–88.

Hoyos, D. (2010). The state of the art of environmental valuation with discrete choice experiments. *Ecological Economics, 69*, 1595–1603.

Hultkrantz, L., & Svensson, M. (2012). The value of a statistical life in Sweden: A review of the empirical literature. *Health Policy*, *108*(2–3), 302–310.

Ian Wallis Associates (IWA) Ltd. (2014). *Car passenger valuations of quantity and quality of time savings* (NZ Transport Agency research report 551). https://www.nzta.govt.nz/resources/research/reports/551/

International Energy Agency (IEA). (2007). *Mind the gap quantifying principal-agent problems in energy efficiency*.

Iragüen, P., & Ortúzar, J. de D. (2004). Willingness-to-pay for reducing fatal accident risk in urban areas: an internet-based Web page stated preference survey. *Accident Analysis and Prevention, 36*, 513–524.

Johannesson, M., Johansson, P.-O., & O'Conor R. M. (1996). The value of private safety versus the value of public safety. *Journal of Risk and Uncertainty*, *13*, 263–275.

Johnston, R. J., Boyle, K. J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T. A., Hanemann, W. M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., & Vossler, C. A. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, *4*(2), 319–405.

Jones-Lee, M. W. (1991). Altruism and the value of other people's safety. *Journal of Risk and Uncertainty*, *4*, 213–219.

Jones-Lee, M. W. (1992). Paternalistic altruism and the value of statistical life. *The Economic Journal*, *102*, 80–90

Jones-Lee, M. W., & Loomes, G. (1995). Scale and context effects in the valuation of transport safety. *Journal of Risk and Uncertainty*, *11*, 183–203.

Jones-Lee, M. W., & Loomes, G. (2015). Final response to Thomas and Vaughan. *Process Safety and Environmental Protection, 94*, 542–544.

Jones-Lee, M., Loomes, G., O'Reilly, D., & Philips, P. (1993). *The value of preventing non-fatal road injuries: Findings of a willingness-to-pay national sample survey* (Transport Research Laboratory contractor report 330). Department of Transport.

Jones-Lee, M. W., Loomes, G., & Philips, P. R. (1995). Valuing the prevention of non-fatal road injuries: Contingent valuation vs. standard gambles. *Oxford Economic Papers*, *47*(4), 676–695.

Kim, S., & Haab, T. C. (2004). *Revisiting bid design issues in contingent valuation* [Paper presentation]. American Agricultural Economics Association Annual Meeting, Denver, Colorado, 1–4 August 2004.

Kling, C. L., Phaneuf, D. J., & Zhao, J. (2012). From Exxon to BP: Has some number become better than no number? *Journal of Economic Perspectives*, *26*(4), 3–26.

Krupnick, A., Alberini, A., Belli, R., Cropper, M., & Simon, N. (1997). *New directions in mortality risk valuation and stated preference methods: Preliminary results*. Working Paper, Resources for the Future.

Leung, J. (2009). *Understanding transport costs and charges. Phase 2 – value of statistical life: A meta analysis. Is the current value of safety for New Zealand too low?* Ministry of Transport.

Le, H., van Geldermalsen, T., Lim, W., & Murphy, P. (2011). Deriving accident costs using willingness-to-pay approaches: A case study for Singapore. In *Australasian Transport Research Forum 2011 Proceedings*.

Lee, P. S. L. (2012). *The application of choice modelling to determine the values of non-market goods and services: a national park case study* [Doctoral dissertation, Massey University]. http://hdl.handle.net/10179/4040

Li, Z., Hensher, D. A., & Rose, J. (2010). *Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence* (Institute of Transport and Logistics Studies working paper ITLS-WP-10-05).

Lindhjem, H., & Navrud, S. (2009). Asking for individual or household willingness to pay for environmental goods? Implication for aggregate welfare measures. *Environmental and Resource Economics*, *43*, 11–29.

Lindhjem, H., Navrud, S., Braathen, N. A., & 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.

Litman, T. (2017). *Valuing transit service quality improvements: Considering comfort and convenience in transport project evaluation*. Victoria Transport Policy Institute.

Marglin, S. A. (1963). The social rate of discount and optimal rate of saving. *Quarterly Journal of Economics*, *77*(1), 95–111.

McDonald, R. L., Chilton, S. M., Jones-Lee, M., & Metcalf, H. (2016). Dread and latency impacts on a VSL for cancer risk reduction. *Journal of Risk and Uncertainty*, *52*(2), 37–161.

McFadden, D., & Train, K. (2017). *Contingent valuation of environmental goods: A comprehensive critique*. Edward Elgar.

Meade, R., & Cheung, L. (2016). *GPS user preferences exploratory work*. Report for the New Zealand Ministry of Transport.

Miller, T. R., & Guria, J. (1991). *The value of statistical life in New Zealand: Market research on road safety*. Ministry of Transport.

Milligan, C., Kopp, A., Dahdah, S., & Montufar, J. (2014). Value of a statistical life in road safety: A benefit-transfer function with risk-analysis guidance based on developing country data. *Accident Analysis and Prevention*, *71*, 236–247.

Ministry of Transport. (2021). *Social cost of road crashes and injuries – June 2020 update*. https://www.transport.govt.nz/assets/Uploads/Social-Cost-of-Road-Crashes-and-Injuries-2020_final.pdf

Mon, E. E., Jomnonkwao, S., Khampirat, B., Satiennam, W., & Ratanavaraha, V. (2018). Willingness to pay for mortality risk reduction for traffic accidents in Myanmar. *Accident Analysis and Prevention*, *118*, 18–28.

Mouter, N., Cabral, M. O., Dekker, T., & van Cranenburgh, S. (2019). The value of travel time, noise pollution, recreation and biodiversity: A social choice valuation perspective. *Research in Transportation Economics, 76*, 100733.

Mouter, N., & Chorus, C. (2016). Value of time – a citizen perspective. *Transportation Research Part A: Policy and Practice*, *91*, 317–329.

Mouter, N., Koster, P., & Dekker, T. (2019). *An introduction to participatory value evaluation* (Tinbergen Institute Discussion Paper TI 2019-024/V).

Mouter, N., van Cranenburgh, S., & van Wee, B. (2017). Do individuals have different preferences as consumer and citizen? The trade-off between travel time and safety. *Transportation Research Part A: Policy and Practice*, *106*, 333–349.

Mouter, N., van Cranenburgh, S., & van Wee, B. (2018). The consumer-citizen duality: Ten reasons why citizens prefer safety and drivers desire speed. *Accident Analysis & Prevention*, *121*, 53–63.

Munro, A. (2005). Household willingness to pay equals individual willingness to pay if and only if the household income pools. *Economics Letters*, *88*(2), 227–230.

New Zealand Productivity Commission. (2014). *Regulatory institutions and practices*.

Nielsen, J. S., Chilton, S., & Metcalf, H. (2019). Improving the risk-risk tradeoff method for use in safety appraisal responses. *Environmental Economics and Policy Studies*, *21*, 61–86.

Niroomand, N., & Jenkins, G. P. (2016). Estimating the value of life, injury, and travel time saved using a stated preference framework. *Accident Analysis and Prevention*, *91*, 216–225.

Niroomand, N., & Jenkins, G. P. (2017). Estimating the value of life and injury for pedestrians using a stated preference framework. *Journal of Safety Research*, *62*, 81–87.

Niroomand, N., & Jenkins, G. P. (2018). A comparison of stated preference methods for the valuation of improvement in road safety. *Economic Analysis and Policy*, *59*, 138–149.

Nisa, R. L. K. (2018). *Scrutinizing consumer-citizen duality towards safety and travel time trade-offs. A case study of on-demand motorbike taxi in Jakarta* [Master's thesis, Delft University of Technology]. http://resolver.tudelft.nl/uuid:53732fc5-0385-49b0-858f-16c3cdcc6b97

NZ Transport Agency. (2018). *Economic evaluation manual*.

Obermeyer, A., & Hirte, G. (2019). *Willingness to pay for road safety: A conceptual study and pilot survey for Germany* [Paper presentation]. 59th ERSA Congress, August 27–30, 2019, Lyon, France.

Olofsson, S., Gerdtham, U. G., Hultkrantz, L., & Persson, U. (2016). *Chained approach vs contingent valuation for estimating the value of risk reduction* (Lund University, Department of Economics Working Paper 2016:34).

Organisation for Economic Co-operation and Development (OECD). (2022). *PPPs and exchange rates*. OECD National Accounts Statistics (database). https://doi.org/10.1787/data-00004-en

Outwater, M. L., Spitz, G., Lobb, J., Campbell, M., Pendyala, R., Sana, B., & Woodford, W. (2010). *TCRP H-37 Characteristics of Premium Transit Services that affect Mode Choice: Summary of Phase 1.* Transit Cooperative Research Program.

Parumog, M., Mizokami, S., & Kakimoto, R. (2006). Value of traffic externalities from attribute-based stated choice. *Transportation Research Record*, *1954*, 52–60.

Penn, J., & Hu, W. (2018). Euthanizing value of a statistical life: Monetizing differences in public perceptions and alternatives. *Applied Economics*, *50*(16), 1824–1836.

Perry, B. (2007). Appendix 2 to *Household incomes in New Zealand: Trends in indicators of inequality and hardship 1982 to 2004*. Ministry of Social Development.

Perry, B. (2019). *Household incomes in New Zealand: Trends in indicators of inequality and hardship 1982 to 2018.* Ministry of Social Development.

Poe, G., Clark, J. E., Rondeau, D., & Schulze, W. D. (2002). Provision point mechanisms and field validity tests of contingent valuation. *Environmental and Resource Economics, 23*(1), 105–131.

Rizzi, L. I., & Ortúzar, J. de D. (2003). Stated preference in the valuation of interurban road safety. *Accident Analysis and Prevention*, *35*, 9–22.

Rizzi, L. I, & Ortúzar, J. de D. (2006a). Estimating the willingness-to-pay for road safety improvements. *Transport Reviews*, *26*(4), 471–485.

Rizzi, L. I., & Ortúzar, J. de D. (2006b). Road safety valuation under a stated choice framework. *Journal of Transport Economics and Policy, 40*(1), 69–94.

Robinson, L. A., & Hammitt, J. K. (2016). Valuing reductions in fatal illness risks: Implications of recent research. *Health Economics*, *25*, 1039–1052.

Robinson, L. A., Hammitt, J. K., & O'Keeffe, L. (2019). Valuing mortality risk reductions in global benefit-cost analysis. *Journal of Benefit Cost Analysis*, 10(Suppl 1), 15–50.

Rose, J. M., & Bliemer, M. C. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews*, *29*(5), 587–617.

Rose, J., Mulley, C., Tsai, C.-H., & Hensher, D. (2013). Will bus travellers walk further for a more frequent service? A stated preference investigation. In *Australasian Transport Research Forum 2013 Proceedings* 2. https://www.australasiantransportresearchforum.org.au/sites/default/files/2013_rose_mulley_tsai_hensher.pdf

Rouwendahl, J., de Blaeij, A., Rietveld, P., & Verhoef, E. (2010). The information content of a stated choice experiment: A new method and its application to the value of a statistical life. *Transportation Research Part B: Methodological*, *44*, 136–151.

Sagoff, M. (1988). *The economy of the Earth. Philosophy, law and the environment.* Cambridge University Press.

Sandorf, E. D. (2019). Did you miss something? Inattentive respondents in discrete choice experiments. *Environmental and Resource Economics*, *73*, 1197–1235.

Scarpa, R., & Rose, J. M. (2008). Designs efficiency for non-market valuation with choice modelling: How to measure it, what to report and why. *Australian Journal of Agricultural and Resource Economics 52*(3), 253–282.

Small, K. (2012). Valuation of travel time. *Economics of Transportation*, *1*, 2–14.

Small, K., Noland, R., & Koskenoja, P. (1995). *Socio-economic attributes and impacts of travel reliability: A stated preference approach* (California PATH research report UCB-ITS-PRR-95-36). Institute of Transportation Studies.

Small, K. A., Winston, C., & Yan, J. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, *73*(4), 1367–1382.

Simon, N. B., Dockins, C., Maguire, K. B., Newbold, S. C., Krupnick, A., & Taylor, L. O. (2019). Policy brief – What's in a name? A search for alternatives to "VSL". *Review of Environmental Economics and Policy*, *13*, 155–161.

Stats NZ. (2019). *Measuring child poverty: Equivalence scale*. https://www.stats.govt.nz/methods/measuring-child-poverty-equivalence-scale

Strand, J. (2007). Public-good valuation and intra-family allocation. *Environmental and Resource Economics*, *38*, 527–543.

Svensson, M., & 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*, 1205–1212.

Swait, J., & Adamowicz, W. (2001). The influence of task complexity on consumer choice: a latent class model of decision strategy switching. *Journal of Consumer Research*, *28*(1), 135–148.

Symonds Travers Morgan (NZ) Ltd. (1997). *Valuation of travel time savings* (Transfund New Zealand Research Report No 93).

Tait, P., Miller, S., Rutherford, P., & Abell, W. (2016). *Non-market valuation of improvements in freshwater quality for New Zealand residents, from changes in stock exclusion policy* (MPI Technical Paper No. 2017/08).

Thomas, P. J., & Vaughan, G. J. (2015a). Testing the validity of the 'value of a prevented fatality' (VPF) used to assess UK safety measures. *Process Safety and Environmental Protection*, *94*, 239–261.

Thomas, P. J., & Vaughan, G. J. (2015b). Reply to the letter to the editor. *Process Safety and Environmental Protection, 93*, 299–306.

Tilahun, N. Y., & Levinson, D. M. (2010). A moment of time: Reliability in route choice using stated preference. *Journal of Intelligent Transportation Systems 14*(3), 179–187.

Train, K. E. (2002). *Discrete choice methods with simulation (vol. 8)*. Cambridge University Press.

Transport for New South Wales. (2020). *Transport for NSW economic parameter values. Version: 2.0.* https://www.transport.nsw.gov.au/news-and-events/reports-and-publications/tfnsw-economic-parameter-values

United Nations Economic Commission for Europe. (2011). *Canberra Group handbook on household income statistics* (2nd ed.).

US Department of Transportation. (1997). *The value of saving travel time: Departmental guidance for conducting economic evaluations.*

US Department of Transportation. (2016). *The value of travel time savings: Departmental guidance for conducting economic evaluations – revision 2 (2016 update).*

US Department of Transportation. (2021a). *Departmental guidance on valuation of a statistical life in economic analysis*. https://www.transportation.gov/office-policy/transportation-policy/revised-departmental-guidance-on-valuation-of-a-statistical-life-in-economic-analysis

US Department of Transportation. (2021b). *Departmental guidance. Treatment of the value of preventing fatalities and injuries in preparing economic analyses. March 2021*. https://www.transportation.gov/resources/value-of-a-statistical-life-guidance

US Environmental Protection Agency. (2022). *Mortality risk valuation*. https://www.epa.gov/environmental-economics/mortality-risk-valuation

van Houtven, G., Sullivan, M. B., & Dockins, C. (2008). Cancer premiums and latency effects: A risk tradeoff approach for valuing reductions in fatal cancer risks. *Journal of Risk and Uncertainty*, *36*, 179–199.

Veisten, K., Flügel, S., Rizzi, L., Ortúzar, J. de D., & Elvik, R. (2013). Valuing casualty risk reductions from estimated baseline risk. *Research in Transport Economics*, *43*, 50–61.

Vincent, M. (2008). *Measurement valuation of public transport reliability* (Land Transport New Zealand Research Report 339). https://www.nzta.govt.nz/resources/research/reports/339

Viscusi, W. K. (2005). *The value of life* (John M Olin Center for Law, Economics and Business discussion paper no. 517).

Vrtic, M. Schuessler, N., Erath, A., & Axhausen, K. W. (2010). The impacts of road pricing on route and mode choice behaviour. *Journal of Choice Modelling*, *3*(1), 109–126.

Waka Kotahi NZ Transport Agency. (2021). *Monetised benefits and costs manual* (Version 1.4, February 2021).

Walker, J. L., Wang, Y., Thorhauge, M., & Ben-Akiva, M. (2018). D-efficient or deficient? A robustness analysis of stated choice experimental designs. *Theory and Decision*, *84*, 215–238.

Wallis, I., Rupp, K., & Alban, R. (2015). *Travel time saving assessment* (NZ Transport Agency research report 570). https://www.nzta.govt.nz/resources/research/reports/570

Wallis, I. P., & Wignall, D. R. (2012). *The benefits of public transport – option values and non-use values* (NZ Transport Agency research report 471). https://www.nzta.govt.nz/resources/research/reports/471

Wang, B., & Hensher, D. (2015). Working while travelling: What are implications for the value of travel time savings in the economic appraisal of transport projects? In *Australasian Transport Research Forum 2015 Proceedings*.

Wardman, M., Chintakayala, V. P. K., & De Jong, G. C. (2016). Values of travel time in Europe: Review and meta-analysis. *Transportation Research Part A: Policy and Practice*, *94*, 93–111.

Wijnen, W., & Stipdonk, H. (2016). Social costs of road crashes: An international analysis. *Accident Analysis and Prevention*, 94, 97–106.

Appendix A: Final survey

WAKA KOTAHI NZ TRANSPORT AGENCY Pg 1

EEM Parameters

Household Survey

Questionnaire Structure – Final questionnaire

1.12.20

WAKA KOTAHI NZ TRANSPORT AGENCY Pg 2

Introduction – Selected Respondent

Thank you for taking part, my name is __________from CBG.

This is a survey for Waka Kotahi NZ Transport Agency; a government agency that looks after the national transport system.

Your answers will help Waka Kotahi to understand people's transport preferences and the kind of transport system that people would value the most. This will help them with planning for the future.

All your answers are confidential and anonymous. You can stop the survey at anytime if you're not comfortable.

Please ask at any point if there is anything you don't understand, or would like me to explain further.

If more information is needed:

Click here

>>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 3

[Further information for interviewer to offer as required - If button clicked on intro. screen]

Who are Waka Kotahi? Waka Kotahi is a Crown entity established under the Land Transport Management Act 2003. The objective of the Agency is to undertake its functions in a way that contributes to an efficient, effective and safe land transport system in the public interest. Each year, the Transport Agency funds innovative and relevant research that contributes to this objective.

What is the research for? Waka Kotahi are trying to better understand the factors which influence travel behaviour and inform decisions about policy or infrastructure.

Who are CBG? CBG is an independent research company and we have been commissioned to conduct this research for Waka Kotahi.

What will happen to my Information? This is a large survey, with around 8,000 households across Aotearoa New Zealand being surveyed. The answers you give will be combined with those of others and analysed and reported together. Your information will be kept confidential. The researchers are governed by the Code of Conduct of the New Zealand Research Association.

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 4

Screeners (asked of all)

If quotas are full or the respondent does not meet criteria (i.e. does not use any of the modes required), they will be screened out here

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 5

These next questions are to help us understand what kind of journeys or trips you tend to make, including what kind of transport you use, what you travel for and how often you travel.

We are not interested in work related travel so please, exclude any trips you were making as part of your job (such as travelling during working hours or transporting goods or people) BUT we do want you to include any trips TO and FROM your place of work.

I'll explain further as we go along.

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 6

Q1a. Thinking about all trips and journeys you made in New Zealand in the last **month** (including all travel, from short distance trips of a few minutes, to long distance trips that could take several hours), have you made any trips using a…

This excludes air travel

Please select all that apply

- ☐ Private vehicle (car, ute, van etc) as a **driver**
- ☐ Private vehicle (car, ute, van etc) as a **passenger**
- ☐ Motorbike/scooter as a driver
- ☐ Bus
- ☐ Train
- ☐ Ferry
- ☐ Walking/running (to go somewhere or do something, not just for exercise)
- ☐ Bicycle (including e-bikes)
- ☐ Other (e-scooter, skateboard, taxi, uber etc) – please specify_____________

[If at least one of the following are not selected: (driver, passenger, motorbike, bus or train) – thank participant for their time and screen them out]

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 7

Now we are interested in how often you make certain types of trips or journeys.

Q1b Still thinking about the last **month**, how often have you made the following types of trips?
Please select one per row

	6-7 days a week	5 days a week	3-4 days a week	1-2 days a week	Once every 2-3 weeks	About once a month	Not in the last month
A regular commuting trip (such as regular trips to/from work/study/other regular activities that you needed to be on time for)	☐	☐	☐	☐	☐	☐	☐
A trip in your local area that is not time dependent (such as shopping, personal business, social visit, recreation)	☐	☐	☐	☐	☐	☐	☐
A trip in your local area that you needed to be on time for (such as a meeting/appointment/something with a specific start/end time)	☐	☐	☐	☐	☐	☐	☐

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 8

[Only display if trip type at least monthly at Q1. Skip if only one mode selected at Qf]

Q2a Thinking about all of the regular commuting trips you've made in the last month, what types of transport have you used?
(such as regular trips to/from work/study/other regular activities that you needed to be on time for)
Please select all that apply

[Only show modes selected at Qf]
- ☐ Private vehicle as a **driver** (car, ute, van etc)
- ☐ Private vehicle as a **passenger** (car, ute, van etc)
- ☐ Motorbike/scooter as a driver
- ☐ Bus
- ☐ Train
- ☐ Ferry
- ☐ Cycling (including e-bikes)
- ☐ Walking/running
- ☐ Other (e-scooter, skateboard etc) – please specify____________

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 9

[Only display if trip type at least monthly at Q1. Skip if only one mode selected at Qf]

Q2b Thinking about all of the local trips that are not time dependent you've made in the last month, what types of transport have you used?
(such as shopping, personal business, social visit, recreation)
Please select all that apply

[Only show modes selected at Qf]

- ☐ Private vehicle as a **driver** (car, ute, van etc)
- ☐ Private vehicle as a **passenger** (car, ute, van etc)
- ☐ Motorbike/scooter as a driver
- ☐ Bus
- ☐ Train
- ☐ Ferry
- ☐ Cycling (including e-bikes)
- ☐ Walking/running
- ☐ Other (e-scooter, skateboard etc) – please specify____________

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 10

[Only display if trip type at least monthly at Q1. Skip if only one mode selected at Qf]

Q2c Thinking about all of the local trips that you needed to be on time for that you've made in the last month, what types of transport have you used?
(such as a meeting/appointment/something with a specific start/end time)
Please select all that apply

[Only show modes selected at Qf]

- ☐ Private vehicle as a **driver** (car, ute, van etc)
- ☐ Private vehicle as a **passenger** (car, ute, van etc)
- ☐ Motorbike/scooter as a driver
- ☐ Bus
- ☐ Train
- ☐ Ferry
- ☐ Cycling (including e-bikes)
- ☐ Walking/running
- ☐ Other (e-scooter, skateboard etc) – please specify____________

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 11

One of the following modes/purpose combinations will be selected based on quotas per participant

	Driver/ motorcyclist	Passenger	Bus	Train
Commuter	1	5	9	12
Local time dependent	2	6	10	13
Local non-time dependent	3	7	11	14

[Only one mode/purpose combination will be selected per respondent. Each respondent will be presented with 5x time/frequency/cost/crowding/traffic games and 5x time/reliability/cost games for the one mode/purpose combination selected]

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 12

Recent trip

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 13

[Show for trip type selected manually by interviewer]

Q3a Thinking about the most recent trip you made as **[insert mode]** when you were **[Insert purpose]** ...
If you can't remember the last trip you made, please think about a recent trip of this type that you make
Note that this might be the return trip (e.g. back home from the shops) or an outbound trip (e.g. to the shops from home) but not both.

Where did you start the trip from? ____________________

What was the destination of the trip?__________________

How long did the trip take?__________________Hours____________________minutes

What was the purpose of the trip?____________________

[Only ask if mode is car driver/ passenger:]
Q3b How many people were in the vehicle, including yourself?
Type in. If travelling alone, type in "1"

____

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 14

[Only ask if Q3b>1. Only ask if mode is car driver/passenger]
Q3c What relationship were the other people in the vehicle to you?
Please select all that apply

- ☐ My child
- ☐ My parent
- ☐ Partner/spouse
- ☐ Other family member (grandparent, cousin, auntie/uncle etc.)
- ☐ Friend (non-family member)
- ☐ My employer
- ☐ My employee
- ☐ Co-worker
- ☐ Other (specify)

[Only ask if mode is car driver/motorbike:]
Q3d On this trip, who paid for your running costs (fuel, electricity, maintenance, road user charges etc.)?
Please select all that apply

- ☐ I did
- ☐ Someone else did
- ☐ My work did
- ☐ A combination (e.g. I shared the cost with work/a partner etc.)

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 15

For the following questions, I want you to imagine you need to take a trip - that is you can't decide not to go, or to go another time. Imagine the transport system you use has changed and you now have two routes, or trip options, to choose from to get to your destination. I will give you information for each route and will ask you to choose which you would prefer to take based on the information you are given.

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 16

Bus/train user section

(only show if bus/train selected at slide 14. Otherwise skip to car driver/passenger section)

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 17

Below is an example of what the screens will look like:

The following questions will be about a trip you are making **as a** [bus/train] **passenger for** [trip purpose]

		Route one	Route two
How often trips are scheduled	Service frequency	Every 15 minutes	Every 10 minutes
How long an average trip on this route takes	Average Travel time	35 minutes	45 minutes
Cost of a ticket	Cost (fare)	$3	$5
How full the bus or train is	Crowding	You will have to stand for the whole trip	You will be able to sit for the whole trip

[Note: the order that attributes are displayed in will be rotated within each respondent]

We would like you to consider what you think you would do if you did have this information in real life. Please assume that any other information you haven't been given is the same for both routes (e.g. the weather is the same for both routes and it is the same time of day)

It is important you read and consider everything before making your choice.

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

PT – Time/frequency/crowding/cost

Pg 18

Below is an example of what the games will look like. Five of these games will be presented per participant

QP1 Which of these two routes would you prefer to take **as a** [bus/train] **passenger**?
This is a trip you are making as a [insert trip type]

	Route one	Route two
Service frequency	Every 15 minutes	Every 10 minutes
Average Travel time	35 minutes	45 minutes
Cost (fare)	$3	$5
Crowding	You will have to stand for the whole trip	You will be able to sit for the whole trip
	Prefer route one	Prefer route two

QP1b Why did you choose that route?

WAKA KOTAHI NZ TRANSPORT AGENCY

PT – Time/frequency/crowding/cost

Pg 19

Below is an example of what the games will look like. Four more of these games will be presented per participant

QP2-5 Which of these two routes would you prefer to take **as a** [bus/train] **passenger**?
This is a trip you are making as a [insert trip type]

	Route one	Route two
Service frequency	Every 15 minutes	Every 10 minutes
Average Travel time	35 minutes	45 minutes
Cost (fare)	$3	$5
Crowding	You will have to stand for the whole trip	You will be able to sit for the whole trip
	Prefer route one	Prefer route two

WAKA KOTAHI NZ TRANSPORT AGENCY

PT – Time reliability

Pg 20

The following questions are also about a trip you are making **as a** [bus/train] **passenger for** [trip purpose]
This time, instead of information about frequency and crowding, you will be shown the variation in travel time over six different trips on the same route
Below is an example of what the screens will look like:

		Route one	Route two
How long ten different trips on the route would typically take	Average travel time	30 minutes	35 minutes
On average how long 6 different trips would take. These trip times occur randomly and could be on any day of the week	Different travel times for 6 trips	28 mins, 28 mins, 28 mins, 28 mins, 33 mins, 33 mins	30 mins, 30 mins, 35 mins, 37 mins, 37 mins, 39 mins
Includes fuel/running costs	Trip cost	$4	$3

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

PT – Time reliability

Pg 21

Below is an example of what the games will look like. Five of these games will be presented per participant

QP6 Which of these two routes would you prefer to take **as a** [bus/train] **passenger**?
This is a trip you are making as a [insert trip type]

	Route one	Route two
Average travel time	30 minutes	35 minutes
Different travel times for 6 trips	28 mins, 28 mins, 28 mins, 28 mins, 33 mins, 33 mins	30 mins, 30 mins, 35 mins, 37 mins, 37 mins, 39 mins
Trip cost	$4	$3
	Prefer route one	Prefer route two

QP6b Why did you choose that route?

WAKA KOTAHI NZ TRANSPORT AGENCY

PT – Time reliability

Pg 22

Below is an example of what the games will look like. Five of these games will be presented per participant

QP7-10 Which of these two routes would you prefer to take **as a** [bus/train] **passenger**?
This is a trip you are making as a [insert trip type]

	Route one	Route two
Average travel time	30 minutes	35 minutes
Different travel times for 6 trips	27, 28, 30, 30, 32, 33	20, 25, 30, 35, 40, 45
Trip cost	$4	$3
	Prefer route one	Prefer route two

WAKA KOTAHI NZ TRANSPORT AGENCY   Pg 23

Car driver/passenger section

(only show if driver/passenger selected at slide 14. If they've already answered the PT section, skip to social safety/risk section)

WAKA KOTAHI NZ TRANSPORT AGENCY   Pg 24

Car – Time/heavy traffic valuation

Below is an example of what the screens will look like:

The following questions will be about a trip you are making **as a** [driver/passenger] **in a private vehicle for** [trip purpose]**.**

		Route one	Route two
How long an average trip will take	Average travel time	45 minutes	30 minutes
Time spent moving slowly in heavy traffic	Heavy traffic	0 minutes Of your travel time is spent in heavy traffic	10 minutes Of your travel time is spent in heavy traffic
Includes fuel/running costs	Cost	$3	$5

[Note: the order that attributes are displayed in will be rotated within each respondent]

The information we give you may be more than you would usually have before making a trip, but we'd like you to consider what you think you would do if you did have this information in real life. Please assume that any other information you haven't been given is the same for both routes (e.g. the weather is the same for both routes and it is the same time of day)

It is important you read and consider everything before making your choice.

WAKA KOTAHI NZ TRANSPORT AGENCY

Car – Time/heavy traffic valuation

Pg 25

[If a driver, show]:

For the following, please imagine that you are...

- paying for the whole running cost yourself .
- [If car]: driving alone in the vehicle (with no passengers)
- [If motorbike]: alone on the motorbike (with no passengers)

[If a passenger, show]:

For the following, please imagine that...

- even though you are a passenger, you have some say in which route you choose
- you are paying for half (50%) of the travel costs. The cost we show is what you will have to pay (your share)
- it is only you and the driver in the vehicle

WAKA KOTAHI NZ TRANSPORT AGENCY

Car – Time/heavy traffic valuation

Pg 26

Below is an example of what the games will look like. Five of these games will be presented per participant

QC1 Which of these two routes would you prefer to take **as a** [driver/passenger]?
This is a trip you are making as a [insert trip type]

	Route one	Route two
Average travel time	**45 minutes**	**30 minutes**
Heavy traffic	**0 minutes** Of your travel time is spent in heavy traffic	**10 minutes** Of your travel time is spent in heavy traffic
Cost	$3	$5

Prefer route one

Prefer route two

QC1b Why did you choose that route?

Car – Time/heavy traffic valuation

Pg 27

Below is an example of what the games will look like. Five of these games will be presented per participant

QC2-5 Which of these two routes would you prefer to take **as a** [driver/passenger]?
This is a trip you are making as a [insert trip type]

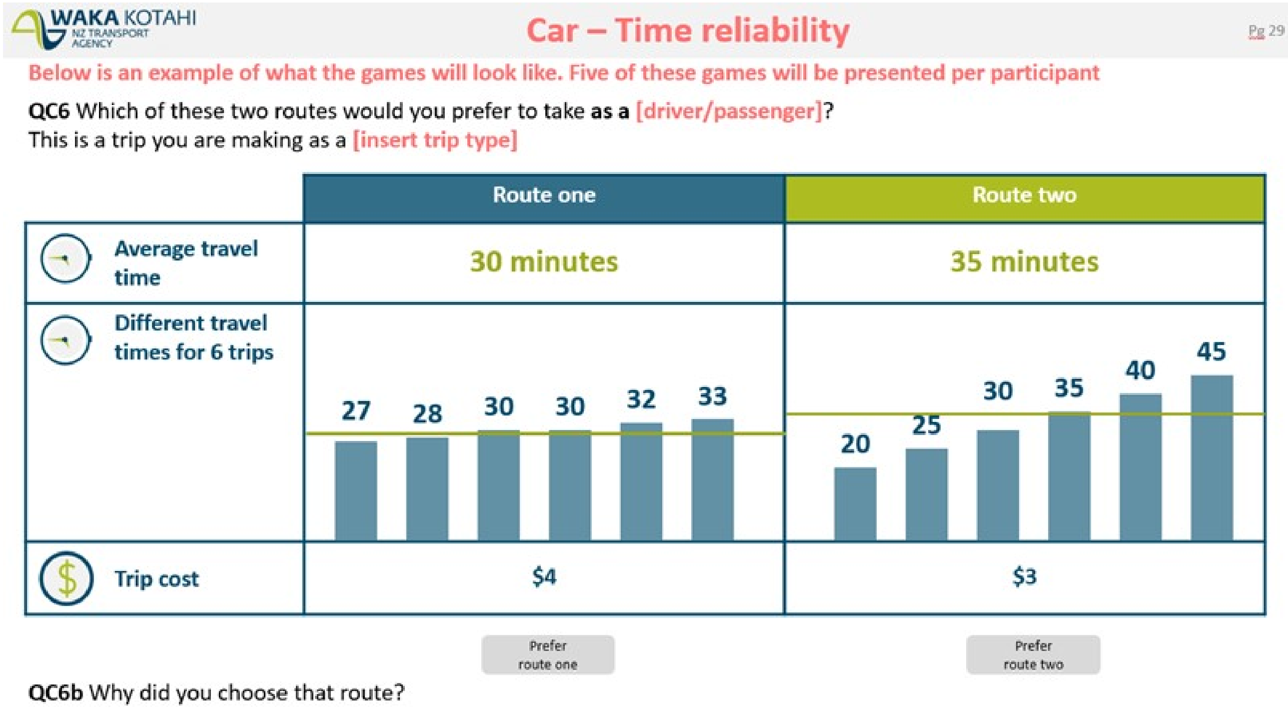
	Route one	Route two
Average travel time	45 minutes	30 minutes
Heavy traffic	0 minutes Of your travel time is spent in heavy traffic	10 minutes Of your travel time is spent in heavy traffic
Cost	$3	$5
	Prefer route one	Prefer route two

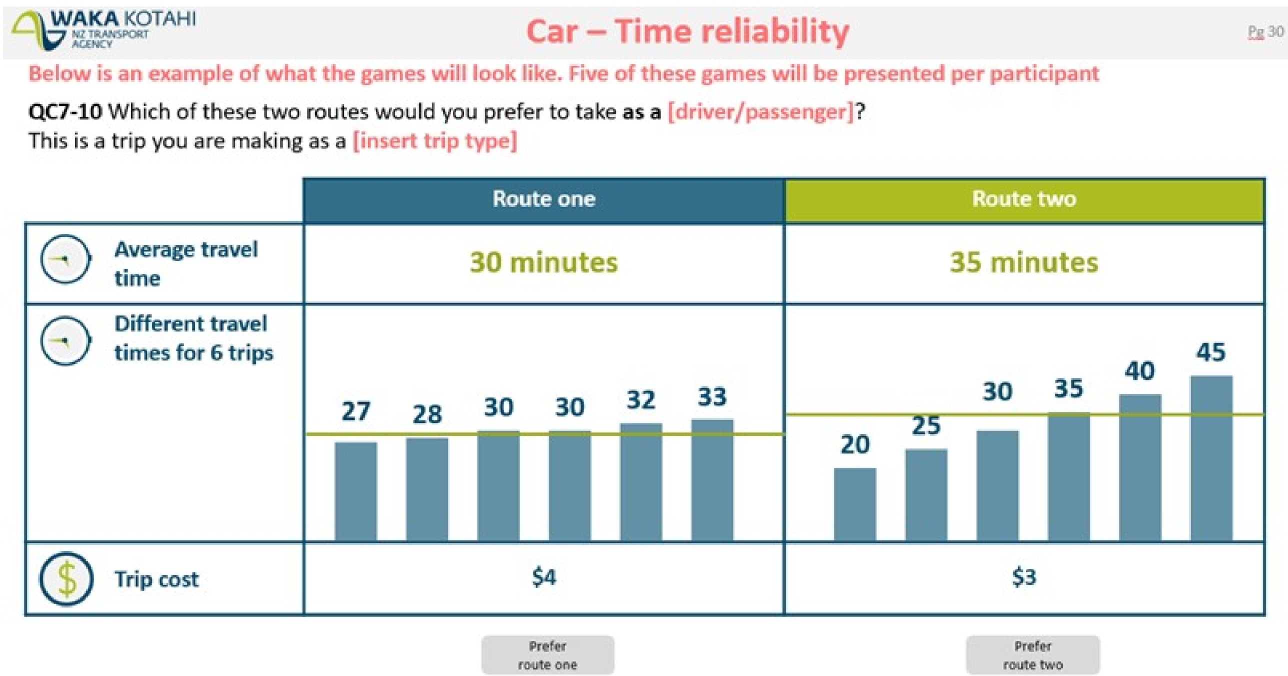
Car – Time reliability

Pg 28

The following questions are also about a trip you are making **as a** [driver/passenger] **in a private vehicle for** [Trip purpose]
This time, instead of time spent in heavy traffic, you will be shown the variation in travel time over six different trips on the same route
Below is an example of what the screens will look like:

[Note: the order that attributes are displayed in will be rotated within each respondent]


WAKA KOTAHI
NZ TRANSPORT AGENCY
Car – Time reliability
Pg 29
Below is an example of what the games will look like. Five of these games will be presented per participant
QC6 Which of these two routes would you prefer to take as a [driver/passenger]?
This is a trip you are making as a [insert trip type]
Route one
Route two
Average travel time
30 minutes
35 minutes
Different travel times for 6 trips
27
28
30
30
32
33
20
25
30
35
40
45
Trip cost
$4
$3
Prefer route one
Prefer route two
QC6b Why did you choose that route?


WAKA KOTAHI
NZ TRANSPORT AGENCY
Car – Time reliability
Pg 30
Below is an example of what the games will look like. Five of these games will be presented per participant
QC7-10 Which of these two routes would you prefer to take as a [driver/passenger]?
This is a trip you are making as a [insert trip type]
Route one
Route two
Average travel time
30 minutes
35 minutes
Different travel times for 6 trips
27
28
30
30
32
33
20
25
30
35
40
45
Trip cost
$4
$3
Prefer route one
Prefer route two

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 31

Social safety/risk section
(Asked of all)

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 32

Now we are going to ask you some questions about investment in road safety projects.

The government could improve safety on the road in a number of ways, including upgrading or building new roads, and changing road features like intersections or lights. However, making roads safer can sometimes have other trade-offs like higher costs or slower journey times.

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 33

For these next questions, I will give you some options the government could invest in and will ask you to choose which you would prefer.

Here's an example of what the questions will look like:

	Current (no changes made)	Investment option one	Investment option two
Deaths (per year)	250	200	250
Serious injuries (per year)	1750	1000	1500
Minor injuries (per year)	6000	5000	6000
Increase in your personal costs (per year)	$0	**$200** more per year	**$100** more per year

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 34

Before we continue, I have some more information about the choices:

Serious injuries

- injuries requiring medical attention or admission to hospital, including fractures, concussion and severe cuts.

Minor injuries

- Injuries other than serious, which require first aid or cause discomfort or pain, including bruising and sprains.

Cost

- This is a personal cost that you would need to pay every year in the future (not a one-off cost)
- Cost increases would apply to every adult in New Zealand
- Cost could be included in income taxes, rates (or increased rent if you live in a rental property), a fuel tax, road user charges or any other tax that you pay directly or indirectly
- This cost is the net amount you need to pay and takes into account any savings there may be to the healthcare system, emergency services or ACC etc. This means that even if costs are saved elsewhere, you will still need to pay the amount displayed on screen

Lastly, these upgrades would be made nationwide to multiple different roads each year, which would include roads in your region and in other regions around the country.

WAKA KOTAHI NZ TRANSPORT AGENCY Pg 35

The following questions are about a nation-wide investment programme

QS1 Which of the following would you prefer?

	Current (no changes made)	Investment option one	Investment option two
Deaths (per year)	250	200	250
Serious injuries (per year)	1750	1000	1500
Minor injuries (per year)	6000	5000	6000
Increase in your personal costs (per year)	$0	$200 more per year	$100 more per year
	Prefer this (no change)	Prefer option 1	Prefer option 2

QS1b Why did you choose that option?

WAKA KOTAHI NZ TRANSPORT AGENCY Pg 36

QS2-5 Which of the following would you prefer?

	Current (no changes made)	Investment option one	Investment option two
Deaths (per year)	250	200	250
Serious injuries (per year)	1750	1000	1500
Minor injuries (per year)	6000	5000	6000
Increase in your personal costs (per year)	$0	$200 more per year	$100 more per year
	Prefer this (no change)	Prefer option 1	Prefer option 2

WAKA KOTAHI NZ TRANSPORT AGENCY Pg 37

QS6 For this question, please rank the following in order of which you think is the highest priority for the government to spend money on when investing in road upgrades

	Reducing deaths (on the roads per year)	**Survey will display a drag-and-drop box where you place each in order from most important on the top to least at the bottom*
	Reducing serious injuries (on the roads per year)	
	Reducing heavy traffic	
	Reducing average travel times	
	Making trips more reliable (meaning trips on the same route are more likely to always take the same amount of time)	

[Note: the order of rows will be rotated between participants]

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY Pg 38

General question section (asked of all)

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 39

Thank you, that's the end of this section. I will now ask you some more general questions about you

Q4a I am now going to read out a list of different types of road crashes. Have you ever been personally involved in a crash where you were...

Please select all that apply

- ☐ In a car as a driver
- ☐ In a car as a passenger
- ☐ On a motorbike/scooter
- ☐ On a bicycle
- ☐ On a bus or train
- ☐ None of the above (I've never been involved in a road crash) [exclusive]

[Only ask if 'yes' to any accident types at Q4a]

Q4b Have you had any injuries as a result of the crashes you have been involved in?

Please select one only

- ☐ Yes
- ☐ No
- ☐ Don't know

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 40

[Only ask if 'yes' to Q4b]

Q4c Were your injuries...

Please select all that apply

- ☐ Minor injuries (with no ongoing affects e.g. cuts, bumps, whiplash)
- ☐ Serious injuries (with ongoing affects e.g. broken bones, psychological distress, major bleeding, paralysis)

Q4d Do you know anyone who has been involved in a crash which resulted in fatalities?

Please select one only

- ☐ Yes
- ☐ No
- ☐ Don't know

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY Pg 41

[Ask of all even if haven't driven a car/motorbike in the last month]

Q5 What type or types of driver's licence do you currently have?
Please select all that apply

☐ Car (full)
☐ Car (restricted)
☐ Car (learners)
☐ Car (overseas)

☐ Motorcycle (full)
☐ Motorcycle (restricted)
☐ Motorcycle (learners)
☐ Motorcycle (overseas)

☐ Heavy vehicle (full)
☐ Heavy vehicle (learners)

☐ No licence (exclusive)

Only allow a max of 1x car, 1x motorbike and 1x heavy vehicle licence type to be selected

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY Pg 42

Q6. Which of the following age groups do you belong to?
Please select one only

1. 16-24
2. 25-34
3. 35-44
4. 45-54
5. 55-64
6. 65+

Q7 Which of the following genders do you identify with?
Please select one only

1. Male
2. Female
3. Gender diverse

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 43

Q8. Which ethnic group or groups do you belong to?
Please select all that apply

- ☐ New Zealand European
- ☐ Māori
- ☐ Samoan
- ☐ Cook Island Maori
- ☐ Tongan
- ☐ Niuean
- ☐ Chinese
- ☐ Indian
- ☐ Other European (Australian, English etc.)
- ☐ Other Pacific Islander (Fijian/Tokelauan etc.)
- ☐ Korean
- ☐ Japanese
- ☐ Malaysian
- ☐ Vietnamese
- ☐ Filipino
- ☐ Other Asian
- ☐ Other (Please specify)____________________
- ☐ Don't know
- ☐ Choose not to answer

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY

Pg 44

Q9. What is your **household** annual income (before tax)?
Please select one only

- ☐ zero income
- ☐ $20,000 or less
- ☐ $20,001 - $30,000
- ☐ $30,001 - $50,000
- ☐ $50,001 - $70,000
- ☐ $70,001 - $100,000
- ☐ $100,001 - $150,000
- ☐ $150,001 or more
- ☐ Don't know
- ☐ Choose not to answer

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY Pg 45

Q10 What is the highest level of education you have achieved?

Please select one only

- ☐ High school (level 1-4 certificate)
- ☐ Diploma (level 5 and 6)
- ☐ Bachelors Degree (level 7)
- ☐ Postgraduate Degree (Honours, Masters or Doctorate)
- ☐ Other (specify)________________

Q11 Which of the following best describes you..

Please select one only

- ☐ Student (full time)
- ☐ Student (part time)
- ☐ Employed (full time)
- ☐ Employed (part time)
- ☐ Employed (casually)
- ☐ Not working for pay
- ☐ Homemaker (full time)
- ☐ Volunteer worker (regular)
- ☐ Retired
- ☐ Unemployed (and seeking work)
- ☐ Other (specify)________________

Qa [Don't ask – location recorded using GPS]

<< >>

WAKA KOTAHI NZ TRANSPORT AGENCY Pg 46

That's all the questions I have for you. Thank you for your time.

Pay incentive/koha

If you have any questions about the research, please contact xxx from CBG on xxx. Leave a card if needed.

Submit

Q12 Don't read – please select any games you feel the participant had comprehension issues with
This includes any games where attributes were consistently misunderstood or ignored entirely)

- ☐ **Car** - Individual route choices (time, traffic and cost)
- ☐ **Public transport –** Individual route choice games (frequency, time, cost, crowding)
- ☐ **Car** – Reliability route choices (time, time over 6 trips, cost)
- ☐ **Public transport** – Reliability route choice (time, time over 6 trips, cost)
- ☐ **Safety** – Investment choices (deaths, injuries, cost)
- ☐ All of the above
- ☐ None of the above (understood everything) [exclusive]

[Note that only the games asked of the participant will be displayed above]

Q13 Don't read – Explain below the issues with any games selected above:

Appendix B: Discrete choice models

Choice modelling begins with the common microeconomic assumption that consumers aim to derive the greatest possible satisfaction (utility) from their limited income or budget constraint (ie, utility maximisation). Accordingly, a respondent's selection within a choice set represents what maximises their utility. Using data on the alternative(s) chosen in a choice task, discrete choice models can be estimated. This involves the estimation of a series of regression-like equations that predict the utility the decision maker assigns to each of the alternatives. The utility is *latent* in the sense it is concealed to the analyst, who must model it indirectly based on a number of observables. The modelled utilities for the alternatives are then compared to produce a probability of each alternative being chosen, based on assumptions about the random component of utility. Discrete choice models involve the simultaneous estimation of several equations, up to the number of alternatives in the data. Unlike linear regression models, however, these equations do not directly predict the observed outcome, which in this case would be the observed choices. Rather, the 'regression-like' equations predict the latent utilities for each of the alternatives, which are then subsequently used to predict the choice outcomes.

Choice modelling draws on two other consumer behaviour theories to model decision making: the Lancasterian consumer theory and random utility theory (Lee, 2012).

- Lancasterian theory suggests consumers derive utility from the attributes of a good as opposed to the good as a whole. Hence, alternatives within a choice set are described in terms of their key attributes.
- Random utility theory proposes an individual's utility can be divided into an observable or measurable component (V_j) and a random, unobservable component (e_j). Assuming these components are additive, we can define utility (U_j) derived from a specific good (j) as:

$$U_j = V_j + e_j \qquad \text{(Equation B.1)}$$

The random element (e_j) of utility is included because evaluators cannot perfectly predict a person's utility, as it is unlikely we can observe or measure every characteristic of the individual, good, or situation that affects decision behaviour. The measurable element of utility (V_j) for good j is commonly referred to as the deterministic component, and can be explained by the attributes of any given alternative (ie, good):

$$V_j = \beta_{0j} + \beta_{1j}.X_{1j} + \beta_{2j}.X_{2j} + \ldots + \beta_{kj}.X_{kj} = \beta X \qquad \text{(Equation B.2)}$$

where:

V_j – represents the measurable/observable utility for alternative j

β_{0j} – represents the average influence of all observed factors on utility

β_{kj} – represents the effect of a unit change in attribute k to utility

X_{kj} – represents the level of attribute k for choice j

Under this definition of utility β_{kj}, the effect of attribute k on choosing option j is assumed to be the same for all respondents. Each alternative has a unique utility function; however, for model identification purposes, the constant term (β_{0j}) for one alternative must be normalised to zero.

If we could directly measure each respondent's utility for a given option (V_j), then we could carry out a standard regression to find the coefficients (β_{kj}) that best fit the observed choices. However, only the respondents' choices are recorded in the data, not V_j itself. Thus, choice modelling uses the actual choices made, and the attributes associated with the alternatives, to predict the probability of an individual choosing a particular alternative (Train, 2002). The probability of an individual selecting alternative j over all other alternatives (J) in a choice set can be denoted as:

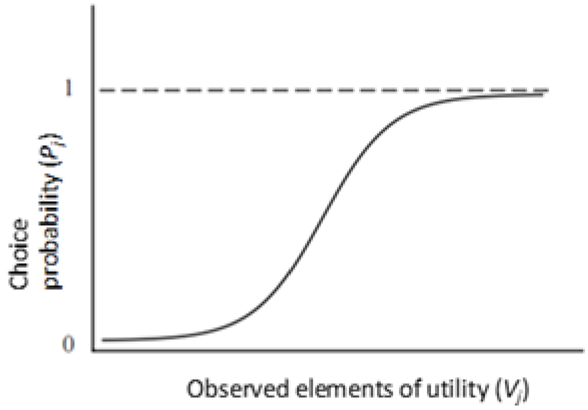
$$Prob_i = Prob\left[(V_j + e_j) \geq (V_J + e_J)\right] \quad \text{(Equation B.3)}$$

When predicting probabilities for separate (discrete) outcomes, such as those in a choice set, an appropriate regression model must be used. Various logit models are commonly employed for this purpose (Scarpa & Rose, 2008; Meade & Cheung, 2016).

B.1 Logit model

The multinomial logit (MNL) model converts the choice frequencies for a given alternative into choice probabilities via the logistic function. The logit probabilistic function is an ‘S’ shaped curve with the minimum and maximum limits of 0 and 1 used to model how the probability of an event may be affected by one or more explanatory variables (Train, 2002). In Figure B.1 we see how the probability of an alternative being chosen increases with the level of utility obtained from the choice.

Figure B.1 Logit model



Using the MNL model, the probability of a respondent selecting option j is the exponential of its observed elements of utility (V_j) divided by the sum of the exponential of observed utility for all options in the choice set (V_J).

$$Prob_j = \frac{exp^{v_j}}{\sum_J exp^{v_J}} \quad \text{(Equation B.4)}$$

B.2 Mixed logit model

The MNL model is limited by its assumption that taste or preference for a given attribute (represented by an attribute’s regression coefficient (β_{kj}) is the same for all respondents. Choices may be better explained by the mixed logit model, also known as the random parameters logit (RPL) model, a model that allows tastes to vary. The mixed logit model redefines the measured element of utility, now V_{ji}, to account for individuals’ (i) heterogeneous preferences for option j:

$$V_{ji} = \beta_{0ji} + \beta_{1ji}.X_{1j} + \beta_{2ji}.X_{2j} + \ldots + \beta_{kji}.X_{kj} = [\beta + \eta_i].X \quad \text{(Equation B.5)}$$

The vector η_i captures individual preference heterogeneity around the vector of mean preferences (β). A *distribution* of taste parameters is produced, rather than a single point estimate for a given attribute. Consequently, mixed logit WTP values are also distributions instead of point estimates. In short, the flexibility of the mixed logit model enables evaluation of how individuals differ in their trade-offs between attributes (Meade & Cheung, 2016).

B.3 Generalised mixed logit

The generalised mixed logit (GMXL) model adds flexibility to the RPL by allowing scale, which is a measure of consistency of choices, to vary across individuals. Following Fiebig et al. (2010), the observed part of the utility function in the GMXL model is:

$$V_{ji} = [\sigma_i \beta + \gamma \eta_i + (1 - \gamma)\, \sigma_i \eta_i].X_{ji} \qquad \text{(Equation B.6)}$$

In this specification, as in the RPL:

- η_i is individual taste heterogeneity
- σ_i is the scale parameter for individual i
- γ 'governs how the variance of residual taste heterogeneity varies with scale in a model that includes both' (Fiebig et al., p. 398) and is between 0 and 1.

Setting $\sigma_i = 1$, and $\gamma = 0$ or $\gamma = 1$, yields the RPL model.

$$V_{ji} = [\beta + \eta_i].X_{ij} \qquad \text{(Equation B.7)}$$

B.4 WTP-space models

In preference-space models the utility function is specified as $V_{ij} = \lambda_i p_{ij} + \beta_i' x_{ij}$, where the price attribute (p_{ij}) and its associated coefficient (λ_i) are separated from the other choice attributes. WTP-space models normalise λ_i to yield $V_{ij} = \lambda_i[p_{ij} + \theta_i' x_{ij}]$, where $\theta_i' = \beta_{ij}/\lambda_i$. Because WTP is defined as β_{ij}/λ_i, θ_i is a vector of WTP estimates for each of the attributes (Hensher et al., 2015).

B.5 NLOGIT specification

NLOGIT specifies the preference-space GMXL model as

$$\begin{aligned} U_{it}(j) &= \boldsymbol{\beta}_i \mathbf{x}_{it,j} + \varepsilon_{it,j} & & \\ \boldsymbol{\beta}_i &= \sigma_i \boldsymbol{\beta} + [\gamma + \sigma_i(1-\gamma)]\, \boldsymbol{\Gamma} \mathbf{w}_i & & 0 \le \gamma \le 1 \\ \sigma_i &= \exp(-\tau^2/2 + \tau \nu_i) & & \nu_i \sim N(0,1) \end{aligned} \qquad \text{(Equation B.8)}$$

Where

$\mathbf{w}_i$ specifies the distribution of random parameters

$\boldsymbol{\Gamma}$ is a variance/covariance matrix of heterogeneity in the attributes

σ_i is individual i's specific scale of the idiosyncratic error

τ is a scale heterogeneity factor.

Hence,

$$\boldsymbol{\beta}_i = \exp(-\tau^2/2 + \tau\nu_i)\boldsymbol{\beta} + [\gamma + \exp(-\tau^2/2 + \tau\nu_i)(1-\gamma)]\, \boldsymbol{\Gamma}\mathbf{w}_i \qquad \mathbf{w}_i \sim N[0,1],\ 0 \le \gamma \le 1 \qquad \text{(Equation B.9)}$$

The RPL and MNL models are nested within the GMXL model

$$\begin{aligned} &\text{RPL:} & \tau &= 0 & \boldsymbol{\beta}_i &= \boldsymbol{\beta} + \boldsymbol{\Gamma}\mathbf{w}_i & \tau &= 0 \rightarrow \gamma \text{ not identified} \\ &\text{MNL:} & \tau &= 0, \boldsymbol{\Gamma} = 0 & \boldsymbol{\beta}_i &= \boldsymbol{\beta} & & \end{aligned} \qquad \text{(Equation B.10)}$$

Appendix C: Detailed results

Table C.1 Safety model results

Model	RPL	RPL	RPL	RPL	GMXL	WTP-space GMXL
Distribution	**Triangular**	**Maximum**	**Lognormal**	**Rayleigh**	**Triangular**	**Triangular**
Random parameters in the utility function						
β_{death} (SE)	0.02399*** (0.00055)	0.01340*** (0.00104)	−4.05805*** (0.03486)	−2.51487*** (0.03684)	0.02494*** (0.00055)	4.16700*** (0.05072)
$\beta_{serious\ injury}$ (SE)	0.00127*** (3.624E-4)	0.00048*** (0.00015)	−6.88787*** (0.04947)	−5.18516*** (0.10419)	0.00146*** (0.3991E-4)	0.19305*** (0.00587)
$\beta_{minor\ injury}$ (SE)	0.00016*** (7.843E-5)	0.00012*** (0.8290E05)	−10.8072*** (0.57448)	−9.43071*** (0.38358)	0.00024*** (0.7735E-5)	0.01960*** (0.00128)
β_{SQ} (SE)	−1.95087*** (0.09347)	−1.97043*** (0.09357)	−1.53546*** (0.08935)	−1.98503*** (0.09360)	−1.04188*** (0.04042)	175.218*** (6.15214)
Non-random parameters in the utility function						
β_{cost} (SE)	−0.00538*** (0.00011)	−0.00434*** (0.00010)	−0.00537*** (0.00009)	−0.00432*** (0.00010)	−0.00581*** (0.8287E-4)	
Distributions of random parameters						
Death (SE)	0.02399*** (0.00055)	0.02482*** (0.00100)	1.34110*** (0.01157)	1.25947*** (0.06234)	0.02494*** (0.00055)	4.16700*** (0.05072)
Serious injury (SE)	0.00127*** (0.3624E-4)	0.00178*** (0.00021)	1.10915*** (0.05028)	1.62001*** (0.16022)	0.00146*** (0.3991E-4)	0.19305*** (0.00587)
Minor injury (SE)	.00016*** (0.7843E-5)	0.6920E-5 (0.4683E-4)	0.23864 (1.64440)	0.17507 (0.17161)	0.00024*** (0.7735E-5)	0.01960*** (0.00128)
SQ (SE)	4.43812*** (0.10368	4.13003*** (0.09310)	4.245605*** (0.09799)	4.07280*** (0.09118)	0.00800 (0.05246)	175.218*** (6.15214)
τ (SE)					1.30172*** (0.01129)	1.07443*** (0.00933)
γ (SE)					0.19302*** (0.04634)	
Log likelihood	−25,059.15	−24,718.73	−25,142.14	−24,734.57	−27,371.05	−28,969.90
k	6	9	9	9	8	7
Pseudo R^2	.226	.237	.224	.236	.155	.105
BIC	1.729	1.706	1.736	1.708	1.889	1.999

Notes: *, **, *** significant at 10%, 5%, 1% levels respectively. BIC = Bayesian information criterion. SE = standard error. SQ = status quo.

For each model there are 28,995 observations from 5,799 individuals.

Safety attribute coefficients are positive because the attributes were negatively coded to enable estimation of distributions defined only over positive values.

For triangular distributions, distribution parameters are identical to random parameters. This offers some economy in presenting subsequent models (SQ has a normal distribution, so the distribution parameter is different to the random parameter).

Table C.2 Public transport service model results

Journey purpose	Commute		Time-dependent		Flexible	
Journey duration	**Short**	**Long**	**Short**	**Long**	**Short**	**Long**
Random parameters in the utility function						
$\beta_{frequency}$ (SE)	0.13123*** (.00941)	0.16562*** (0.01658)	0.06331*** (0.01066)	0.07286*** (0.01173)	0.14160*** (0.01579)	0.11333*** (0.01775)
$\beta_{seated\ time}$ (SE)	0.16725*** (.01734)	0.09443*** (0.02442)	0.08493*** (0.01596)	0.11562*** (0.01924)	0.13112*** (0.02811)	0.11665*** (0.02945)
$\beta_{standing\ time}$ (SE)	0.20259*** (0.02355)	0.21602*** (0.03715)	0.13663*** (0.01700)	0.21762*** 0.02795)	0.20702*** (0.03628)	0.21443*** (0.05334)
β_{SQ} (SE)	−0.35506** (0.15788)	−0.18842 (0.34178)	−0.07500 (0.24125)	0.58686** (0.27732)	−0.02423 (0.26165)	0.045041 (0.37753)
Distributions of random parameters						
SQ (SE)	0.14834 (1.14845)	0.19711 (0.89032)	0.05474 (0.52656)	0.00787 (1.13858)	0.10203 (1.24921)	0.76438 (0.56412)
τ (SE)	1.22325*** (0.32386)	1.83832*** (0.32490)	1.66611*** (0.16861)	1.40432*** (0.31351)	1.56303*** (0.32757)	1.07617*** (0.36237)
Log likelihood	−668.63	−538.33	−589.82	−267.69	−478.86	−188.27
Restricted log likelihood	−791.35	−580.22	−672.89	−329.21	−527.83	−221.36
N	1,105	725	1,060	460	815	345
k	7	7	7	7	7	7
Pseudo R^2	.155	.072	.123	.187	.093	.150
BIC	1.229	1.513	1.133	1.204	1.200	1.143

Notes: *, **, *** significant at 10%, 5%, 1% levels respectively. BIC = Bayesian information criterion. SE = standard error. SQ = status quo.

Distributions: frequency, seated time, standing time – constrained triangular; SQ – normal.

Triangular distribution parameters are identical to random parameters.

Table C.2 (continued): Public transport service model results

Journey purpose	Commute	Time-dependent	Flexible	All
Journey duration	**All**	**All**	**All**	**All**
Random parameters in the utility function				
$\beta_{\text{frequency}}$ (SE)	0.13779*** (0.00802)	0.07709*** (0.00827)	0.13545*** (0.00995)	0.12367***. (0.00443)
$\beta_{\text{seated time}}$ (SE)	0.13563*** (0.01515)	0.09504*** (0.01417)	0.13053*** (0.01632)	0.12623*** (0.00832)
$\beta_{\text{standing time}}$ (SE)	0.19820*** (0.01995)	0.14146*** (0.01728)	0.21278*** (0.02125)	0.19664*** (0.01090)
β_{SQ} (SE)	−10.24660* (0.14558)	0.16422 (0.17639)	−0.07979 (0.17625)	−0.03018 (0.07713)
Distributions of random parameters				
SQ (SE)	0.07965 (0.97626)	0.01921 (0.97885)	0.38364 (0.50401)	0.17683 (0.31554)
τ (SE)	1.29524*** (0.15572)	1.55687*** (0.18315)	1.45668*** (0.17711)	1.38177*** (0.07826)
Log likelihood	−1,216.43	−865.30	−669.05	−2,764.16
Restricted log likelihood	−1,373.04	−1,002.51	−750.00	−3,126.08
N	1,830	1,520	1,160	4,510
k	7	7	7	7
Pseudo R^2	.114	.137	.108	.116
BIC	1.342	1.153	1.172	1.231

Notes: *, **, *** significant at 10%, 5%, 1% levels respectively. BIC = Bayesian information criterion. SE = standard error. SQ = status quo.

Distributions: frequency, seated time, standing time – constrained triangular; SQ – normal.

Triangular distribution parameters are identical to random parameters.

Table C.3 Public transport reliability model results

Journey purpose	Commute		Time-dependent		Flexible	
Journey duration	Short	Long	Short	Long	Short	Long
Random parameters in the utility function						
$\beta_{\text{mean travel time}}$ (SE)	0.26162*** (0.01733)	0.22185*** (0.02919)	0.12667*** (0.02480)	0.10971*** (0.03749)	0.51651*** (0.05834)	0.28824*** (0.03511)
$\beta_{\text{standard deviation of time}}$ (SE)	0.25262*** (0.02374)	0.27204*** (0.03523)	0.19412*** (0.02300)	0.20496*** (0.03326)	0.26031*** (0.05080)	0.30032*** (0.043640
β_{SQ} (SE)	0.08048 (0.11559)	−0.53887** (0.23563)	0.00376 (0.15662)	0.15568 (0.24523)	−0.03773 (0.27653)	0.38629 (0.25030)
Distributions of random parameters						
SQ (SE)	0.13159 (0.53374)	0.81832** (0.38770)	0.00712 (1.17047)	0.17555 (0.96231)	0.10604 (1.34474)	1.33912*** (0.46552)
τ (SE)	1.41267*** (0.20165)	1.42325** (0.66363)	1.62664*** (0.16199)	1.46317*** (0.49937)	1.59961*** (0.57130)	1.03501** (0.52538)
Log likelihood	−719.079	−539.991	−618.916	−292.563	−503.401	−199.658
Restricted log likelihood	−792.724	−572.403	−670.485	−329.478	−527.716	−221.662
N	1,105	725	1,060	460	815	345
K	6	6	6	6	6	6
Pseudo R^2	.093	.057	.077	.112	.046	.099
BIC	1.318	1.513	1.185	1.307	1.257	1.202

Notes: *, **, *** significant at 10%, 5%, 1% levels respectively. BIC = Bayesian information criterion. SE = standard error. SQ = status quo.

Distributions: mean travel time, standard deviation of travel time – constrained triangular; SQ – normal.

Triangular distribution parameters are identical to random parameters.

Table C.3 (continued): Public transport reliability model results

Journey purpose	Commute	Time-dependent	Flexible	All
Journey duration	**All**	**All**	**All**	**All**
Random parameters in the utility function				
$\beta_{\text{mean travel time}}$ (SE)	0.27493*** (0.01113)	0.09383*** (0.02043)	0.43128*** (0.03563)	0.28245*** (0.00691)
$\beta_{\text{standard deviation of time}}$ (SE)	0.24414*** (0.01356)	0.20362*** (0.01916)	0.38278*** (0.05507)	0.22637*** (0.00872)
β_{SQ} (SE)	−0.02035 (0.07692)	0.13398 (0.13345)	−0.06665 (0.22898)	−0.00772 (0.05028)
Distributions of random parameters				
SQ (SE)	0.01313 (0.58556)	0.07225 (0.66872)	0.86304 (0.55444)	0.07614 (0.23844)
τ (SE)	1.24640*** (0.18615)	1.82032*** (0.16562)	1.61584*** (0.33052)	1.32446*** (0.12751)
Log likelihood	−1,261.59	−913.04	−713.13	−2,914.46
Restricted log likelihood	−1,369.16	−1,001.91	−749.91	−3,121.98
N	1,830	1,520	1,160	4,510
k	6	6	6	6
Pseudo R^2	.079	.089	.049	.066
BIC	1.389	1.214	1.245	1.297

Notes: *, **, *** significant at 10%, 5%, 1% levels respectively. BIC = Bayesian information criterion. SE = standard error. SQ = status quo.

Distributions: mean travel time, standard deviation of travel time – constrained triangular; SQ – normal.

Triangular distribution parameters are identical to random parameters.

Table C.4 Private transport service model results

Journey purpose	Commute			Time-dependent			Flexible		
Journey duration	**Short**	**Medium**	**Long**	**Short**	**Medium**	**Long**	**Short**	**Medium**	**Long**
Random parameters in the utility function									
$\beta_{\text{free-flow time}}$ (SE)	0.64481*** (0.01815)	0.62255*** (0.06971)	0.51044*** (0.03158)	0.51691*** (0.01282)	0.67883*** (0.08170)	0.49304*** (0.03917)	0.68313*** (0.01985)	0.62486*** (0.06335)	0.57189*** 90.04798)
$\beta_{\text{heavy traffic time}}$ (SE)	1.22885*** (0.03517)	0.99431*** (0.10990)	0.92998*** (0.05403)	1.03842*** (0.02807)	0.97684*** (0.10613)	0.96211*** (0.07530)	1.24707*** (0.04103)	0.90373*** (0.08692)	0.93400*** (0.06900)
β_{SQ} (SE)	−0.10723 (0.12348)	−0.84455 (0.72318)	−0.23555 (0.65692)	−0.05864 (0.10343)	0.52401 (0.71897)	−1.00141 (0.80516)	0.07127 (0.13208)	0.36899 (0.62154)	−2.57875*** (0.78293)
Distributions of random parameters									
SQ (SE)	0.02212 (0.27240)	1.33931 (3.36310)	0.99934 (2.64118)	0.96792*** (0.18136)	0.75409 (1.84309)	1.00663 (3.34501)	0.97011*** (0.19821)	0.30108 (3.39313)	1.00016 (3.69597)
τ (SE)	1.22636*** (0.02725)	0.64815 (0.47035)	0.99655*** (0.24687)	1.23227*** (0.02079)	1.49600*** (0.23566)	1.25133*** (0.26851)	1.26970*** (0.03569)	1.19424** (0.53146)	1.00543*** (0.25292)
Log likelihood	−4,184.22	−675.18	−894.20	−5,401.18	−522.35	−824.83	−4,024.98	−546.66	−631.089
Restricted log likelihood	−4,967.66	−715.92	−988.36	−6,480.67	−578.14	−888.56	−4,646.51	−603.17	−702.02
N	6,930	970	1,320	9,415	815	1,275	7,125	860	960
k	6	6	6	6	6	6	6	6	6
Pseudo R^2	.158	.057	.095	.167	.096	.072	.134	.094	.101
BIC	1.211	1.411	1.369	1.150	1.303	1.308	1.133	1.292	1.333

Notes: *, **, *** significant at 10%, 5%, 1% levels respectively. BIC = Bayesian information criterion. SE = standard error. SQ = status quo.

Distributions: free-flow time, heavy traffic time – constrained triangular; SQ – normal.

Triangular distribution parameters are identical to random parameters.

Table C.4 (continued): Private transport service model results

Journey purpose	Commute	Time-dependent	Flexible	All
Journey duration	**All**	**All**	**All**	**All**
Random parameters in the utility function				
$\beta_{\text{free-flow time}}$ (SE)	0.51534*** (0.01059)	0.45805*** (0.00809)	0.59972*** (0.01494)	0.51982*** (0.00633)
$\beta_{\text{heavy traffic time}}$ (SE)	0.95448*** (0.02028)	0.86838*** (0.01692)	1.00462*** (0.02727)	0.95677*** (0.01242)
β_{SQ} (SE)	–	–	−0.33418** (0.15233)	–
Non-random parameters in the utility function				
β_{SQ} (SE)	0.08847*** (0.02550)	0.02045 (0.02322)	–	0.02749* (0.01415)
Distributions of random parameters				
SQ (SE)	–	–	0.27863 (0.63606)	–
τ (SE)	1.04936*** (0.02389)	1.22306*** (0.02610)	1.12161*** (0.03329)	1.06543*** (0.01578)
Log likelihood	−5,798.74	−6,785.31	−5,234.45	−17,875.57
Restricted log likelihood	−6,674.28	−7,947.96	−5,952.79	−20,575.24
N	9,220	11,505	8,945	29,670
k	5	5	6	5
Pseudo R^2	.131	.146	.121	.131
BIC	1.260	1.181	1.173	1.206

Notes: *, **, *** significant at 10%, 5%, 1% levels respectively. BIC = Bayesian information criterion. SE = standard error. SQ = status quo.

Distributions: free-flow time, heavy traffic time – constrained triangular; SQ – normal.

Triangular distribution parameters are identical to random parameters.

Table C.5 Private transport reliability model results

Journey purpose	Commute			Time-dependent			Flexible		
Journey duration	**Short**	**Medium**	**Long**	**Short**	**Medium**	**Long**	**Short**	**Medium**	**Long**
Random parameters in the utility function									
$\beta_{\text{mean travel time}}$ (SE)	0.52567*** (0.01073)	0.81004*** (0.085290)	0.070885*** (0.05415)	0.58732*** (0.00763)	0.46528*** (0.02299)	0.50687*** (0.02630)	0.65679*** (0.01224)	0.63368*** (0.04629)	0.69593*** (0.06639)
$\beta_{\text{standard deviation of time}}$ (SE)	0.41244*** (0.01751)	0.52682*** (0.07549)	0.34282*** (0.04891)	0.31548*** (0.01152)	0.33433*** (0.05089)	0.17560*** (0.04271)	0.43704*** (0.01915)	0.28843*** (0.06089)	0.00803 (0.11130)
β_{SQ} (SE)	−0.36710 (0.07995)	−0.13346 (0.64919)	0.74757 (0.98805)	–	−0.15830 (0.31419)	1.34300** (0.64505)	−0.33092*** (0.08391)	0.27563 (0.38897)	−2.09332*** (0.88829)
Non-random parameters in the utility function									
β_{SQ} (SE)	–	–	–	0.08321*** (0.02900)	–	–	–	–	–
Distributions of random parameters									
SQ (SE)	0.00206 (0.34157)	0.46457 (3.23508)	0.31686 (7.09007)	–	0.10550 (1.31128)	0.97347 (2.33364)	0.09880 (0.31645)	0.86751 (1.02612)	0.98695 (3.62028)
τ (SE)	0.95095*** (0.07160)	1.35485*** (0.30506)	1.20583*** (0.24115)	0.96882*** (0.04097)	1.16900*** (0.28839)	1.59929*** (0.25229)	0.98149*** (0.06844)	1.41767*** (0.27183)	1.36011*** (0.30559)
Log likelihood	−4,579.38	−669.67	−927.25	−5,791.34	−521.59	−838.21	−4,228.32	−546.95	−650.78
Restricted log likelihood	−4,965.80	−720.69	−989.70	−6,470.3	−578.00	−892.17	−4,638.85	−604.13	−704.75
N	6,950	975	1,320	9,455	815	1,280	7,140	865	970
k	6	6	6	5	6	6	6	6	6
Pseudo R^2	.078	.071	.063	.105	.098	.060	.088	.095	.077
BIC	1.321	1.392	1.419	1.227	1.301	1.324	1.188	1.285	1.360

Notes: *, **, *** significant at 10%, 5%, 1% levels respectively. BIC = Bayesian information criterion. SE = standard error. SQ = status quo.

Distributions: mean travel time, standard deviation of travel time – constrained triangular; SQ – normal.

Triangular distribution parameters are identical to random parameters.

Table C.5 (continued): Private transport reliability model results

Journey purpose	All	Commute	Time-dependent	Flexible	Flexible	Flexible	Flexible
Journey duration	**All**	**All**	**All**	**All**	**All**	**All**	**All**
Random parameters in the utility function							
$\beta_{\text{mean travel time}}$ (SE)	0.69077*** (0.00649)	0.63983*** (0.01149)	0.58420*** (0.00690)	0.66374*** (0.00025)	0.66973*** (0.00143)	0.67207*** (0.01262)	0.63423*** (0.00557)
$\beta_{\text{standard deviation of time}}$ (SE)	0.39974*** (0.00693)	0.44211*** (0.01546)	0.35055*** (0.00923)	0.49964*** (0.00033)	0.49253*** (0.00170)	0.45772*** (0.01486)	0.46074*** (0.00614)
β_{SQ} (SE)	–	–	–	−0.63829*** (0.00229)	−0.06461*** (0.00319)	−1.19420*** (0.61800)	3.04820*** (0.53072)
Non-random parameters in the utility function							
β_{SQ} (SE)	0.18028*** (0.01509)	0.16082*** (0.02643)	0.02021 (0.02538)	–	–	–	–
Distributions of status quo random parameters							
Distribution type	na	na	na	Triangular	Gamma	Weibull	Lognormal
SQ (SE)	–	–	–	−0.63829*** (0.00229)	Fixed	0.36869 (0.33486)	0.99310 (1.27649)
τ (SE)	0.95666*** (0.01897)	0.93110*** (0.04123)	0.78055*** (0.02606)	2.04558*** (0.00147)	2.15601*** (0.00769)	0.81347*** (0.04324)	0.81814*** (0.02090)
Log likelihood	−19,098.70	−6,277.93	−7,271.91	−5,421.32	−5,436.04	−5,500.58	−5,587.18
Restricted log likelihood	−20,582.99	−6,681.13	−7,948.94	−5,952.39	−5,952.39	−5,952.39	−5,952.39
N	29,770	9,245	11,550	8,975	8,975	8,975	8,975
k	5	5	5	5	5	6	6
Pseudo R^2	.072	.060	.085	.089	.087	.076	.061
BIC	1.284	1.360	1.261	1.210	1.214	1.228	1.248

Notes: *, **, *** significant at 10%, 5%, 1% levels respectively. BIC = Bayesian information criterion. SE = standard error. SQ = status quo.

Distributions: mean travel time, standard deviation of travel time – constrained triangular; SQ – normal.

Triangular distribution parameters are identical to random parameters.