Values of time and reliability in passenger and freight transport in The Netherlands

Report for the Ministry of Infrastructure and the Environment

SIGNIFICANCE, VU UNIVERSITY AMSTERDAM, JOHN BATES SERVICES
In collaboration with TNO, NEA, TNS NIPO and PanelClix

Project 08064, 1 November 2013

Revised final version

significance
quantitative research

**John
Bates**

Contents

Management Samenvatting vii
Executive Summary xiii
Acknowledgements xix

CHAPTER 1 **Introduction** **1**
1.1 Study objective 1
1.2 Overview of the methodology 2
1.3 Contents of this report 3

CHAPTER 2 **Theoretical models for values of time and reliability** **5**
2.1 Mean-dispersion models versus scheduling models 5
  2.1.1 Base models without variability 5
  2.1.2 Mean-dispersion models 6
  2.1.3 Scheduling models 6
  2.1.4 Combined mean-dispersion / scheduling models 7
2.2 Absolute versus relative models 7
2.3 Models in preference space and in WTP space 8
2.4 Models in logWTP space 9
2.5 Prospect theory 9
2.6 Advanced MNL models with diminishing sensitivity for higher base levels and small changes 10
  2.6.1 Introducing diminishing sensitivity for higher base levels 10
  2.6.2 Allowing different values for small / large changes 11
  2.6.3 Including a dispersion variable 12
2.7 Multinomial Logit models versus Mixed Logit models 13
2.8 Continuous distribution mixed logit models versus Latent Class models 13
2.9 Cross-sectional models versus panel models 14
2.10 Non-parametric techniques 14

FREIGHT TRANSPORT **15**

CHAPTER 3 **Freight transport questionnaire** **17**
3.1 A priori expectations for the VOT of shippers and carriers 17
3.2 Set-up of the questionnaire 19

3.3 SP experiments for shippers and carriers (excl. carriers using sea and inland waterways) 19
3.3.1 Description of the experiments 20
3.3.2 Description of the attributes 22
3.4 SP experiments for carriers using sea and inland waterways 23

CHAPTER 4 **Freight transport data** **25**
4.1 Number of respondents 25
4.2 Data selection and data quality 26
4.2.1 Treatment of the dominant alternative 26
4.2.2 Outliers 27
4.2.3 Transport characteristics 29
4.2.4 Trading 30
4.3 RP data 30

CHAPTER 5 **Freight transport data analysis** **33**
5.1 Initial model tests 33
5.1.1 Initial model specifications 33
5.1.2 Initial MNL results 34
5.1.3 Results for mean/median transport time and for mean/median scheduling terms 35
5.1.4 Results for relative models 36
5.1.5 Results for logWTP models and prospect theory 37
5.1.6 Mixed logit results 38
5.1.7 Jack-knife procedure 38
5.1.8 Results from the RP data 38
5.1.9 Differences between containerised and non-containerised transport 39
5.2 Selecting the final model specification 39
5.3 Synthesis per segment 41
5.3.1 Carriers-Road (including own account shippers) 41
5.3.2 Carriers-Rail transport 44
5.3.3 Carriers-air transport 45
5.3.4 Carriers-inland waterways transport 45
5.3.5 Carriers-Sea transport 47
5.3.6 Shippers that contract out 47
5.4 Final VOT results 49
5.4.1 The new VOTs 49
5.4.2 Discussion on the use of factor costs and trade-off ratios 52
5.4.3 Comparison against the 2003/2004 VOTs 54
5.4.4 Comparison against the international literature 56
5.5 Final VOR results 56
5.5.1 The new VORs 56
5.5.2 Comparison against previous values 58

5.6 Validation of freight transport VOT/VOR outcomes against practice 60

PASSENGER TRANSPORT 61

CHAPTER 6 **Passenger transport questionnaire 63**
6.1 Set-up of the questionnaire 63
6.2 SP experiments for passenger transport (excluding recreational navigation) 64
    6.2.1 Description of the experiments...64
    6.2.2 Description of the attributes in private car transport............................65
    6.2.3 Description of the attributes in public transport 66
    6.2.4 Description of the attributes in air transport..67
6.3 SP experiments for recreational navigation 68
    6.3.1 Description of the SP experiments...68
    6.3.2 Description of the attributes in recreational navigation........................69
6.4 Dominant questions...69

CHAPTER 7 **Passenger transport data 71**
7.1 The 2009 data (Panelclix panel) 71
    7.1.1 Recruitment 71
    7.1.2 Number of respondents...73
7.2 The 2011 data (en-route recruitment) 74
    7.2.1 Recruitment 74
    7.2.2 Number of respondents...76
    7.2.3 Additional questions..77
7.3 Data selection and data quality...79
    7.3.1 Treatment of the dominant alternative 79
    7.3.2 Outliers..80
    7.3.3 Trip characteristics 83
    7.3.4 Trading..84
7.4 RP data 84

CHAPTER 8 **Passenger transport data analysis 87**
8.1 Results for MNL mean-dispersion models...87
8.2 Results for advanced MNL mean-dispersion models 89
8.3 Results for advanced MNL scheduling models 93
    8.3.1 Scheduling models in WTP space..93
    8.3.2 Scheduling models in logWTP space...94
8.4 Results for advanced MNL mean-dispersion models with socio-economic interactions 95
8.5 Additional tests 99
    8.5.1 Differences between experiments...99
    8.5.2 Member of internet panel..100
    8.5.3 Different choices in November 2009...101

8.5.4 Type of unreliability 101
8.5.5 Number of transfers 102
8.5.6 Impact of exclusions 102
8.5.7 Gains and losses 103
8.6 Mixed logit results 104
8.6.1 Introduction 104
8.6.2 Continuous distribution panel mixed logit 104
8.6.3 Discrete distribution panel mixed logit – fixed points of the distribution, but estimated class probabilities 105
8.6.4 Discrete distribution panel mixed logit – estimated points and class probabilities of the distribution 105
8.7 Discussion on the use of MNL or panel-LC mixed logit models 109
8.8 Results for the RP data 112
8.9 Expansion of the survey outcomes for passengers using OViN 112
8.10 The employers' component in the value of time for business travel 113
8.11 Final VOT and VOR results after expansion using OViN 116
8.12 Comparison with the previous Dutch national VOT survey 118
8.12.1 Comparison with current CBA values 118
8.12.2 Re-analysis of 1997 data using new methodology 118
8.13 Comparison with the existing VOT literature 121
8.14 Comparison to the existing VOR literature 124

CHAPTER 9 **Further use of SP data in transport forecasting models 127**
9.1 The P-side and Q-side of reliability 127
9.2 The current transport models 128
9.3 The use of SP experiments in freight and passenger transport forecasting models and in Cost-Benefit Analysis 128

CHAPTER 10 **Summary and conclusions 131**

REFERENCES 137
Reference List 139

APPENDICES 147
Appendix A: Design tables 149
Appendix B: Re-analysis of the 1997 and 2009/2011 data for passenger transport 163

Management Samenvatting

Achtergrond en doelstelstelling

In Nederland worden voorstellen voor grote infrastructuurprojecten beoordeeld met behulp van een kosten-baten analyse. Hierbij worden de effecten van het project voorspeld door verkeers- en vervoersmodellen.

Vaak zijn de tijdwinsten in personen- en goederenvervoer de grootste batenpost. Voor de omzetting van reistijd- en transporttijdwinsten in geldeenheden bestaan er officiële reistijdwaarderingen ('values of time'). In het personenverkeer en –vervoer zijn deze gebaseerd op gegevens uit in 1997 gehouden enquêtes onder reizigers. De huidige waarden in het goederenvervoer zijn gebaseerd op gegevens uit enquêtes onder vervoerders en verladers in 2003/2004.

Een verder potentieel belangrijke batenpost zou kunnen bestaan uit de verbetering van de betrouwbaarheid van de reis- en transporttijden. Als deze tijden voorspelbaarder worden, dan kunnen reizigers en goederen makkelijker op het gewenste tijdstip op hun bestemming aankomen. Dit betekent dat er bij vertrek een kleinere buffer aangehouden kan worden om toch op tijd aan te komen

Op dit moment hebben we in Nederland uitsluitend voorlopige waarden om betrouwbaarheidswinsten om te zetten in geldeenheden. Deze betrouwbaarheids-waarderingen ('values of reliability') zijn gebaseerd op vuistregels, niet op empirisch materiaal uit waarderingsstudies. Het is echter belangrijk om accurate waarden voor betrouwbaarheid te hebben, omdat zonder zulke waarden de baten van diverse transport-, milieu- en verkeersveiligheidsprojecten onderschat kunnen worden.

Het doel van dit project was om de officiële waarden voor tijd in zowel het personenverkeer en- vervoer als het goederenvervoer te actualiseren, en om voor het eerst officiële waarden te bepalen voor betrouwbaarheid gebaseerd op empirisch onderzoek. Als maatstaf voor onbetrouwbaarheid gebruiken we de standaardafwijking van de reistijd, omdat andere mogelijke maatstaven aanzienlijker moeilijker in te bouwen zijn in de nationale en regionale transportmodellen.

Stated Preference onderzoek

Deze studie begon met het ontwerpen van vragenlijsten voor reizigers, vervoerders en verladers. Deze bevatten ieder drie 'stated preference' (SP) experimenten. In deze experimenten worden respondenten gevraagd om te kiezen tussen steeds twee hypothetische alternatieven voor een door hen gemaakte reis of een door of voor hen uitgevoerd transport. De hypothetische alternatieven worden daarbij beschreven in termen van reistijd, reiskosten en betrouwbaarheid. Voor de presentatie van betrouwbaarheid aan de respondenten wordt niet de standaardafwijking gebruikt, omdat deze indicator niet goed zou worden begrepen door de respondenten. In plaats daarvan presenteren we vijf reistijden voor ieder keuze—alternatief. Deze vijf reistijden zijn allemaal even waarschijnlijk.

In alle drie de SP experimenten wordt de respondenten gevraagd om afwegingen te maken tussen verbeteringen/verslechteringen in reistijden en -kosten, en in een deel van de experimenten ook tussen veranderingen in de betrouwbaarheid van reistijd. Voor respondenten namens goederenvervoersbedrijven in de binnenvaart en de zeevaart wordt een andere en innovatieve context voort het experiment gebruikt om tot voor deze sectoren relevante en realistische afwegingen te komen. Het betreft hier een afweging tussen wachttijden (voor een sluis of brug, of voor in- en uitladen aan de havenkade), betrouwbaarheid van de wachttijden en transportkosten.

In het goederenvervoer zijn er 812 succesvol afgeronde interviews verkregen met vervoerders en verladers, gebruik makend van interviewers die de respondenten ondervroegen middels een vragenlijst op een laptop (CAPI: 'computer-assisted personal interviewing').

In het personenverkeer en –vervoer zijn er in 2009 5.760 interviews verkregen met behulp van een bestaand internetpanel van respondenten. Initiële modellen die op deze data werden geschat lieten veel kleinere reistijdwaarderingen zien dan de waarden uit 1997 (na correctie voor inflatie en reële inkomensstijging). Deze verschillen konden niet afdoende verklaard worden door verschillen in de sociaal-economische samenstelling van de steekproef, de kenmerken van de gemaakte verplaatsingen en het ontwerp van de SP experimenten. Dit deed vermoeden dat de lagere reistijdwaarderingen te wijten waren aan de verschillende methoden die in 1997 en 2009 werden gebruikt om respondenten te werven.

Om deze reden zijn in 2011 nieuwe data verzameld, waarbij dezelfde methode voor het werven van respondenten werd gebruikt als in 1988 en 1997. Toen werden reizigers op benzinestations, parkeerplaatsen langs de snelweg en in de stad, treinstations en bushaltes gevraagd om deel te namen aan het onderzoek. In 1988 en 1997 werd een papieren vragenlijst gestuurd naar het adres dat de respondent had opgegeven. In 2011 is dat vervangen door een weblink die naar het emailadres van de respondent werd gestuurd. Het aantal succesvol afgeronde interviews dat op deze wijze is verzameld in 2011 is 1.430. Modeluitkomsten voor de reistijdwaarderingen op basis van de 2011 data bleken beter te sporen met de waarden uit 1997 dan die op basis van de 2009 data, en zijn in het vervolg van het onderzoek als leidend gebruikt bij het bepalen van de aanbevolen waarden. De respondenten in het internetpanel zijn namelijk minder represenatief voor de gemiddelde reiziger dan respondenten die in 2011 zijn geworven.

Uitkomsten voor het goederenvervoer

Voor het goederenvervoer zijn discrete keuzemodellen geschat op de SP gegevens. Voor het niet-wegvervoer gebruiken we daarbij relatieve modellen, waarbij reistijd, reiskosten en betrouwbaarheid worden uitgedrukt als fractie van de opgegeven waarden voor het daadwerkelijk uitgevoerde transport. Om absolute waarden in geldeenheden voor tijd en betrouwbaarheid te krijgen worden additionele data over transportkosten per uur (de zogenaamde 'factorkosten') gebruikt, die RWS-DVS beschikbaar heeft gesteld. De uitkomsten voor de waardering van transporttijd staan in Tabel E1.

De nieuwe waarden voor weg en spoor zijn van dezelfde orde van grootte als de waarden uit 2003/2004 (na correctie voor inflatie) en sporen ook met de internationale literatuur. Voor binnenvaart en zeevaart verkrijgen we nu hogere en plausibelere waarden dan in 2003/2004.

In Tabel E2 staan de uitkomsten voor betrouwbaarheid in het goederenvervoer. De nieuwe waarden komen duidelijk lager uit dan de eerdere voorlopige waarden voor Nederland, die gebaseerd waren op vele veronderstellingen die destijds niet empirisch getest konden worden. Aan de andere kant komen de nieuwe waarden voor betrouwbaarheid veel dichter in de buurt bij recente empirische waarden in het Verenigd Koninkrijk en Noorwegen.

Tabel E1: Transporttijdwaarderingen voor goederenvervoer (Euro/uur per voertuig of vaartuig, prijsniveau 2010)

	Weg	Spoor	Lucht	Binnenvaart	Zee
Container	*[2-40t truck]:* 59	*[complete trein]:* 880	*Niet van toepassing*	*[schip wachtend voor havenkade]:* 98 *[schip wachtend voor sluis/brug]:* 340	*[schip wachtend voor havenkade]:* 760
Niet-container	*[2-15t truck]:* 23 *[15-40t truck]:* 44 *[totaal niet-container]:* 37	*[bulk]:* 1200 *[wagenlading trein]:* 1100 *[totaal niet-container]:* 1200	*[compleet vrachtvliegtuig]:* 13000	*[schip wachtend voor havenkade]:* 65 *[schip wachtend voor sluis/brug]:* 300	*[schip wachtend voor havenkade]:* 830
Totaal	*[2-40t truck]:* 38	*[complete trein]:* 1100	*[compleet vrachtvliegtuig]:* 13000	*[schip wachtend voor havenkade]* 69 *[schip wachtend voor sluis/brug]:* 300	*[schip wachtend voor havenkade]:* 820

Voetnoten:
- Al deze waarden zijn de samengevoegde waarderingen van verladers en vervoerders, na afronding.
- De waarden voor spoor betreffen een trein (niet een wagen).
- De waarden voor binnenvaart en zee betreffen een schip.
- Alle waarden zijn exclusief BTW.

Tabel E2: Betrouwbaarheidswaarderingen voor goederenvervoer (Euro per uur standaardafwijking, per voertuig of vaartuig, prijsniveau 2010)

	Weg	Spoor	Lucht	Binnenvaart	Zee
Container	*[2-40t truck]:* 4	*[complete trein]:* 100	*Niet van toepassing*	*[schip wachtend voor havenkade]:* 18 *[schip wachtend voor sluis/brug]:* 27	*[schip wachtend voor havenkade]:* *45*
Niet-container	*[2-15t truck]:* 34 *[15-40t truck]:* 6 *[totaal niet-container]:* 15	*[bulk]:* 260 *[wagenlading trein]:* 240 *[totaal niet-container]:* 250	*[compleet vrachtvliegtuig]:* 1600	*[schip wachtend voor havenkade]:* 25 *[schip wachtend voor sluis/brug]:* 25	*[schip wachtend voor havenkade]:* *110*
Totaal	*[2-40t truck]:* 14	*[complete trein]:* 200	*[compleet vrachtvliegtuig]:* 1600	*[schip wachtend voor havenkade]* 24 *[schip wachtend voor sluis/brug]:* 26	*[schip wachtend voor havenkade]:* *100*

Voetnoten:
- Al deze waarden zijn de samengevoegde waarderingen van verladers en vervoerders, na afronding.
- De waarden voor spoor betreffen een trein (niet een wagen).
- De waarden voor binnenvaart en zee betreffen een schip.
- Alle waarden zijn exclusief BTW.

Uitkomsten voor het personenverkeer en -vervoer

Voor het personenverkeer en –vervoer hebben we geavanceerde discrete keuzemodellen (de zogenaamde ‘panel Latent Class’ modellen) geschat, waarbij de reistijdwaardering afhangt van de opgegeven reistijd en reiskosten van de werkelijke reis, van de verschillen in tijd en kosten die in de SP werden aangeboden en van kenmerken van de respondent (zoals opleiding, inkomen, leeftijd en samenstelling van het huishouden). We houden hierbij ook rekening met niet-waargenomen verschillen in de reistijdwaardering van personen en het feit dat onze SP steekproef een panelkarakter heeft (oftewel deze steekproef heeft meerdere observaties per respondent).

De nieuwe aanbevolen reistijdwaarderingen zijn berekend door weging van de steekproef van respondenten, en wel zodanig dat de verdeling van de reistijden in de verplaatsingen in het nationale mobiliteitsonderzoek OViN 2010 wordt weerspiegeld. De aldus verkregen waarden voor tijd en betrouwbaarheid staan in respectievelijk tabel E3 en E4, de betrouwbaarheidsratio’s (verhouding tussen betrouwbaarheidswaardering en reistijdwaardering) staan in tabel E5.

De nieuwe reistijdwaarderingen zijn doorgaans niet heel verschillend van de waarden die nu in KBA worden gebruikt in Nederland (gebaseerd op het SP onderzoek van 1997). Over het algemeen sporen ze ook goed met de uitkomsten van een internationale meta-analyse en de recente VOT studies in Zweden en Noorwegen.

Tabel E3: Reistijdwaarderingen voor personenverkeer en -vervoer (Euro/uur, prijsniveau 2010)

	Auto	Trein	Bus, tram, metro	Alle vervoer over land	Luchtvaart	Pleziervaart
Woon-werk	9,25	11,50	7,75	9,75		
Zakelijk werknemer	*12,75*	*15,50*	*10,50*	*13,50*	*85,75*	
Zakelijk werkgever	*13,50*	*4,25*	*8,50*	*10,50*	*-*	
Zakelijk	26,25	19,75	19,00	24,00	85,75	
Overig	7,50	7,00	6,00	7,00	47,00	8,25
Alle motieven	9,00	9,25	6,75	8,75	51,75	8,25

Voetnoot:
- Alle waarden zijn afgerond naar het dichstbijzijnde veelvoud van € 0,25.
- Alle waarden zijn inclusief BTW.

Tabel E4: Betrouwbaarheidswaarderingen voor personenverkeer en -vervoer (Euro/uur, prijsniveau 2010)

	Auto	Trein	Bus, tram, metro	Alle vervoer over land	Luchtvaart	Pleziervaart
Woon-werk	3.75	4.75	3.25	4.00		
Zakelijk werknemer	*14,50*	*18,00*	*12,00*	*15,50*	*56,00*	
Zakelijk werkgever	*15,50*	*4,75*	*9,75*	*12,25*	*-*	
Zakelijk	30,00	22,75	21,75	27,75	56,00	
Overig	4,75	4,50	3,75	4,50	30,75	0
Alle motieven	5,75	5,50	3,75	5,25	33,75	0

Voetnoot:
- Alle waarden zijn afgerond naar het dichstbijzijnde veelvoud van € 0,25.
- Alle waarden zijn inclusief BTW.

Tabel E5: Betrouwbaarheidsratios voor personenverkeer en -vervoer (Euro/uur, prijsniveau 2010)

	Auto	Trein	Bus, tram, metro	Alle vervoer over land	Luchtvaart	Pleziervaart
Woon-werk	0,4	0,4	0,4	0,4		
Zakelijk	1,1	1,1	1,1	1,1	0,7	
Overig	0,6	0,6	0,6	0,6	0,7	0

Voetnoot:
- Alle waarden zijn afgerond naar het dichstbijzijnde veelvoud van 0,1

Executive Summary

Background and aim of the study

In the Netherlands, proposals for large transport infrastructure projects are evaluated using a cost-benefit analysis. In such an analysis, impacts of the project are predicted using transport models and, if needed and possible, converted into monetary units.

For many projects, the time savings in passenger and freight transport are the most important benefit. For the conversion of travel and transport time savings, official values of time are available. For passenger transport, the current values are based on a survey among travellers carried out in 1997; for freight transport, the current values are based on a survey among shippers and carriers carried out in 2003/2004.

Another potentially important benefit of a project could be the improvement in travel time reliability. If travel and transport times would become more predictable, travellers and agents in freight transport would find it easier to arrive at the destination at the preferred moment and therefore reduce their safety margins in departure time. At the moment, in the Netherlands we only have provisional values to convert reliability gains into monetary units. These reliability values are based on rules of thumb and not supported by empirical evidence obtained from valuation studies. However, it is very important to have accurate values for reliability, since the benefits of some transport investments –and potentially also of some environmental and safety investments– may otherwise be underestimated.

This project was carried out to update the official values of time in both passenger and freight transport in the Netherlands and to deliver the first values of reliability based on an empirical foundation. We use the standard deviation of travel time as the measure for reliability, since all other possible measures of reliability would be much harder to incorporate in national and regional transport models.

SP surveys

Questionnaires have been designed for interviewing travellers, shippers and carriers. These interviews contain three stated preference (SP) experiments. In these experiments, respondents are asked to choose between two hypothetical alternatives for a trip or transport they actually made. The hypothetical alternatives are described in terms of travel time, travel costs and reliability. The reliability is not presented in the form of the standard deviation, because such an indicator is not well understood by respondents, but in the form of five travel times, which are all equally likely to happen.

In all stated preference experiments, respondents were asked to trade between improvements and deteriorations of travel time and travel cost, and in some experiments also between changes in reliability and arrival time. Only for respondents from firms transporting goods by inland waterways or by sea, we used a different and innovative choice context. Since for these respondents the standard choice context was not realistic, they were asked to trade between waiting times (for a lock or bridge or to be loaded or unloaded at a quay), reliability of these waiting times and the total transport costs.

For freight transport, 812 interviews were successfully carried out in 2010 with shippers and carriers, using computer-assisted face-to-face interviewing.

For passenger transport, 5,760 interviews were collected in 2009 using an existing internet panel. Initial models estimated on these data showed values of time that were much lower than the inflation- and income-corrected 1997 values. These differences could not be fully explained by differences in the socio-economic composition of the sample, attributes of the trips or differences in the design of the SP experiments. It was therefore thought that the lower values of time were caused by the different way in which respondents were recruited in 2009 compared to 1997.

Therefore, additional data was collected in 2011 using the same method of recruiting respondents as used in 1988 en 1997. In those years travellers at petrol stations/service areas, parking garages, railway stations and bus stops were asked to participate in a survey. In 1988 and 1997 a paper-based questionnaire was sent by mail to the address provided by the respondent. In 2011 a web link to the internet questionnaire was sent to the respondent's email address. In this way, 1430 interviews were successfully collected. Model results for the VOT based on the 2011 data were clearly more in line with the 1997 values than those for 2009. The 2011 data have therefore been treated as leading in the derivation of the final recommended values. The respondents in the internet panel are, presumably, less representative for the average travel than those who were recruited in the 2011 survey.

Results for freight transport

For freight transport, discrete choice models were estimated on the SP data. For the non-road models we use relative models, in which the attributes are measured relative to the observed levels. To obtain absolute money values of time and reliability from these models, additional data on the transport costs per hour (the so-called 'factor costs') are required. These values were provided by RWS-DVS. The resulting VOTs are shown in Table E1.

The new values of time for road and rail are of the same order of magnitude as the (inflation-corrected) values from 2003/2004 and are also compatible with the international literature. For inland waterway transport and sea transport we now obtain higher and more plausible values of time than in 2003/2004.

Table E2 presents the results for the freight transport value of reliability. The new values of reliability for freight transport are clearly lower than the provisional Dutch values, which were based on many assumptions that could not be tested at the time, However, the new values are much closer to recent empirical findings in the UK and Norway.

Table E1: Values of time for freight transport (Euro/hour per vehicle or vessel, price level 2010)

	Road	Rail	Air	Inland waterways	Sea
Container	*[2-40t truck]:* 59	*[full train]:* 880	*Not applicable*	*[ship waiting for a quay]:* 98 *[ship waiting for a lock/bridge]:* 340	*[ship waiting for a quay]:* 760
Non-container	*[2-15t truck]:* 23 *[15-40t truck]:* 44 *[all non-container]:* 37	*[bulk]:* 1200 *[wagonload train]:* 1100 *[all non-container]:* 1200	*[full freighter aircraft]:* 13000	*[ship waiting for a quay]:* 65 *[ship waiting for a lock/bridge]:* 300	*[ship waiting for a quay]:* 830
All	*[2-40t truck]:* 38	*[full train]:* 1100	*[full freighter aircraft]:* 13000	*[ship waiting for a quay]:* 69 *[ship waiting for a lock/bridge]:* 300	*[ship waiting for a quay]:* 820

Notes:
- All these values are combined values from shippers and carriers and were obtained after rounding off.
- The values for rail are for a train (not a wagon).
- The values for inland waterways and sea refer to a ship.
- These values do not include VAT.

Table E2: Values of reliability for freight transport (Euro/hour per vehicle or vessel, price level 2010)

	Road	Rail	Air	Inland waterways	Sea
Container	*[2-40t truck]:* 4	*[full train]:* 100	*Not applicable*	*[ship waiting for a quay]:* 18 *[ship waiting for a lock/bridge]:* 27	*[ship waiting for a quay]:* 45
Non-container	*[2-15t truck]:* 34 *[15-40t truck]:* 6 *[all non-container]:* 15	*[bulk]:* 260 *[wagonload train]:* 240 *[all non-container]:* 250	*[full freighter aircraft]:* 1600	*[ship waiting for a quay]:* 25 *[ship waiting for a lock/bridge]:* 25	*[ship waiting for a quay]:* 110
All	*[2-40t truck]:* 14	*[full train]:* 200	*[full freighter aircraft]:* 1600	*[ship waiting for a quay]:* 24 *[ship waiting for a lock/bridge]:* 26	*[ship waiting for a quay]:* 100

Notes:
- All these values are combined values from shippers and carriers and were obtained after rounding off.
- The values for rail are for a train (not a wagon).
- The values for inland waterways and sea refer to a ship.
- These values do not include VAT.

Results for passenger transport

For passenger transport, we estimated advanced discrete choice models (so-called panel Latent Class models) that allow the values of time to depend on the actual travel time and travel cost, on the size of the time and cost changes offered in the SP experiments and on other attributes of the respondents (e.g. education, income, age, household composition). We also account for unobserved value of time differences in the population and for the fact that our estimation sample is a panel, i.e. that we have multiple observations from each respondent.

The recommended values of time were calculated by weighting the sampled respondents to represent the distribution of time travelled in the trips recorded in the national travel survey OViN 2010. The resulting values of time and reliability are displayed in Table E3 and E4, respectively.

In most cases, the new values of time are not very different from the values that are used in CBA in The Netherlands at the moment (based on the SP survey of 1997). Generally speaking, they also provide a good match to the values from an international meta-analysis and the values obtained in the recent national value of time studies in Sweden and Norway.

Table E3: Values of time for passenger transport (Euro/hour, price level 2010)

	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Commute	9.25	11.50	7.75	9.75		
Business employee	*12.75*	*15.50*	*10.50*	*13.50*	*85.75*	
Business employer	*13.50*	*4.25*	*8.50*	*10.50*	-	
Business	26.25	19.75	19.00	24.00	85.75	
Other	7.50	7.00	6.00	7.00	47.00	8.25
All purposes	9.00	9.25	6.75	8.75	51.75	8.25

Notes:
- All values are rounded off to the nearest multiple of € 0.25
- These values include VAT.

Table E4: Values of reliability for passenger transport (Euro/hour, price level 2010)

	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Commute	3.75	4.75	3.25	4.00		
Business employee	*14.50*	*18.00*	*12.00*	*15.50*	*56.00*	
Business employer	*15.50*	*4.75*	*9.75*	*12.25*	-	
Business	30.00	22.75	21.75	27.75	56.00	
Other	4.75	4.50	3.75	4.50	30.75	0
All purposes	5.75	5.50	3.75	5.25	33.75	0

Notes:
- All values are rounded off to the nearest multiple of € 0.25
- These values include VAT.

Table E5: Reliability ratios for passenger transport

	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Commute	0.4	0.4	0.4	0.4		
Business	1.1	1.1	1.1	1.1	0.7	
Other	0.6	0.6	0.6	0.6	0.7	0

Notes:

- All values are rounded off to the nearest multiple of 0.1

Acknowledgements

We would like to thank the following persons for providing useful comments in one or more stages of this project:

- Pim Warffemius (min I&M[1]-KiM);
- Pauline Wortelboer (min I&M-KiM);
- Arjen ’t Hoen (min I&M-KiM);
- Jan van der Waard (min I&M-KiM);
- Christa Glasbergen (min I&M-DGB/ OV en Spoor);
- Theo Janssen/Rik Valom (min I&M-DGB/ Maritiem);
- Henk van Mourik (min I&M-DGB/ Wegen en Verkeersveiligheid);
- Anneke de Wit (min I&M-DGB/ Luchtvaart);
- Francis Cheung (Rijkswaterstaat-DVS);
- Toon van der Hoorn (Rijkswaterstaat-DVS);
- Bas Turpijn (Rijkswaterstaat-DVS);
- Sabine Visser (CPB Netherlands Bureau for Economic Policy Analysis);
- Gerbert Romijn (CPB Netherlands Bureau for Economic Policy Analysis);
- Evert Jan v/d Kaa (Delft University of Technology);
- Sytze Rienstra (Syconomy);
- Peter Bonsall (University of Leeds);
- Jonas Eliasson (Royal Institute of Technology, Stockholm);
- Maria Börjesson (Royal Institute of Technology, Stockholm);
- Eric Molin (Delft University of Technology);
- Harry Timmermans (Eindhoven University of Technology);
- Caspar Chorus (Delft University of Technology);
- Bert van Wee (Delft University of Technology);
- Hugh Gunn (HGA);

---

[1] Min I&M is the Dutch Ministry of Infrastructure and the Environment.

- Andrew Daly (RAND Europe & University of Leeds);
- Mogens Fosgerau (Danish Technical University);
- Farideh Ramjerdi (TØI Institute of Transport Economics, Oslo);
- Kay Axhausen (ETH Zürich);
- David Hensher (University of Sydney).

The research team:

- Significance: Gerard de Jong (project leader), Marco Kouwenhoven, Matthijs Joosten & Derk Wentink;
- VU University: Paul Koster, Erik Verhoef, Vincent van den Berg & Yin-Yen Tseng
- John Bates Services: John Bates;
- TNO: Lori Tavasszy, Jannette van Staalduinen & Walther Ploos van Amstel;
- NEA: Jarl Schoemaker;
- TNS NIPO;
- PanelClix.

CHAPTER 1 Introduction

1.1 Study objective

The objective of this project for the Dutch Ministry of Infrastructure and the Environment, was:

> *to provide values of time and travel time reliability (variability) for passenger and freight transport by mode that can be used in cost-benefit analysis (CBA) of transport projects.*

These monetary values are used to convert impacts of a transport project, such as shorter travel times for passengers and freight transport in minutes, into money units, so that these time savings can be incorporated in the CBA of these projects, together with impacts such as investment and maintenance costs, safety and emission impacts.

Values of time studies have been done before in the Netherlands, most recently in 1997 for passenger transport (Hague Consulting Group, 1998) and in 2003/2004 for freight transport (RAND Europe, SEO and Veldkamp/NIPO, 2004). However, the value of travel time reliability has never been measured in a formal value of time study; i.e. a study meant to produce values for actual policy making.[2] Including benefits from travel time variability improvements in CBAs is important, since otherwise the benefits from transport investment and possible also environmental and safety investments may be underestimated.[3] This study therefore needs to update the official values of time in both passenger and freight transport in the Netherlands and to deliver values of reliability based on empirical foundation.

In earlier projects (RAND Europe 2004, Hamer et al. 2005, HEATCO 2006), it was decided that the variability of transport time should be measured by the standard deviation of the travel time distribution. The main reason behind this choice was the assessment that including travel time variability in transport forecasting models would be quite difficult, and that using the standard deviation would be the easiest option. Any formulation that

---

[2] though provisional values have been set (Hamer et al., 2005)

[3] The introduction of 80 km/h on certain parts of the highway network may lead to a lower average speed and hence costs related to travel time losses, but this may be compensated by benefits due to higher reliability of travel times. Also, some incident management investments may lead to additional benefits due to improved travel time reliability.

would go beyond the standard deviation of travel time (or its variance) would be asking too much from the national and regional models (as they might look like in a few years from now) that are regularly used in CBA in The Netherlands.[4]

The modes for freight transport covered in this project are:

- road;
- rail;
- air;
- inland waterways and
- sea transport.

The freight values of time and reliability that are required for use in CBA need to refer to an average vehicle by mode and should be on an hourly basis. If possible and if relevant, a segmentation by type of commodity and containerisation (yes/no) should be made.

For passenger transport, the modes covered are:

- car;
- bus, tram, metro and train[5];
- airplane[6] and
- recreational navigation[7].

The monetary values required for passenger transport should be for a traveller, travelling for a certain travel purpose, per hour (possibly, if relevant, distinguishing between different socio-economic groups, e.g. income groups).

1.2 Overview of the methodology

The project was carried out by a consortium of Significance (project leader), VU University Amsterdam and John Bates Services, in collaboration with TNO, NEA, TNS NIPO and PanelClix.

To meet the study objectives, the project was divided into the following phases:

---

[4] Nevertheless, other specifications than the one where unreliability is measured as the standard deviation of transport time, such as the scheduling model (see Chapter 2), have been tried as well, to see which specification performs best on the data obtained. If a scheduling model did a better job in explaining the data, it would still be possible, under certain conditions, to calculate a standard deviation of transport time from the estimated scheduling coefficients (also see Chapter 2).

[5] Train includes conventional train services as well as high speed rail (however, it will turn out that we do not have enough observations to give separate high-speed rail VOTs, see Chapter 8).

[6] In the previous national VOT surveys of 1988 and 1997, airplane was not included.

[7] In the previous national VOT surveys of 1988 and 1997, recreational navigation was not included. However, a VOT for this mode is regularly needed, especially for the appraisal of proposed locks and bridges. Therefore, it was included in this survey.

1. **Design of questionnaires**
   In this phase, the questionnaires for interviewing travellers, shippers and carriers that were designed in a previous project (Significance et al., 2007), were updated. Through pilot interviews, these questionnaires were tested and improved.

   These questionnaires contained several choice experiments on the values of time and reliability that related to hypothetical alternatives for a trip actually made or a transport actually carried out (by varying costs, time and reliability). Such experiments are called 'Stated Preference' (SP) experiments. These SP experiments were designed in the previous project as well (Significance et al., 2007). The questionnaires also contain some 'Revealed Preference' (RP) questions on observed choices in transport. The questionnaires and the SP and RP experiments were designed such that the results meet the demands for usage in Cost-Benefit Analyses.

2. **Main survey field work**
   The interviews with shippers and carriers were carried out as computerised face-to-face interviews for freight by TNS NIPO. The interviews with travellers were carried out as internet interviews with respondents recruited either from the existing PanelClix panel (2009 survey) or recruited en-route by TNS NIPO for passenger transport (2011 survey).

3. **Analysis of the interview data**
   This phase includes estimation of models explaining the choices made by the respondents between the alternatives offered. The values of time and reliability are then derived from the estimated coefficients of these models, in conjunction with other available data (e.g. on transport costs per hour by mode in freight transport, or on the distribution of minutes travelled across various socio-economic groups, travel purposes, modes and trip lengths for weighting the passenger transport results).

Project memoranda and draft reports in this project have been assessed by a project group from the Ministry of Infrastructure and the Environment, also including the CPB Netherlands Bureau for Economic Policy Analysis. Furthermore, all reports, including this final report, have been reviewed by two external reviewers (Prof. Dr. Bert van Wee and Dr. Caspar Chorus of Delft University of Technology). The questionnaires used were discussed with sector representatives (and adapted when necessary) before using them in the pilot. The resulting values of time and reliability for freight transport were reviewed by freight and logistics experts from TNO (Prof. Walther Ploos van Amstel and Prof. Dr. Lori Tavasszy).

1.3 Contents of this report

This report starts with a general chapter on the theoretical models for the value of time and reliability (Chapter 2).

After this, a block of three chapters concerning freight transport follows:

- A short description of the questionnaires and the SP experiments for freight (Chapter 3). The full survey design is explained in detail in Significance et al. (2007) and in Appendix A.
- A description of the data sets that we collected (Chapter 4)
- A report of the analyses carried out (Chapter 5) containing:
    - a description and justification of the methods used in the analysis of the main survey data;
    - the direct results of the analysis in the form of the estimated coefficients for attributes such as transport cost, transport time and transport time reliability;
    - the final results in terms of recommended values of time and reliability; and
    - a comparison of the outcomes with values in the previous national value of time studies in The Netherlands and the international literature.

Then, we have a block of three chapters concerning passenger transport (Chapters 6- 8) with similar contents as for freight.

Chapter 9 contains reflections and recommendations on how to use these monetary values in combination with passenger and freight transport forecasting models (that currently do not include transport time reliability). And finally, a summary and conclusions can be found in Chapter 10.

The final versions of the questionnaires are available upon request.

CHAPTER 2 Theoretical models for values of time and reliability

In this section we present models from the literature that take account of the influence of both transport time and reliability. There is much more literature on the value of time (VOR) and the value of reliability (VOR) in passenger transport than in freight transport. This section is therefore largely based on passenger transport literature, but we believe that these concepts are also relevant for goods movements.

2.1 Mean-dispersion models versus scheduling models

In the literature on valuing reliability/variability of travel time in passenger transport, two model specifications are used in most cases (see de Jong et al., 2004; Batley et al., 2008; OECD, 2010; Carrion and Levinson, 2012): the mean-dispersion approach and the scheduling approach. They differ in the terms that are included in the utility function that is estimated on the SP data (for an introduction to the random utility model, see McFadden (1978) and Ben-Akiva and Lerman (1985)).

2.1.1 Base models without variability

These models have a simple utility function with a cost and a time term:

$$U = \beta_C \cdot C + \beta_T \cdot T \qquad [1]$$

where:

β_C = Cost coefficient (to be estimated)

C = Travel or transport cost

β_T = Time coefficient (to be estimated)

T = Travel or transport time

The value of time can be calculated by dividing the time coefficient by the cost coefficient:

$$VOT = \frac{\beta_T}{\beta_C} \qquad [2]$$

2.1.2 Mean-dispersion models

In these models the utility function also includes some measure of the dispersion (spread) of the travel time distribution, usually the standard deviation or the variance[8]:

$$U = \beta_C \cdot C + \beta_T \cdot T + \beta_R \cdot \sigma \qquad [3a]$$

or:

$$U = \beta_C \cdot C + \beta_T \cdot T + \beta_R \cdot \sigma^2 \qquad [3b]$$

where:

β_R = Reliability coefficient (to be estimated)

σ = Standard deviation of the travel or transport time distribution

σ^2 = Variance of the travel or transport time distribution

The value of time can still be calculated by equation [2]. The value of reliability is calculated in a similar way:

$$VOR = \frac{\beta_R}{\beta_C} \qquad [4]$$

2.1.3 Scheduling models

This specification defines (un)reliability as the number of minutes that one will arrive earlier or later than preferred (the schedule delay terms or scheduling terms 'Early' and 'Late' below). This can also be defined as the number of minutes that one will depart earlier or later than preferred. This specification can be based on the scheduling model (departure time choice model) developed by Vickrey (1969) and Small (1982). A related alternative scheduling model that starts from the utility at the origin and the destination location over time was presented by Vickrey (1973) and Tseng and Verhoef (2008). The utility function has the following form:

$$U = \beta_C \cdot C + \beta_T \cdot T + \beta_{Early} \cdot Early + \beta_{Late} \cdot Late \qquad [5]$$

where:

β_{Early} = Coefficient on early arrival (to be estimated)

Early = Schedule delay early (in number of minutes earlier than preferred)

β_{Late} = Coefficient on late arrival (to be estimated)

Late = Schedule delay late (in number of minutes later than preferred)

The value of arriving early (*ValEarly*) and of arriving late (*ValLate*) can be calculated as follows:

---

[8] In the literature, these models are usually called 'mean-variance models', which might be confusing given that most applications include the standard deviation, not the variance. We therefore use a more general name.

$$ValEarly = \frac{\beta_{Early}}{\beta_C} \qquad [6a]$$

$$ValLate = \frac{\beta_{Late}}{\beta_C} \qquad [6b]$$

There is a theoretical equivalence relation (under certain assumptions) between the Vickrey/Small scheduling approach and an approach using the mean and the standard deviation of travel time (Bates et al, 2001, Fosgerau and Karlström. 2010). There is also an equivalence relation between the Vickrey/Tseng/Verhoef scheduling model to a model with the mean and the variance of travel time (Fosgerau and Engelson, 2011). Therefore, it is theoretically possible to calculate a dispersion measure (and hence a VOR) from a departure time choice model. The best approach will depend on how one can obtain the best empirical data and which model would fit best in the transport forecasting model system that is used (Börjesson et al., 2011).

2.1.4 Combined mean-dispersion / scheduling models

In principle, it is possible to estimate both a coefficient on the standard deviation (or variance) and on the schedule delay terms. Such a utility function looks like:

$$U = \beta_C \cdot C + \beta_T \cdot T + \beta_R \cdot \sigma + \beta_{Early} \cdot Early + \beta_{Late} \cdot Late \qquad [7]$$

In practice, it is often difficult to find significant coefficients for both β_R and for $\beta_{Early} / \beta_{Late}$. However, the design of our SP experiments is such that we have independent design parameters that control for the standard deviation of the travel time and for the (most likely) arrival time. Therefore, we might be able to find significant values for all coefficients.

Note that the interpretation of the VOR derived from a pure mean-dispersion model might be slightly different from the interpretation of the VOR from the combined model, since all scheduling costs in [3a] might be included in the VOR, whereas they are not included in the VOR derived from [7].

2.2 Absolute versus relative models

All models in Section 2.1 can be called *absolute* models, since their utility functions contain terms of a coefficient multiplied by the absolute value of each parameter. Another specification that can be used is the *relative* model. Such models were used (for all the modes) in the Dutch freight VOT studies of 1992 and 2003/2004 to cope with the heterogeneity in the typical transports in the SP data.

In a relative model, all attributes are expressed as ratios relative to their base value. So, the utility of a fractional change of each attribute is estimated. However, this cannot be done for the scheduling terms (early and late), since it is not sensible to define a fraction of an arrival time[9]. Therefore, we only present the relative equivalent of the mean-dispersion model:

$$U = \beta_C^{rel} \cdot \frac{C}{C_0} + \beta_T^{rel} \cdot \frac{T}{T_0} + \beta_R^{rel} \cdot \frac{\sigma}{\sigma_0} \qquad [8]$$

where:

C_0 = Base value of the travel or transport cost (BaseCost)

T_0 = Base value of the travel or transport time (BaseTime)

σ_0 = Base value of the standard deviation of the travel or transport time distribution

The ratio of the estimated time coefficient to the estimated cost coefficient can be treated as a trade-off ratio that indicates how relative changes in time are traded off against relative changes in costs.

$$TR = \frac{\beta_T^{rel}}{\beta_C^{rel}} \qquad [9]$$

By multiplying this ratio by the transport cost per hour for a mode (or vehicle type within a mode), the so-called ‘factor costs’, we obtain the VOT (and similarly the VOR):

$$VOT = TR \cdot FactorCost \qquad [10]$$

2.3 Models in preference space and in WTP space

The utility functions in Section 2.1 can be called models in *preference space*, since the coefficients to be estimated are all in utility units. The VOT must be calculated by dividing two coefficients, see equation [2]. These functions in preference space are mathematically equivalent to the following models in in *willingness to pay (WTP) space,* where the VOT and VOR variables are estimated directly.

The equivalent of the mean-dispersion model [3a] in WTP space reads:

$$U = \beta_C \cdot (C + VOT \cdot T + VOR \cdot \sigma) \qquad [11]$$

And the equivalent of the scheduling model [5] in WTP space is:

$$U = \beta_C \cdot (C + VOT \cdot T + ValEarly \cdot Early + ValLate \cdot Late) \qquad [12]$$

Note that equation [11] can be rewritten to allow for the direct estimation of the reliability ratio *RR*:

$$U = \beta_C \cdot [C + VOT \cdot (T + RR \cdot \sigma)] \qquad [13]$$

---

[9] a fraction of a *deviation* from the desired arrival time could, however, be included, if the initial deviation would be unequal to zero. But for travellers who arrive at their desired moment, this would imply division by zero)

2.4 Models in logWTP space

The Danish, Norwegian and Sweden Value-of-Time surveys were estimated in logWTP space, i.e. they used logarithmic utility functions in their estimation processes (e.g. Fosgerau, 2006a,b, Börjesson et al., 2011, Börjesson and Eliasson, 2011). This was done since they discovered that their SP observations were better reproduced when the error-terms were multiplicative rather than additive (in other words: the size of the error-terms depended on the time and cost levels). Using logarithmic utilities weights the information of all respondents in a different way than normal utility functions do.

For instance in a mean-dispersion model this gives:

$$U = \lambda \cdot \log\left(C + VOT \cdot T + VOR \cdot \sigma\right) \qquad [14]$$

where:

λ = scale parameter

It is noteworthy that the consumer surplus calculated from a model estimated in logWTP space is different from the conventional consumer surplus and is much more complicated (Fosgerau & Bierlaire, 2009).

2.5 Prospect theory

Prospect theory (Kahneman and Tversky, 1979, 1992; van de Kaa, 2008) says that:

- the valuation of an attribute depends on the current value of that attribute, i.e. it depends on the reference alternative (the situation as observed now): reference dependence
- there will be a difference in the valuation of gains and losses in an attribute: losses are valued more negatively: loss aversion
- there will be a difference in the valuation at different values of an attribute (e.g. between a short and a long transport): size dependence.

If such relationships are found in SP data, the researcher can try to control for the influence of these on the final VOT and VOR results (de Borger and Fosgerau, 2008). The loss aversion hypothesis in particular states that gains are valued differently from losses. It is common (see earlier references to the Scandinavian VOT studies) to estimate separate values for the VOT for (also see Figure 1):

- the Willingness-to-Pay (WTP), in which a time improvement combined with a cost increase is evaluated against the current situation,
- the Willingness-to-Accept (WTA), in which a time deterioration in combination with a cost reduction is evaluated against the current situation,
- an Equivalent Gain (EG), in which a time improvement (and current cost) is evaluated against a cost reduction (and current time), and
- an Equivalent Loss (EL) in which a time deterioration (and current cost) is evaluated against a cost increase (and current time).

Theory predicts that the WTP is smaller than the value for EG and EL, and these values are again smaller than the WTA: WTP < EG, EL < WTA.

There is also literature on non-linear weighting of probabilities by respondents (Hensher et al., 2011a,b; Hensher and Li, 2012). We have not used these methods in this study, since they have not yet been used to generate official VOTs and VORs[10] and a number of issues need to be resolved within academic research first. Furthermore, Koster and Verhoef (2012) found in an application for morning peak travellers that the difference in terms of costs between linear and non-linear weighting of probabilities was only 3%.

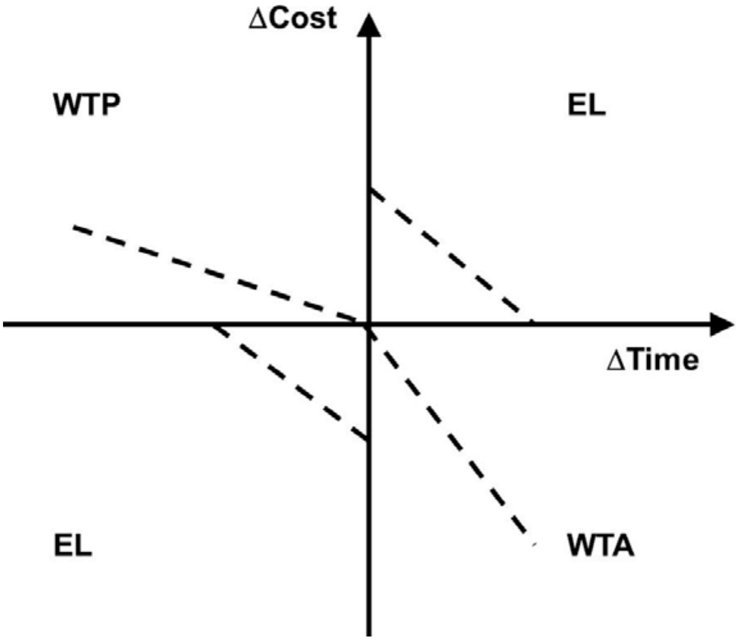


Figure 1: Cost and time differences presented in the SP

2.6 Advanced MNL models with diminishing sensitivity for higher base levels and small changes

2.6.1 Introducing diminishing sensitivity for higher base levels

From earlier work (e.g. Gunn, 2001; Mackie et al., 2003; Daly et al., 2011; Stathopoulos and Hess, 2011) it is known that the VOT can be strongly dependent on the current level of the travel time and travel cost of the respondent, just as is claimed by prospect theory. These BaseCost (C_0) and BaseTime (T_0) values are used in the SP experiments around which the time and cost levels are varied. It can be expected that both levels are correlated.

---

[10] The official Danish, Swedish and Norwegian VOTs and VORs are not based on these methods, although there have been some applications to Scandinavian data recently, see e.g. Hjorth and Ramjerdi (2011).

We therefore try to include the BaseTime and BaseCost dependency in the utility specification. Following earlier authors (e,g, Mackie et al., 2003; Stathopoulos and Hess, 2011), we try a power law dependence of both the BaseCost and BaseTime:

$$U = \beta_C \cdot \left(C \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} + VOT_{ref} \cdot T \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \right) \qquad [15]$$

where C_0 and T_0 are the individual's base cost and base time, and C_{ref} and T_{ref} are arbitrary reference values.

The VOT now depends on C_0 and T_0:

$$VOT(C_0, T_0) = \left. \frac{\partial U / \partial T}{\partial U / \partial C} \right|_{T=T_0, C=C_0} = VOT_{ref} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} / \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} \qquad [16]$$

From the utility definition in equation [15] it can be seen that choosing different reference points C_{ref} and T_{ref}, will influence the values of β_C and VOT_{ref}, but the utility U itself remains indifferent to this choice. Therefore, the $VOT(C_0, T_0)$ in equation [16] is also not dependent on C_{ref} and T_{ref}, so and hence we can choose them freely.

2.6.2 Allowing different values for small / large changes

The most recent VOT studies (e.g. De Borger and Fosgerau (2008) for the Danish data and Börjesson and Eliasson (2011) for the Swedish data) allow for different VOTs for small and large time savings offered in the SP (but recommend using a single value for a large and a small time savings offered by transport projects). We also want to correct for the influence of the size of the time and cost changes offered in the SP.

The previous utility function [15] is equivalent to

$$U = \beta_C \cdot \left((C_0 + \Delta C) \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} + VOT_{ref} \cdot (T_0 + \Delta T) \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \right) \qquad [17]$$

Note that from the point of view of binary choice C_0 and T_0 can be taken out of this utility function without affecting the coefficients (since they will appear in the utility of both alternatives), so we can reduce it to:

$$U = \beta_C \cdot \left(\Delta C \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} + VOT_{ref} \cdot \Delta T \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \right) \qquad [18]$$

Some of the dependency on BaseTime might be caused by a difference in valuation between small and large changes. The design of the SP experiment is such that larger changes in time and cost were offered for larger values for BaseTime. In the next step of the model development, we therefore take into account that small changes in time and costs might be valued differently from large changes. So, we add an exponent to the ΔC and ΔT-terms to investigate whether there is any change in sensitivity further away from the base (current) values of cost and time. Therefore, we include factors that can take this changing sensitivity into account:

$$U = \beta_C \cdot \left(\Delta C \cdot \left(\frac{|\Delta C|}{\Delta C_{ref}} \right)^{\kappa_C} \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} + VOT_{ref} \cdot \Delta T \cdot \left(\frac{|\Delta T|}{\Delta T_{ref}} \right)^{\kappa_T} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \right) \tag{19}$$

where again the factors ΔC_{ref} and ΔT_{ref} do not affect the estimates in any decisive way.

This equation can be rewritten as:

$$U = \beta_C \cdot \left(\operatorname{sgn}(\Delta C) \cdot \left(\frac{|\Delta C|}{\Delta C_{ref}} \right)^{\gamma_C} \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} \cdot \Delta C_{ref} + VOT_{ref} \cdot \operatorname{sgn}(\Delta T) \cdot \left(\frac{|\Delta T|}{\Delta T_{ref}} \right)^{\gamma_T} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \cdot \Delta T_{ref} \right) \tag{20}$$

where $\kappa_C = \gamma_C + 1$ and $\operatorname{sgn}(\Delta C) = \frac{\Delta C}{|\Delta C|}$ and similar for sgn(ΔT). This "sign"-function is necessary, since we can only take a fractional exponent of a positive number. Now, the VOT, obtained by differentiation, will be dependent on C_0 and T_0, but also on ΔC and ΔT:

$$VOT(C_0, T_0, \Delta C, \Delta T) = VOT_{ref} \cdot \frac{\gamma_T \cdot \left(\frac{|\Delta T|}{\Delta T_{ref}} \right)^{\gamma_T - 1} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T}}{\gamma_C \cdot \left(\frac{|\Delta C|}{\Delta C_{ref}} \right)^{\gamma_C - 1} \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C}} \tag{21}$$

Note that the $VOT(C_0, T_0, \Delta T, \Delta C)$ is independent of the chosen reference values for T_{ref}, C_{ref}, ΔT_{ref} and ΔC_{ref}, since

$$VOT_{ref} \propto \frac{\Delta T_{ref}{}^{\gamma_T - 1} \cdot T_{ref}{}^{\lambda_T}}{\Delta C_{ref}{}^{\gamma_C - 1} \cdot C_{ref}{}^{\lambda_C}} \tag{22}$$

2.6.3 Including a dispersion variable

By analogy, we extend the above analysis to a mean-dispersion model, by adding the standard deviation σ with an exponent on the reliability level as well:

$$U = \beta_C \cdot \left(\operatorname{sgn}(\Delta C) \cdot \left(\frac{|\Delta C|}{\Delta C_{ref}} \right)^{\gamma_C} \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} \cdot \Delta C_{ref} + VOT_{ref} \cdot \operatorname{sgn}(\Delta T) \cdot \left(\frac{|\Delta T|}{\Delta T_{ref}} \right)^{\gamma_T} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \cdot \Delta T_{ref} + VOR_{ref} \cdot \sigma \cdot \left(\frac{\sigma_0}{\sigma_{ref}} \right)^{\lambda_R} \right) \tag{23}$$

Since one of the objectives of this research is to determine values of reliability in terms of standard deviations or reliability ratios for use in practical CBA, we prefer using a mean-dispersion model, where all reliability influences will be incorporated in the standard deviation term, over a pure scheduling model (where a reliability ratio can only be calculated indirectly) or a mixed scheduling and dispersion model (where the standard deviation only captures part of the reliability influences). However, estimation results for a pure scheduling model will be reported in Section 8.3

2.7 Multinomial Logit models versus Mixed Logit models

The MNL models that we have discussed so far can be extended to include interaction variables, e.g. for characteristics of the person, firm, trip or the shipment to account for 'observed heterogeneity'. These models with interaction variables take the differences between respondents of different characteristics such as sex, age, income, etc. into account. But not all characteristics that influence the estimates can be taken into account, e.g. because no information on these characteristics is available. Also, two respondents with exactly the same characteristics could have different model coefficients. The MNL results are incorrect if these individual coefficients are related to other factors in the models (travel time for instance).

Mixed logit (ML) models allow for random taste variation between respondents ('unobserved heterogeneity') as well. These models use the same utility functions as the MNL models. However, they assume that one (or more) of the coefficients do not have a fixed value, but have an underlying distribution. This can be either a continuous distribution or a discrete distribution. In case of a continuous distribution usually a specific statistical distribution is employed such as normal or lognormal. Then, the ML procedure estimates both the mean and the standard deviation of this distribution. When the standard deviation is not significantly different from 0, there is no significant taste heterogeneity in the data and therefore it can always easily be tested whether mixed logit models are significantly better than MNL models.

The estimation of continuous distribution mixed logit models requires repeated sampling from a statistical distribution; therefore, it takes much longer to estimate than MNL.

2.8 Continuous distribution mixed logit models versus Latent Class models

A serious drawback of the continuous distribution mixed logit models is that the distribution of tastes may be misspecified (Fosgerau, 2006; Fosgerau and Bierlaire, 2007). This means that the shape of the real distribution of tastes is very different from the distribution that is imposed by the researcher. This can lead to biased estimates.

There are two ways to get around this problem. First, non- or semi-parametric estimation can be used, where no arbitrary assumptions are made about the continuous distribution of tastes. Second, a more common approach is to use latent class (LC) or finite mixture models. These assume discrete distributions for certain coefficients, but without imposing a particular shape. The result is a 'histogram' with class probabilities, and corresponding estimated values for the coefficients. A latent class model with one class is equivalent to a standard MNL model. We therefore can apply statistical tests to test for heterogeneity and determine the number of latent classes in the data.

Latent class models are thus a special case of mixed logit models: they are ML models with a discrete distribution for one or more of its coefficients.

2.9 Cross-sectional models versus panel models

Most MNL models applied to SP data assume that the errors relating to the different observations of an individual are unrelated. Such models are called ‘cross-sectional’ since they analyse each choice in isolation and do not account for the fact that respondents make a sequence of choices. If errors are correlated over the choices of the same respondent, or the size of the error is correlated to one of the explanatory variables, this may result in a biased estimate in the MNL model.

In principle, it is possible to account for this in an MNL model by modelling the choice sequence of each respondent rather than modelling all choices separately, but in most cases this is rather unpractical given the large number of possible sequences.

Panel mixed logit models can account for taste heterogeneity and repeated measurements, by using a common random component for the same individual and keeping the tastes of an individual constant over a series of choices.

2.10 Non-parametric techniques

The logit model assumes an extreme value distribution of the error term as opposed to non-parametric estimation techniques that do not make assumptions on the statistical distribution. The extreme value assumption is a common assumption because logit models are less ‘data hungry’ than non-parametric estimation techniques. Furthermore, non-parametric panel data models are very complex to understand and estimate. The parametric logit model has the advantage that it does not suffer from potential identification problems (Chamberlain, 2010). Non-parametric techniques have been used in the recent VOT studies in Denmark (Fosgerau, 2006a,b), Sweden (Börjesson and Eliasson, 2011) and Norway (Ramjerdi et al, 2010).

Sometimes non- or semi-parametric techniques are used within mixed logit models (Börjesson et al., 2012a). In those cases, the extreme value assumption for the error term is still being made, however, the non-parametric technique is used to determine which shape of the distribution of (one of) the coefficients is appropriate (see for example: Bastin et al. 2010; Fosgerau and Bierlaire, 2007).

FREIGHT TRANSPORT

CHAPTER 3 Freight transport questionnaire

3.1 A priori expectations for the VOT of shippers and carriers

The main difference between the freight and the passenger surveys are the specific instructions on whether to include considerations related to the cargo and related to the transport services for shippers and carriers. The interviews carried out in freight transport refer to a typical transport (shipment)[11], and the VOTs that we obtain will therefore initially be per shipment. However, we shall convert this into values per vehicle/vessel, which is the most appropriate unit for use in CBA.

We make the following assumptions (a priori hypotheses) on the freight value of time (Table 1). These hypotheses were supported by experts at an expert meeting which took place at Schiphol airport on 25 October 2004 (Hamer et al. (2005), de Jong et al. (2009)).

Table 1: Hypothesis on the aspects that freight respondents include their VOT

	VOT related to cargo	VOT related to vehicles and staff
Carrier	Not included	Proportional to factor cost
Own account shippers	Interest, deterioration, disruption of production, out of stock	Proportional to factor cost
Shipper that contracts out	Interest, deterioration, disruption of production, out of stock	Not included

Carriers are in the best position to give the VOT that is related to the costs of providing transport services. If the transport time decreased, vehicles and staff would be released for other transports, so there would be vehicle and labour cost savings. Results in the Netherlands and other countries so far (see de Jong, 2008) indicate that the VOT that is

---

[11] Shippers that contract their transport out were asked to select a "typical transport (using a prescribed mode) that is regularly carried out for your firm by a carrier (this is a shipment that is representative for your firm in terms of packaging, distance, destination, etc." Carriers (and shippers with own transport) were asked to select a "typical transport (using a prescribed mode) that is regularly carried out by your firm (this is a shipment that is representative for your firm in terms of packaging, distance, destination, etc." We found that for most carriers this typical transport was equivalent to a loading unit (such as a container or truckload), whereas for shippers the typical transports were mostly shipments in the sense of an amount of goods that leave a sending firm (e.g. manufacturer) for a receiver at the same time.

related to the transport services is more or less equal to the vehicle and labour cost per hour (the 'factor cost'), at least for road transport.

Shippers that contract out are most interested in another VOT, that is the VOT that is related to the goods themselves. This includes the interest costs on the capital invested in the goods during the time that the transport takes (only important for high-value goods, but we did not impose a definition of high value on the respondents), a reduction in the value of perishable goods during transit, but also the possibility that the production process is disrupted by missing inputs or that customers cannot be supplied due to lack of stock. The latter two arguments are also (possibly even more so) important for the value of transport time reliability (VOR).

Shippers with own account transport can give information on both the VOT that is related to the costs of providing transport services and the VOT that is related to the goods themselves. If both VOT components are properly distinguished, for this group the VOT related to the cargo and that related to the vehicles and staff can be added, as can the carrier VOT and shipper (contract out) VOT to obtain the overall VOT for use in societal cost-benefit analysis.

In this study VOTs are sought that include both components (transport services-related and goods-related), since in CBAs for transport projects in The Netherlands the user benefits of savings in vehicle and staff cost are included in the time savings of the project. Previous studies have not tried to disentangle the two VOT components, but in the current study we will obtain estimates for both components separately.

Of course there may be exceptions to the general pattern depicted in Table 1, but in the freight questionnaires we steer the shippers that contract out only to answer on the components they generally know most about (bottom-left), and likewise for carriers (top-right). We do this by giving very explicit instructions and explanations to get clearly defined component values from each type of agent. In other words, we:

1. Explain to all respondents that the changes in time, costs and reliability are generic: these apply to all carriers using the same infrastructure, and are not competitive advantages for their specific firm.
2. Explain to carriers (and logistics service providers) that a shorter transport time might be used for other transports: the staff and vehicles/vessels can be released for other productive activities. A higher reliability means that the carrier can be more certain about such re-planning/re-scheduling. Also explain that they do not have to take into account what would happen (deterioration, disruption of production process, running out of stock, etc.) to the goods if they were late.
3. Explain to the shippers that contract out that they only have to take into account what would happen (deterioration, disruption of production process, running out of stock, etc.) to the goods if the delivery were late (whether these things would occur and how important they are was left to the respondent).
4. Explain to shippers with own account transport that they have to take all of this (=cargo and vehicle) into account.

3.2 Set-up of the questionnaire

The freight transport questionnaire consisted of the following parts:

1. questions regarding the firm;
2. selection of a typical transport (see Section 3.1) and questions on the attributes of this transport, such as transport time and costs. These values are used as base levels for the attribute levels presented in the SP experiments.
3. questions on the availability of other modes for this transport and what the attribute levels would be for that mode (as basis for estimating an RP model). For the carriers this referred to a different route rather than a mode.
4. SP experiment 1 (transport time versus transport cost)
5. Introduction of variable transport times and SP experiment 2a (usual transport time, variation in transport times, most likely arrival time, and transport costs). Note that carriers using inland waterways or sea did not participate in this experiment.
6. SP experiment 2b (same as 2a without the variation in most likely arrival time).
7. questions in which the shippers or carriers were asked to evaluate the choices they made in the experiments.

The full questionnaires are available on request.

The SP experiments were set-up dependent on the type of transport. For road, rail and air transport, respondents took part in the three experiments (1, 2a and 2b). These are described in Section 3.3. For inland waterways and sea, however, respondents took part in only two experiments, which were slightly different in nature.

3.3 SP experiments for shippers and carriers (excl. carriers using sea and inland waterways)

The statistical design and the choice of presentation format for reliability are described in Significance et al. (2007) and Tseng et al. (2009). In this section we repeat the most important aspects of the design.

The list of attributes for the three SP experiments is shown in Table 2.

Table 2: List of attributes in SP experiments for shippers and carriers (excluding carriers using sea and inland waterways)

Attribute	Experiment 1	Experiment 2a	Experiment 2b
Usual transport time	√	√	√
Transport cost	√	√	√
Reliability, i.e. five possible transport times		√	√
Five possible arrival times		√*	√#
Departure time		√^	√^

Notes:

* : the most likely arrival time (which we define as the second and third possible arrival time, which were always set to the same value) vary according to the table as described in the design report. The other arrival times are calculated from this most likely arrival time and the five possible transport times

\# : the most likely arrival time is always equal to the expected arrival time of the typical transport as described by the respondent. Hence, this attribute is fixed and is not varied according to any design table. The other arrival times are calculated from this most likely arrival time and the five possible transport times.

^ : the departure time is calculated from the most likely arrival time and usual transport time. It is not varied according to any design table.

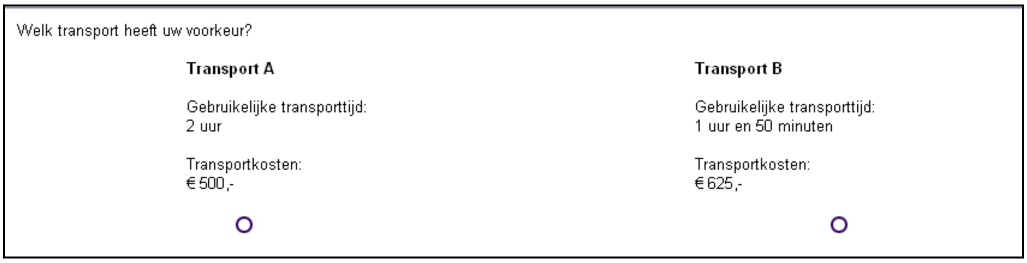
3.3.1 Description of the experiments

The SP data for all shippers and for the carriers in road, rail and air transport contain three SP experiments, with 19 pairwise choices in total. All experiments present within-mode choices. This means that both alternatives of each pairwise choice concern the same mode, which is the same as the mode used for the recent transport.

The interviews were carried out face-to-face by an interviewer visiting the respondent. The interviewer used a computer for the questionnaire. All SP experiments were generated during the interview and used previous answers regarding the typical transport. During the SP experiments, the alternatives were shown on screen to the respondent. The example choice pairs below are actual screen shots.

Experiment 1

The first experiment consisted of six SP within-mode choice pairs with each alternative being described by two attributes: transport time and transport cost. An example in is given in Figure 2. The attributes are discussed in more detail in Section 3.3.2.



Welk transport heeft uw voorkeur?

Transport A	**Transport B**
Gebruikelijke transporttijd: 2 uur	Gebruikelijke transporttijd: 1 uur en 50 minuten
Transportkosten: € 500,-	Transportkosten: € 625,-
○	○

Figure 2: Example of SP question of experiment 1 for shippers and carriers (excluding carriers using sea and inland waterways)

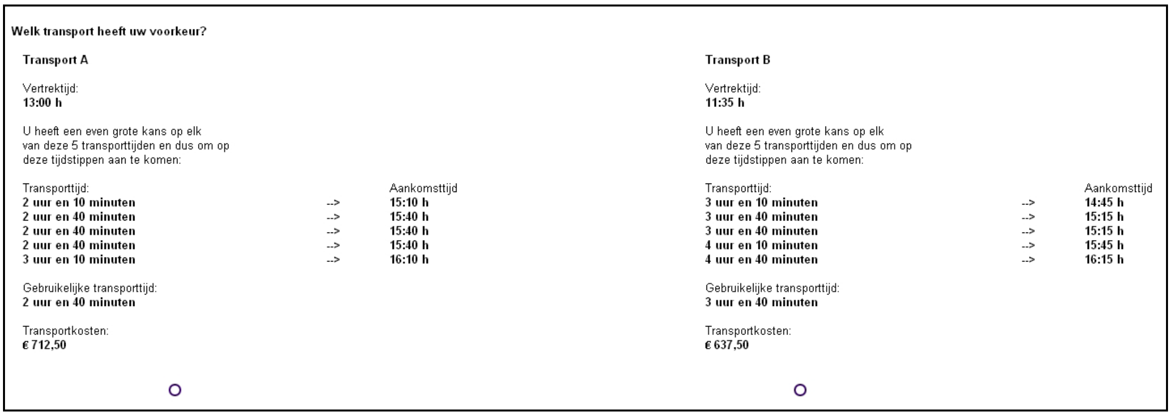
The usual transport time ("Gebruikelijke transporttijd" in Dutch) varied roughly between -14% and +20% around the base time, i.e. the transport time of the typical transport as described by the respondent before. The actual time variations were taken from the design table as given in Significance et al. (2007), which are repeated in Appendix A. Note that the order of the SP questions within each experiment was randomized.

The attribute levels of the transport cost ("Transportkosten" in Dutch) were 85%, 95%, 100%, 110% and 125% of the base cost. This is different from the attribute levels as described in Significance et al. (2007), since we discovered in the pilot stage that the base costs (or more precisely, the base cost per minute of transport time) were fluctuating too much to allow the usage of a table with bands of fixed absolute cost steps. The underlying design (i.e. which attribute levels were shown in which SP question remained unchanged, see also Appendix A).

Experiment 2a

The next experiment consisted of six choice pairs with each alternative being described by four attributes: transport time, transport cost, reliability, and arrival time (departure times are presented as well, but these are not independent of reliability/arrival time). Experiment 2a is based on a statistical design in which expected arrival time is varied (for details see Significance et al., 2007). An example of experiment 2a is given in Figure 3.

The five possible (and equally likely) transport times are constructed such that the second and third are always the same and are equal to the "usual transport time". So, the usual transport time is always the most likely transport time. The second and third possible arrival times are then the most likely arrival time (i.e. the "usual arrival time"). Furthermore, the five possible transport times were asymmetric (for most levels), i.e. the longest transport time differed much more from the usual transport time than the shortest time.



Welk transport heeft uw voorkeur?

Transport A

Vertrektijd:
13:00 h

U heeft een even grote kans op elk
van deze 5 transporttijden en dus om op
deze tijdstippen aan te komen:

Transporttijd:		Aankomsttijd
2 uur en 10 minuten	-->	**15:10 h**
2 uur en 40 minuten	-->	**15:40 h**
2 uur en 40 minuten	-->	**15:40 h**
2 uur en 40 minuten	-->	**15:40 h**
3 uur en 10 minuten	-->	**16:10 h**

Gebruikelijke transporttijd:
2 uur en 40 minuten

Transportkosten:
€ 712,50

○

Transport B

Vertrektijd:
11:35 h

U heeft een even grote kans op elk
van deze 5 transporttijden en dus om op
deze tijdstippen aan te komen:

Transporttijd:		Aankomsttijd
3 uur en 10 minuten	-->	**14:45 h**
3 uur en 40 minuten	-->	**15:15 h**
3 uur en 40 minuten	-->	**15:15 h**
4 uur en 10 minuten	-->	**15:45 h**
4 uur en 40 minuten	-->	**16:15 h**

Gebruikelijke transporttijd:
3 uur en 40 minuten

Transportkosten:
€ 637,50

○

Figure 3: Example of SP question of experiment 2a for shippers and carriers (excluding carriers using sea and inland waterways)

Experiment 2b

The final experiment has seven choice pairs with each alternative being described by the same four attributes as in experiment 2a. Though the choice pairs of both experiments look the same, the underlying statistical design is different. For experiment 2b the most likely arrival time is kept constant (see Significance et al. (2007) for a motivation). This implies that the second and third possible arrival times are always the same for both alternatives.

The sixth out of the seven choice pairs is always a dominant choice, i.e. a choice pair for which all attributes of one of the alternatives is always equal or worse compared to the other alternative. This choice pair is used as a check; it is not included in model estimation. This dominant choice is discussed in more detail in Section 6.4.

3.3.2 Description of the attributes

Below we describe the attributes used.

- **Transport cost**

  For *carriers and own account shippers* this refers to door-to-door[12] transport costs (fuel, staff, depreciation and maintenance of equipment used, administration, insurance, fines paid to the client, social security payments and taxes charged, but not including VAT), including possible transshipment costs, but excluding initial loading and final unloading.

  For *shippers* that contract out transport services it is the price paid for the door-to door transport services, including for transshipments, if any, not including VAT.

  Unlike the 2003/2004 freight VOT survey, factor cost calculations were not used in the questionnaire. In 2003/2004, for respondents who did not give transport costs/prices, we used the standard factor costs calculations (NEA et al, 2003) for the reference level. In the new survey we urge the respondents to give their best possible estimate of the cost/price. If they cannot give any estimate, there was no point in continuing the interview.

  The current (observed) level for the transport cost is referred to as "*base cost*".

- **Transport time**

  This refers to the (one-way) trip from door-to-door (including transfer time and the average delay the respondent normally encounters for this transport).

  The current (observed) level for the transport time is referred to as "*base time*".

- **Reliability**

  This is the variability of total door-to-door transport time. This attribute is presented as a series of five possible transport times, described only verbally (see Tseng et al. (2009) for a justification).

---

[12] For carriers (incl. logistics service providers), this may refer to transport to/from a port, airport or railway terminal.

- **Departure / arrival time**
  This is the departure / arrival time at the gate of the sender/receiver (arrival time is presented for each of the five possible transport times). The base level for the departure time is the observed (real) departure time. The base level for the arrival time is the expected arrival time, which is the time at which the respondent expected the transport to arrive when the transport began. This is not necessarily equal to the preferred arrival time (which is defined as the arrival time the respondent would choose if there would be no congestion).

3.4 SP experiments for carriers using sea and inland waterways

For transporters in sea and inland waterways transport, discussions with professionals from the sector led us to choose a different setting. Here we have used a setting where a ship is waiting for a lock, bridge or to be loaded/unloaded at a quay in the port. The attributes used are presented in Table 3.

Table 3: List of attributes in SP experiments for carriers using sea and inland waterways

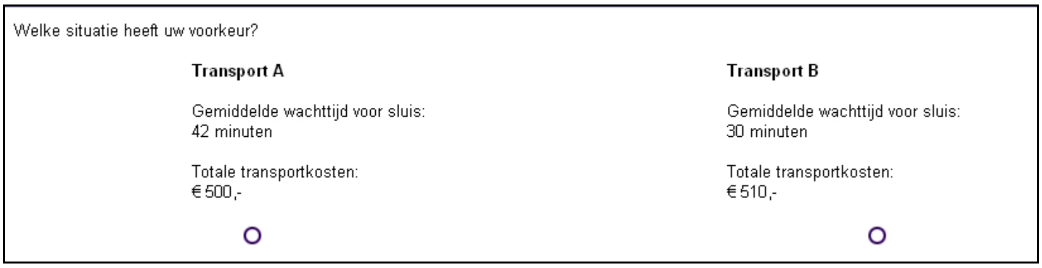
Attribute	Experiment 1	Experiment 2b
Average waiting time	√	√
Transport cost	√	√
Reliability, i.e. five possible waiting times		√

Note that no departure and arrival times were presented in the choice alternatives.

The SP experiments are:

Experiment 1

Six choice pairs with average waiting time at the lock/bridge/quay and total transport cost (in case of lock/bridge) or cost of using the quay. An example of this lock/bridge experiment is given in Figure 4.



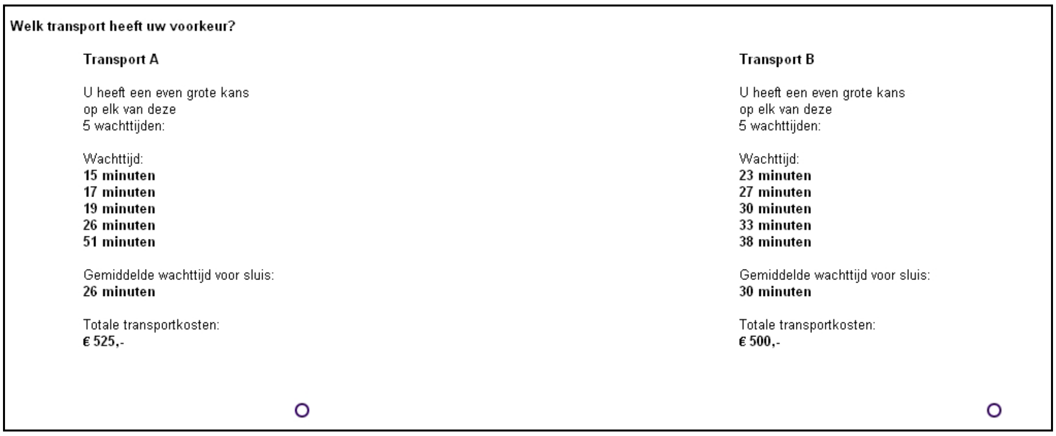
Welke situatie heeft uw voorkeur?

Transport A	**Transport B**
Gemiddelde wachttijd voor sluis: 42 minuten	Gemiddelde wachttijd voor sluis: 30 minuten
Totale transportkosten: € 500,-	Totale transportkosten: € 510,-
○	○

Figure 4: Example of SP question of experiment 1 for carriers using sea and inland waterways

Experiment 2b

Six choice pairs plus a dominant choice pair (the latter is not used in estimation) with waiting time, variability of waiting time (presented by five possible waiting times) and transport cost. An example is shown in Figure 5.



Welk transport heeft uw voorkeur?

Transport A	**Transport B**
U heeft een even grote kans op elk van deze 5 wachttijden:	U heeft een even grote kans op elk van deze 5 wachttijden:
Wachttijd: **15 minuten** **17 minuten** **19 minuten** **26 minuten** **51 minuten**	Wachttijd: **23 minuten** **27 minuten** **30 minuten** **33 minuten** **38 minuten**
Gemiddelde wachttijd voor sluis: **26 minuten**	Gemiddelde wachttijd voor sluis: **30 minuten**
Totale transportkosten: **€ 525,-**	Totale transportkosten: **€ 500,-**
○	○

Figure 5: Example of SP question of experiment 2b for carriers using sea and inland waterways

The average waiting time presented was always the average of the five possible waiting times. This differs from the experiment 2a/2b for shippers and other carriers, since the asymmetric distribution of their possible transport times caused the average transport time always to be different from the usual (i.e. most likely) transport time. The complete design can be found in Appendix A.

No experiment 2a with additional variation in arrival time was done, since representatives from the inland waterway transport and maritime sectors had told us in pre-interviews that an initial design with three experiments would take too long for their sectors, and that a time scheduling context would not generally be recognised.

CHAPTER 4 Freight transport data

4.1 Number of respondents

Shipper and carrier firms were recruited from existing registers of firms (e.g. from Chambers of Commerce) and approached (mostly by phone) to seek firms that were prepared to participate in the interviews. The subsequent interviews were carried out as face-to-face interviews where a professional interviewer visited the firm and the questions were shown on a laptop computer.

Table 4 shows the number of respondents for each of the questionnaire types (by means of different colours – see below) and for each mode. Note that Table 4 also distinguishes between container and non-container transport. This distinction is one of the requirements specified in the tender documents (since previous studies (e.g. RAND Europe et al., 2004) have found VOT differences between these segments or have explained why these could occur (e.g. TNO-Inro and MuConsult, 2002)).

Table 4: Number of freight respondents by (sub)segment

		Road	Rail	Air	Inland water-ways	Sea	Total
Container	**Carrier**	35	10	0	16	18	**79**
	Own account shipper	10	2	0	0	0	**12**
	Contract out shipper	41	14	0	18	80	**153**
Non-container	**Carrier**	131	5	19	69	12	**236**
	Own account shipper	36	0	0	0	0	**36**
	Contract out shipper	162	19	44	22	49	**296**
Total		**415**	**50**	**63**	**125**	**159**	**812**

Note: the questionnaire types are indicated by a shading colour:

🟦 Questionnaire type A – carrier (road, rail, air)

🟧 Questionnaire type B – shipper that contracts out (all modes)

🟨 Questionnaire type C – own account shipper (road, rail, air)

🟩 Questionnaire type D – inland waterways and sea transport carriers

These numbers should be compared to the targets set at the beginning of the study. In Table 5 both numbers are indicated using a format such as 45/50, where 45 indicates the total number of respondents and 50 indicates the original target.

Table 5: Comparison of the number of freight respondents with targets

		Road	Rail	Air	Inland water-ways	Sea	Total
Container	**Carrier/own account shipper**	45/50	26/20	0/0	16/40	18/20	**244/230**
	Contract out shipper	41/50			18/20	80/30	
Non-container	**Carrier/own account shipper**	167/50	24/30	19/30	69/40	12/30	**568/290**
	Contract out shipper	162/50		44/20	22/20	49/20	
Total		**415/200**	**50/50**	**63/50**	**125/120**	**159/100**	**812/520**

Note that the targets for the total number of respondents by mode have been amply met, though some sub-segment targets have been short. After a first data analysis and after consultation with the client it was decided that this data set was sufficient for the present study. Data from all respondents (including respondents over and above the required target number) were included at the start of the analysis.

4.2 Data selection and data quality

4.2.1 Treatment of the dominant alternative

Each series of 7 SP questions in experiment 2B contained a dominant choice pair. This is a pair of SP alternatives where each attributes of one alternative is worse than the same attribute of the other alternative (see Section 3.2). Table 6 shows the share of persons in each of the segments that choose the dominated alternative (the alternative that was inferior on all presented attributes; the other alternative then is dominant). The percentages of respondents choosing the dominated alternative are generally in line with expectations (12.3% over all respondents).

Table 6: Share per segment choosing the dominated alternative

		Road	Rail	Air	Inland water-ways	Sea
Container	**Carrier/own account shipper**	22.2%	3.9%	-	6.3%	5.6%
	Contract out shipper	2.4%			16.7%	11.3%
Non-container	**Carrier/own account shipper**	13.8%	8.3%	36.8%	8.7%	16.7%
	Contract out shipper	9.3%		11.4%	18.2%	20.4%

We tested models with and without these respondents for the largest segments (road carriers/own account shippers and contract out shippers for all modes) and for these largest segments the results were inconclusive: sometimes deleting the respondents that choose the dominated alternative improves the t-ratios (because one obtains a ‘better’ selection of respondents), but sometimes the t-ratios go down (because the sample size decreases). Since the sample size for the smaller segments is often already quite small, we decided to keep the respondents that choose the dominated alternative in all models for freight transport.

4.2.2 Outliers

In the analysis own account is combined with carriers because initial estimations showed that no separate models for own account could be estimated, and hence shippers here means “contract out shippers”. For each segment, we have investigated the base characteristics (base time, base cost, weight of the shipment) of the typical transports as described by the respondents. Based on this analysis, it was decided that some respondents had to be excluded from further analysis.

- **Carriers road (including own account shippers)**
  Containerised transport: we have 45 respondents in the raw sample (see Table 4). Two outliers on base time (> 6 days, i.e. 8640 minutes) were excluded. One respondent with strange levels for base time, cost, and weight has been removed. Furthermore, three interviews with weight of the shipment of more than 40 tonnes were discarded. The reason for excluding these observations is that 40 tonnes is the maximum amount of cargo that can be transported by a single lorry. Transports over 40 tonnes, thus, would entail the use of more than one lorry, which would result in difficulties in the cost estimations. This leaves 39 respondents in the sample.

  Non-containerised transport: we have 167 respondents in the raw sample (see Table 4). One outlier on base time (> 6 days, i.e. 8640 minutes) and two outliers on base costs (< 10 euro) were excluded. Furthermore, seven interviews with an unknown weight of the transport or with a weight of more than 40 tonnes were discarded. This leaves 157 respondents in the sample.
- **Carriers rail (including own account shippers)**
  Table 4 shows that we only have 17 interviews. This is due to the fact that there are not many firms providing rail freight services in the Netherlands. There were no implausible outliers.
- **Carriers air**
  For this mode there are 19 interviews, all with carriers (non-containers). One outlier interview with a base time of less than 1000 minutes but a base cost of more than 100,000 euro was removed, because this did not seem to be internally consistent.
- **Carriers inland waterways**
  The 85 respondents participated in either a quay experiment (46 respondents) or in a bridge/lock-experiment (39 respondents). After deleting two respondents with a missing base cost, five respondents with a base cost below 10 euro, one respondent with a base cost above 100,000 euro and one respondent with an unknown weight of the shipment (nor was the TEU known), we have 76 respondents left.

- **Carriers sea**

  The data set contains 30 respondents. 29 of them indicated that they (sometimes) had to wait before the quay and participated in a similar experiment. Only one of them indicated that he sometimes had to wait for a lock / bridge and did such an experiment. This respondent was removed from the data, since the value of waiting time before a bridge/lock might differ from the waiting time before a quay. One other respondent has been removed because he indicated a base cost of only 1 euro. This leaves 28 respondents in the data set.

- **Shippers (all modes)**

  Containerised transport: We interviewed 153 shippers that shipped their freight using containers. We excluded 8 road shippers: shippers with transport times equal to 10 minutes (1 respondent), with transport times more than 6 days (6 respondents) and with transport costs of 1 euro (1 respondent). We also excluded two inland waterways shippers: one respondent with a transport time of 30 minutes and one respondent with an unknown shipment weight. Finally, we excluded three sea shippers: two shippers with a transport time of 60 minutes or less and one shipper with an unknown shipment weight. So, the final sample contains 140 shippers using containers.

  Non-containerised transport: 296 shippers were interviewed that shipped their freight without using containers. Fifteen of them had a shipment cost of less than 10 euro (all modes). One respondent had an unknown shipment cost. Two road shippers had a transport time of more than 6 days and three sea shippers had a transport time of less than 60 minutes. One road shipper had a shipment weight of 422,000 tonnes and a further 25 respondents did not know their shipment weight. All these 47 respondents were excluded, leaving 249 shippers with a non-containerised typical transport in the final data set

In total 88 respondents were excluded (11%). The remaining numbers of respondents are given in Table 7.

Table 7: Number of respondents by target (sub)segment after selection

		Road	Rail	Air	Inland water-ways	Sea	Total
Container	**Carrier**	30	10	-	14	18	**72**
	Own account shipper	9	2	-	0	0	**11**
	Contract out shipper	33	14	-	16	77	**140**
Non-container	**Carrier**	125	5	18	62	10	**220**
	Own account shipper	32	0	0	0	0	**32**
	Contract out shipper	135	16	40	19	39	**249**
Total		**364**	**47**	**58**	**111**	**144**	**724**

4.2.3 Transport characteristics

The number of interviews collected did not allow estimation of separate models for each subsegment, i.e. for each cell in Table 7. Some subsegments were combined, other subsegments were split on the basis of the weight of the shipment. The final subsegments for which separate model were estimated and their number of respondents are displayed in Table 8, see Chapter 5 for a full description.

Table 8: Number of respondents by analysis segment (after selection)

	Road	Rail	Air	Inland water-ways	Sea
Carrier & own account shipper	Container 0-2t: 5 2-40t: 34 Non-cont. 0-2t: 42 2-15t: 65 15-40t: 50	17	18	Quay: 40 Lock/bridge: 36	Quay: 28
Contract out shipper	Container: 140 Non-cont.: 249				

For each of these 12 analysis segments, Table 9 displays the characteristics of the transport time (the so-called BaseTime), transport cost (BaseCost) and shipment weight.

Table 9: Transport time, transport cost and weight in the freight survey

		Carriers										Shippers	
		Road					Rail	Air	Inl.waterw.		Sea	All modes	
		Container		Non-container			All	All			Quay		
		0-2t	*2-40t*	*0-2t*	*2-15t*	*15-40t*			Quai	Lock / bridge		Cont.	Non-cont.
BaseTime	min	35	15	15	30	15	150	150	10	60	60	65	30
	max	480	2010	2880	6660	6300	10800	6360	1320	5940	2160	128610	59550
	median	60	180	80	165	150	2880	1650	60	480	75	15195	240
	average	139	336	382	464	424	3277	2318	127	1153	267	23613	1770
	stdev	191	453	783	947	988	2762	2155	211	1536	431	25714	4792
BaseCost	min	19	58	12	40	30	105	300	200	20	50	1	10
	max	450	1750	1300	2200	2000	23000	12000	34000	5545	250000	80000	50000
	median	60	271	85	350	310	1800	1750	2000	131	5000	1113	250
	average	143	415	185	479	404	4867	3431	3495	446	34243	3087	1587
	stdev	179	376	248	442	375	6600	3751	5402	1094	70224	8709	6029
Weight	min	0.02	2.5	0.01	2.5	18	5	0	100	28	0	1	0
	max	2	30	2	15	37	1680	10	25000	8000	85000	650	20000
	median	1	20	0.5	8	24.5	20	0.80	1000	1400	2000	18	3
	average	0.9	17	0.7	8	26	265	2.6	2057	1821	7541	40	182
	stdev	0.9	8.3	0.6	4	5	570	3.3	3958	1525	17143	103	1356

4.2.4 Trading

For each of the experiments, we checked whether respondents were trading the attributes when making their choices or whether they always chose the cheapest or the fastest trip. As can be seen from Table 10 only 12.4% of the respondents are non-traders in experiment 1. However, many of these respondents are trading time and cost in the other experiments, so, overall only 3.5% of the respondents are non-traders. Note that in experiments 2a and 2b the fraction of non-traders seems higher. However, in those experiments the cost or time attribute is sometimes equal in both alternatives, so the number of choices for which the time or cost attribute differs is less than six and therefore, non-trading occurs more often. Note that these non-traders have been kept in the analysis.

Table 10: Trading behaviour

	Exp. 1	Exp. 2a	Exp. 2b	All exp.
Always chose cheapest trip	10.0%	34.4%	37.8%	3.5%
Always chose fastest trip	2.4%	2.5%	5.4%	0.0%
Trading time/cost attributes	87.7%	63.1%	56.8%	96.5%

4.3 RP data

In their interviews, the carriers provided attribute values for the selected typical transport. Furthermore, they were asked whether there was a (realistic) alternative route for the typical transport (this was only done for carriers in road, rail and air transport; in the lock/bridge/quay experiments for inland waterways and sea transport, the carriers were not asked about an alternative route). If so, they were asked to provide the distance in km, the total transport time and within this total transport time the time spent in congested conditions.

Only 57 carriers (out of 188) answered positively to question on the availability of an alternative route. This RP data set therefore is much smaller than the SP data set, since we only have 1 observation per interview for the RP and many carriers have not provided an alternative route. One observation with a travel time for the alternative route of 1 minute was removed. The median travel time for the chosen route then is 330 minutes and for the alternative route it is 360 minutes.

The shippers were also asked to give base levels of time and cost for the typical transport for their firm. Additionally they were asked whether they could have used another mode for this transport. For those who affirmed, we asked which mode this would be and for an estimate of the transport time and transport cost.

In total only 59 shippers out of 450 answered all the questions on the alternative mode. The median transport time for the chosen mode is 3,600 minutes (about 2.5 days) and for the alternative mode it is 735 minutes. Often the chosen mode is non-road transport and the alternative mode is road transport. The median cost is 470 euro for the chosen and 495 for the alternative mode.

A somewhat larger sample could have been obtained by taking all observations on shippers, linking the origin and destination address to existing transport networks for road, rail and inland waterway transport, and calculating transport time and cost for the chosen and non-chosen modes by skimming those networks. This would be a major effort, which is probably not worthwhile because the number of observations would still be small (a few hundred at most). Moreover, one would be imposing the availability of other modes in situations where the shippers themselves regard these as non-available, and there would be a lack of precision (most existing freight transport models in the Netherlands use large zones, so that matching addresses to this is an error-prone process).

CHAPTER 5 Freight transport data analysis

5.1 Initial model tests

5.1.1 Initial model specifications

Initial models have been developed for all segments in a step-wise approach. The main conclusions from each step are discussed in this section.

We start by estimating multinomial logit (MNL) models because these are the simplest discrete choice models, which already contain all kinds of possibilities for extensive model specifications, and because such models were also estimated in the earlier Dutch national freight VOT studies of 1992 (Hague Consulting Group et al., 1992) and 2003/2004 (RAND Europe et al., 2004), so that we can compare the outcomes against previous results.

For these initial MNL models we used the combined mean-dispersion / scheduling specification as described in Section 2.1.4. This specification allows us to test to which degree the standard deviations or the early and late scheduling penalties are picking up the influence of variability: which specification is better and is there maybe a need to include both to capture the full effect of reliability of travel time? Can the scheduling terms pick up all effects by looking at the consequences of longer and shorter travel times on the arrival times or is there an additional influence that is related to stress, anxiety, image, etc.? In the end, all variability influences should be included in a reliability ratio, based on the standard deviation (see de Jong et al., 2009).

The standard deviation in equation [7] is taken to be the square root of the variance of the five possible transport times as presented in each alternative. This is the so-called population standard deviation:

$$\sigma^2 = \frac{1}{5} \cdot \sum_{i=1}^{5} \left(t_i - t_{mean}\right)^2 \qquad [24]$$

in which t_i is one of the five possible transport times and

$$t_{mean} = \frac{1}{5} \cdot \sum_{i=1}^{5} t_i \qquad [25]$$

The transport time in equation [7] is taken to be the "usual" transport time as presented in each alternative. This is not always the same as the mean transport time as calculated from equation [25]. Section 5.1.3 discusses the differences between these two times.

For the scheduling terms in equation [7] we calculated the early and late arrival terms with respect to the base arrival time. This is the time at which the respondent expected the transport to arrive when the transport began. This is also not necessarily the preferred arrival time (which is defined as the arrival time the respondent would choose if there would be no congestion). In this respect our model differs from the Vickrey/Small scheduling model that looks at scheduling penalties centred on the preferred arrival time. We selected the base arrival time here because the data that we obtained on the preferred arrival times contained some missing values and some implausible outliers and we did not want to exclude any more data. Moreover, in the SP experiments, the earlier and later arrivals are centred round the base arrival time. Also, for more than half of the respondents, the base time equals the preferred arrival time. Finally, we tested models with the preferred arrival time instead of the base time for carriers (road, rail, and air) and found that the results were very similar, but generally slightly better for the base time. The presented arrival time is the arrival time as presented in the SP (for which we have five values for each choice alternative in experiment 2a and 2b.).

Relative models are discussed in Section 5.1.4 and models in logWTP space are presented in Section 5.1.5. In Section 5.1.6, we will discuss the estimation of more complex models, especially mixed logit models. These types of models require more observations than MNL models, and even for larger samples do not always lead to proper convergence. The MNL models that we estimate first can then serve as a fall-back option in the case that more sophisticated models do not lead to significant improvements.

Most estimations were made using BIOGEME (Bierlaire, 2003 and Bierlaire, 2008), which allows for more flexible non-linear formulations than ALOGIT (ALOGIT Software and Analysis Ltd., 2007). Some models have also been estimated using ALOGIT. These estimation results were always identical to the BIOGEME results. Jack-knife estimations were done using ALOGIT (which contains special commands for Jack-knife estimation, unlike BIOGEME where one would have to specify a batch-job to do Jack-knife estimation), except for some models in logWTP space where we built our own Jack-knife procedure.

5.1.2 Initial MNL results

First, we estimated models for each experiment separately.

- Models estimated on data from experiment 1 only did not have statistically significant[13] VOTs for any segment. This is potentially worrying, but we have to keep in mind that this is only one third of the data. We also need to keep in mind that experiment 1 only has two attributes: time and cost, whereas the other experiments have more attributes. The previous freight VOT study (RAND

---

[13] When the term "significant" is used in this report, we always mean significant at the 95% confidence level, i.e. the t-ratio of a coefficient (which is its value divided by its standard deviation from the estimation) is either below -1.96 or above 1.96.

Europe et al., 2004) also contained more attributes in the SP than just time and costs.

- Experiment 2a gave significant VOTs for the segments Carriers-road and Shippers, which are the two segments with the largest number of respondents. For both segments, we also found significant values for the VOR as well. In both cases, the value for the VOT was comparable to the value of the VOR, which implies a reliability ratio close to 1.
- Estimates from experiment 2b were similar to those for experiment 2a. Here, the VOT and VOR of the Carriers-Air segment are just significant. The VOT for the Carriers-Inland waterways and Carriers-Sea are significant as well (recall that respondents in these segments did not participate in experiment 2a).
- In the next model, we pooled data from all three experiments. Scale factors were added to account for possible differences in error levels between the experiments (a higher scale means a lower variance for the observed component of utility). The scale of experiment 2a was set to 1. For Carriers-Inland waterways and Carriers-Sea, the scale of experiment 2b was set to 1. For the other segments, the scale of experiment 2b is never significantly different from the scale of experiment 2a, which is not surprising since the alternatives are presented with the same variables in both experiments.

  We find significant VOTs (before Jack-knifing) for Carriers-Road, Carriers-Inland waterways, Carriers-Sea, and Shippers. We also find significant VORs (before Jack-knifing) for Carriers-Road (non-container), Carriers-Air, and Shippers. These values are slightly higher than the VOTs. We did not find any significant scheduling terms.

5.1.3 Results for mean/median transport time and for mean/median scheduling terms

In SP experiments 2a and 2b (also see Chapter 3) both a “usual transport time” and five possible transport times were presented. For all respondents except for the Carriers-Inland waterways and Carriers-Sea (which were using a different questionnaire) the “usual transport time” did not equal the mean of the five possible transport times, rather it was the median (and also the statistical mode) of the five possible transport times. For application in a CBA framework, the mean is more suitable than the median. A median would need to be converted to a mean for application in CBA.

A similar issue holds for the arrival time: both the mean and median arrival time can be used in the utility function. In theoretical papers, usually the mean arrival time is used (as is the mean transport time).

In order to explore these issues, we first estimated a model based on data from experiment 2a only, using the mean travel time instead of the median travel time. When comparing the resulting estimates with those from the model that uses median transport time (as was used for the initial MNL models discussed in the previous subsection), we conclude that the VOT does not change, but that the VORs are much lower and in some cases they are no longer significant. This can be understood from the design. If an alternative contains the first (lowest) level of the reliability, the mean and median transport time are equal to each other. For higher levels, the mean and median start to deviate. The difference between the mean and median is highly correlated with the level of the reliability and hence, with the standard deviation of the five possible travel times.

So, what is real? It is not clear which representation (mean or median) will have driven the choices. Did respondents calculate the mean travel time from the five possible transport times, or did the presented usual travel time influence their choice?[14] The log-likelihood for the Carriers-Road segment improves by 0.6 points, indicating that a model using mean transport time predicts the choices (slightly) better. However, for Shippers, the log-likelihood deteriorates with 0.3 points. So, this is not conclusive. Since using the mean or median does not influence the VOT, we also cannot use the results from experiment 1 (in which this issue does not exist) as a reference. Note that this issue also exists in reality: the distribution of travel times is usually asymmetric with a tail towards longer times. So, it is also a matter of how you want to define the VOR: does this include or exclude this effect? In Section 5.2, we will make a final choice in this matter.

In a subsequent model, we also used mean arrival times to determine the Arriving Early and Arriving Late terms in the utility functions. Again, this does not influence the VOTs, but the VORs are reduced even further. None of the scheduling terms is significant.

We repeated this analysis for experiment 2b. When using median arrival times, both arrival times have equal levels for arriving early/late due to the design and they do not appear in the utility function. But when using mean arrival times, this is no longer true and coefficients can be estimated. A significant value for arriving late for the Carriers-Road segment is found, though its value (≈ 215 €/hr) is unrealistically high compared to the VOT (≈ 13 €/hr). For the Carriers-Inland waterways segment, a significant value for arriving early was found (t-ratio = 2.0, i.e. on the edge of being significant), but it has an unexpected positive sign. Furthermore, we note that in that segment only the possible waiting times were presented, but no departure and arrival times. So, it is less clear what the reference point for the arrival time is, and what should be called "early" and "late". Therefore, it is difficult to interpret these terms in the Carriers-Inland waterways and Carriers-Sea segments.

When combining the experiments using the mean transport times and mean scheduling terms, we reach similar conclusions. VOTs remain about the same, the VORs are about half their original size when using mean transport times, and they are about zero when using mean scheduling terms as well. Scheduling terms remain insignificant and the log-likelihood sometimes improves and sometimes deteriorates, but the differences are always small.

5.1.4 Results for relative models

In the next step, we estimated relative models, as explained in Section 2.2.

While we have the base costs and times, we did not ask the respondent about their base level of the standard deviation of their real transport times. Therefore, since there are five

---

[14] Note that the five possible transport times were presented before the "usual" transport time was presented, see Figure 3.

levels for the standard deviation in the design, we use the middle one as the base (reference) level.

Since we do not use scheduling terms in this relative model approach, we do not have to decide between using the mean or median arrival time. However, there remains an issue concerning the use of mean or median transport time. Therefore, we have tested both: using the median transport time and using the mean transport time.

When the relative model results are compared with those from the absolute model, we see enormous improvements in both the log-likelihood and the rho-squares for all segments, except for Carriers-Inland waterways where the improvement is relatively small, and for Carriers-Sea where the relative model is worse than the previous models. The time coefficients are now significant at the 95% level in all cases, except Carriers-Air where it is only significant at the 90% level. The coefficients on reliability are significant in all cases, except Carriers-rail and Carriers-Sea.

5.1.5 Results for logWTP models and prospect theory

We have not analysed the structure of the error-terms as De Borger and Fosgerau (2008) did in their logWTP analysis. This requires many extra parameters, and our sample size of freight data is too small for this. However, we re-estimated absolute models using either medians or means in logWTP space (see Section 2.4). Since the argument of a log-function must always be positive, the estimates for the VOT, VOR, Arriving Early and Arriving Late were constrained to be positive.

For the Carriers-Road, Carriers-Rail and Carriers-Air segments, the goodness-of-fit parameters improved by a large amount, but they were still not as good as for the relative models. For the Carriers-Inland waterways and Carriers-Sea segments, these new models also outperformed the relative models. For Shippers, we could not estimate a sensible logarithmic model since most coefficients had to be constrained to zero.

In addition, we tried a different model specification based on so-called pseudo-utilities for choosing the quickest and the cheapest alternative as were used by De Borger and Fosgerau (2008). This did not result in credible estimates, therefore we did not continue with this approach. Note that we could only use experiment 1 data for this test, since this method only works for a pure time/cost experiment.

Next, we tried to estimate separate values for WTP, WTA, EG and EL as suggested by prospect theory (see Section 2.5), but we were not successful. For none of the segments were all four values significant. And those that were significant were not significantly different from each other. It is likely that the data size for freight is too small to determine these effects. It may also be that this is due to the intrinsic difficulty to jointly estimate these parameters (Avineri & Bovy, 2008).

Models with a dependence on base time and cost (and on the time and cost changes offered in the SP, see Section 2.6) were also tested for freight transport as well but did not lead to significant improvements over the models without such effects.

5.1.6 Mixed logit results

Mixed logit models (see Section 2.7) were tried for the largest segments in freight transport (for the smaller segments, it was clear beforehand that there would not be enough observations for mixed logit estimation):

- Carriers in road transport – 196 respondents
- Shippers (all modes) – 389 respondents

For mixed logit estimation, even these segments are small. The models tested were specified in logWTP space. For the VOT and the VOR in the model specification, a statistical distribution was used. Several statistical distributions were tried including the one-sided uniform and the lognormal distributions, which always lead to a coefficient with the expected sign (unlike the normal distribution that can give negative VOTs). This did not lead to acceptable, stable models. This result was expected because of the small sample size. The number of observations in freight, even for the largest segments, is just not sufficient for a mixed logit model.

To account for the repeated measurements problem in the SP data (multiple observations on the same respondent, which in the standard logit model are assumed to be independent) and possibly other errors (such as heteroskedasticity, skewness), the Jack-knife method was applied.

5.1.7 Jack-knife procedure

A well-known problem with SP data, which is also relevant in this project, is that of repeated measurements. This means that the data contains several observations on the same respondent (firm) for several choice situations. Standard estimation of a MNL model on this data treats this variation in exactly the same way as variation between individuals. This leads to an overestimation of the t-ratios: variables seem to have a significant effect where this is sometimes not really the case. There are basically two ways of solving this problem:

- By including individual-specific components (stochastic terms that have the same value per individual); this can be included in a mixed logit model.
- By applying the Jack-knife procedure (or a related re-sampling method).

The Jack-knife method re-samples from the original sample by deleting a small number of observations each time. For each re-sample, statistics (e.g. estimated coefficients and standard errors) are calculated. The Jack-knife estimates are computed as averages of the re-sample statistics. This method is very computer-time intensive and has only been used for the best MNL models.

5.1.8 Results from the RP data

For the 57 RP observations on carriers we tried to estimate a MNL model for the choice between the observed route and the stated alternative route. One observation with a travel time for the alternative route of 1 minute was removed. Also after removing outliers in terms of transport distance and segmentation by base transport time, the number of observations proved insufficient for estimating significant time and distance (cost) coefficients.

Estimation of a binary mode choice model on the 59 observations for shippers did not lead to significant coefficients with the expected sign for time and cost. The same negative result was obtained after removing outliers in terms of time and cost. This sample is just too small for model estimation.

Since RP models could not be estimated on the RP data gathered, joint SP/RP estimation was not carried out either: the RP information does not enrich the SP information, so there is no point in adding the RP to the SP data for estimation.

5.1.9 Differences between containerised and non-containerised transport

The absolute model based on medians in WTP space, the absolute model based on means in WTP space, the relative model based on means, and the absolute model based on means in logWTP space were all estimated for containerised and non-containerised shipments separately.

For Carriers-Road, we find significantly higher values for containerised shipments. For Carriers-Rail, we could not find a significant difference, but this is probably related to the small sample size (with only 5 respondents in the non-containerised sub-segment). The same is true for Carriers-Sea. The relative models gave some indications for a different ratio time change versus cost change for the Carriers-Inland waterways segment.

5.2 Selecting the final model specification

In the previous section, various model specifications have been tried. It is clear that there is no approach that works best for all segments. Therefore, we have made the following decisions in the model development before proceeding:

1. We use MNL models, since the mixed logit estimates were not stable. A Jack-knife procedure will be applied to the final MNL models to correct the estimates and their t-ratios for so-called panel effects. This is the effect that each respondent has made multiple choices within one survey and these choices might be correlated.
2. We prefer models based on mean transport time over median transport time. This approach makes sure that the difference between mean and median transport time is included in the evaluation and therefore there is no problem in practical applications (such as in CBA's). It may underestimate the effect of uncertain transport times, but it certainly does not overestimate this effect (which a median transport time approach might do).
3. We do not use mean scheduling terms, since this completely removes any significant valuation of the standard deviation of possible transport times, but does not produce significant scheduling terms. We have seen that the scheduling terms are usually not significant and therefore, they will be constrained to zero in a later phase in the model development anyhow. In those cases, having mean or median scheduling terms is no longer an issue anyway.
4. Only the Carriers-Road and Shippers segments are large enough to determine separate VOTs and VORs for containerised and non-containerised transports. We have already seen indications that these types of transports have different VOTs.

5. We prefer the relative models for Carriers-Rail, Carriers-Air, Carriers-Inland waterways and Carriers-Sea. We know from discussions with future users of these valuations in practical CBA that it is important to have a VOT that depends on the weight of the shipment. It is known from practice that the VOT of a small shipment is much lower than the VOT for a large shipment. Given the large variation in shipment sizes for the modes, it is not very practical to use a uniform VOT for all shipments with a certain mode. The sample sizes in these segments are so low that it is not possible to properly estimate VOTs that are weight-dependent. Using a ratio between the value of a 1% time change versus a 1% cost change makes it possible to have different VOTs for different shipment sizes, since one can multiply this ratio with different levels for the cost per hour of transporting shipments of different size, to derive their VOT. The relative model was also used (for all modes) in the final results of the national freight VOT studies of 1992 and 2003/2004.
6. The Carriers-Road segment is large enough to estimate separate values for the VOT for some weight classes. The range of shipment sizes (measured as shipment weights) is limited in this segment (ranging from a small package to a full truck) and can suitably be categorised in two or three classes that can be used in practical applications (package, small truck, large truck). This leaves the choice between a model estimated in WTP space and in logWTP space. We will discuss this further below.
7. For shippers, we tried estimating models per mode, but the differences were not significant. Since the resulting models include the modes inland waterways, sea, air and rail, we have chosen relative models for shippers (as for the models for these modes for carriers).
8. The selected models are relative models for all modes except for road-carriers, where we use absolute models. This may seem inconsistent, but we need to take into account that for road transport there is much less heterogeneity in freight vehicle types than for all the other modes. Therefore for road transport (carriers), dealing with heterogeneity through relative models is less important. Moreover, for some of these other modes (e.g. inland waterways), more distinctions are required in the VOTs and VORs than can be given by models on our limited survey data. When using relative models, these distinctions can be brought in through the factor costs. Absolute models have the advantage that no data on factor costs are required to obtain VOTs and VORs (and these are also the models used in passenger transport).

5.3 Synthesis per segment

5.3.1 Carriers-Road (including own account shippers)

Based on the decisions listed above, we used the model using mean travel time and median scheduling terms as the base model and split the data set into three subsegments:

- containers with 2 tonnes < shipment weight;
- non-containers with 2 < shipment weight ≤ 15 tonnes;
- non-containers with 15 < shipment weight ≤ 40 tonnes.

For containers, splitting between the two weight categories used for non-containers did not lead to a significant improvement of the loglikelihood value.

The two segments (container and non-container, i.e. 47 respondents in total) with shipment weights ≤ 2 tonnes are disregarded, since for CBA we need values for complete vehicles. Moreover, for typical transports below 2 tonnes, the time and reliability coefficients were not significant (only cost was). Shipments > 40 tonnes had already been excluded. So we are left with shipments that correspond to the size of a small or large truck.

We estimated absolute models (the attributes are in euros or minutes), both in utility or WTP space and in logWTP space. The logWTP space models performed clearly better in terms of loglikelihood value and were selected. Insignificant variables have been removed and the models were re-estimated without these.

As noted earlier, initial analysis revealed that the carrier and own account VOTs and VORs were not significantly different. Although this is contrary to the expectations in Table 1, it is probably due to the small sample for the latter group. Consequently these groups were combined in the analysis: road own account shippers are treated as carriers.

The commodity type, value of the goods and the value of density of the goods (value per unit of weight) were all tried as segmentation variables, but no clear patterns emerged. When estimating different models for raw and intermediate goods versus final products, some of the VOT and VOR estimates were not significant, and others were clearly less significant than when distinguishing by weight class and container/non-container. The same result was obtained when distinguishing by distance class (below and above 100 km.) This also goes for segments distinguishing whether the goods had to be delivered at a specific time (or within a specific time interval) or not.

Finally, the sample was segmented into a number of classes relating to the observed time (base time) and cost (base cost). In addition, segmentation based on the ratio of base time and base cost was tested, in order to see if differences in cost per unit of time matter. However, this did not lead to significant differences and intuitive values for the VOT and VOR.

The most important segmentation variables clearly were the weight of the shipment and container/non-container. Of course, the costs of providing transport services are related to the shipment weight, though the relationship is not purely linear. With increasing shipment weight, the transport costs (especially the vehicle-related cost) go up every time a larger vehicle (or higher number of vehicles, but this possibility has been ruled out here) is required to transport the goods. The unit costs (transport costs per tonne) however go down with increasing shipment size. So we can expect that the transport cost related component of the VOT will increase with shipment size (weight in tonnes), but less than proportionally.

The final models are shown in middle columns of Table 11 to Table 13. All estimated coefficients are clearly significant. The t-ratios below (from BIOGEME) are so-called 'robust' t-ratios, which allow for non-severe misspecification errors (Bierlaire, 2008). To correct for repeated measurements here (we have no less than eighteen observations for the same respondent, which are likely to be correlated) the right columns contain the estimated coefficients after the Jack-knife procedure has been applied. Since BIOGEME contains no readily available set-ups for Jack-knife estimation, and since ALOGIT (that does have such set-ups) does not allow for the estimations in logWTP space, we performed the Jack-knife procedure ourselves by re-estimating each segment with N respondents N times, each time leaving out one respondent. From the N estimates for each coefficient, a biased-corrected estimate with standard error can be calculated (see DeTar 2002 for an overview of equations).

The estimation directly yields VOTs and VORs in euro/hour. The VOT is clearly higher for the heavier segment (non-containers). The VOR was only significant for non-containers/2-15 tonnes. The reliability ratio (reliability coefficient to time coefficient, or VOR to VOT) here is 1.6. A further discussion on the VOTs and VORs will follow later.

Table 11: Estimated coefficients and t-ratios (in brackets) for MNL model for carriers and own account shippers in road transport using containers

Segment	Road – container Truck 2 -40 tonnes		Road – container Truck 2 -40 tonnes Jack-knife	
File name	CaRoa111-23.F12		CaRoa111-23.j12	
Experiments used	1, 2a and 2b		1, 2a and 2b	
Observations	612		612	
Respondents	34		34	
Final log (L)	-347.0		-347.0	
D.O.F.	2		2	
Rho²(0)	0.182		0.182	
	Value	(T-ratio)	Value	(T-ratio)
Lambda (Cost)	-11.93	(-9.6)	-11.69	(-6.3)
VOT	47.11	(7.0)	45.97	(3.2)

Note:

- Utility used:

$$U = \lambda \cdot \log(C + VOT \cdot T)$$

- VOT is the monetary value of a change of one hour in transport time, in Euro per vehicle.

Table 12: Estimated coefficients and t-ratios (in brackets) for MNL model for carriers and own account shippers in road transport, non-containers, 2 – 15 tonnes

Segment	Road – non-container Truck 2 - 15 tonnes	Road – non-container Truck 2 - 15 tonnes Jack-knife
File name	CaRoa111-5d.F12	CaRoa111-5d.j12
Experiments used	1, 2a and 2b	1, 2a and 2b
Observations	1170	1170
Respondents	65	65
Final log (L)	-683.6	-683.6
D.O.F.	3	3
Rho²(0)	0.156	0.156
	Value (T-ratio)	Value (T-ratio)
Lambda (Cost)	-8.938 (-10.8)	-8.747 (-6.2)
VOT	19.14 (3.7)	18.49 (2.6)
VOR	30.66 (4.6)	29.62 (2.6)
	Derived value	Derived value
Reliability ratio	1.60 (2.8)	1.60 (1.8)

Note:

- Utility used:

$$U = \lambda \cdot \log\left(C + VOT \cdot T + VOR \cdot \sigma\right)$$

- VOT is the monetary value of a change of one hour in transport time, in Euro per movement.
- VOR is the monetary value of a change of an hour in the standard deviation of transport time, in Euro per movement

Table 13: Estimated coefficients and t-ratios (in brackets) for MNL model for carriers and own account shippers in road transport, non-containers 15- 40 tonnes

Segment	Road – non-container Truck 15 - 40 tonnes	Road – non-container Truck 15 - 40 tonnes Jack-knife
File name	CaRoa111-6c.F12	CaRoa111-6c.j12
Experiments used	1, 2a and 2b	1, 2a and 2b
Observations	900	900
Respondents	50	50
Final log (L)	-517.6	-517.6
D.O.F.	2	2
Rho²(0)	0.170	0.170
	Value (T-ratio)	Value (T-ratio)
Lambda (Cost)	-11.03 (-10.7)	-10.79 (-6.7)
VOT	38.24 (6.1)	36.87 (3.3)

Notes, see previous table.

For expansion of the sample used in estimation, Ton and Tavasszy (2010) proposed to use commodity class (especially distinguishing NSTR Chapter 9: other goods) and distance class. We do not have the NSTR code of the typical transport (only a description from the respondent himself/herself) and the distinction between raw materials, intermediate products and final products. Final products is the category that comes closest to NSTR9 (other goods), though it is a much broader category. As discussed above, our selected model does not include a distinction by commodity type or distance class (this leads to less significant VOTs and VORs, sometimes even insignificant ones). Given that our best estimates for road transport (carriers and own account shippers) do not distinguish by type of good and distance class, expansion of the estimation results in terms of these variables will not be necessary.

5.3.2 Carriers-Rail transport

Segmenting on the basis of observed weight or value density did not lead to significant VOTs and VORs. The best model was a relative model (cost and time expressed as relative to their base level), using mean travel time (not median). This model has significant time and cost coefficients. Reliability and scheduling variables were not significant and removed from the final model. The implied trade-off ratio that can be calculated from the time and cost coefficients is 0.33. This means that a 10% increase in transport time is for the rail carriers equivalent to 3.3% increase in total transport costs. This ratio can be multiplied by the transport cost per hour for a wagon or train (separately, if required, by type of train) to give the value of time (for the carriers: the part of the VOT that is related to supplying transport services) for a wagon or train per hour.

In the Jack-knife method (using ALOGIT) we used the maximum number of repetitions (equal to the number of respondents), each time we use a subsample where one respondent has been removed. So since we have 17 respondents, we estimate 17 models, each on a different sample of 16 respondents. As expected, the coefficient values do not change much as a result of the Jack-knife and only the t-ratios go down. The trade-off ratio remains unchanged.

Table 14: Estimated coefficients and t-ratios (in brackets) for MNL model for carriers and own account shippers in rail transport

Segment	Rail		Rail – Jack-knife	
File name	CaRai205.F12		Rail202.j12	
Experiments used	1,2a and 2b		1,2a and 2b	
Observations	306		306	
Respondents	17		17	
Final log (L)	-157.7		-157.7	
D.O.F.	2		2	
Rho^2(0)	0.257		0.257	
	Value	(T-ratio)	Value	(T-ratio)
BetaCost (relative)	-10.24	(-7.5)	-9.742	(-4.6)
BetaTime (relative)	-3.341	(-3.6)	-3.157	(-2.6)
	Derived value		Derived value	
Trade-off ratio time vs cost	0.326	(4.2)	0.324	(3.3)

Note:

- Utility used:

$$U = \beta_C^{rel} \cdot \frac{C}{C_0} + \beta_T^{rel} \cdot \frac{T}{T_0}$$

- Relative Cost: impact of a change in cost (relative to base cost) on utility
- Relative Time: impact of a change in time (relative to base time) on utility

5.3.3 Carriers-air transport

Here too, the relative model (using mean travel time) worked best. The scheduling variables are not significant, but the VOTs and VORs are (before correcting for repeated measurements). Here we obtain an implied trade-off ratio for time versus cost of 0.43 and one for reliability versus cost of 0.11 (also reliability versus time: 0.26).

In the Jack-knife procedure (with 18 runs), the coefficient on the reliability becomes insignificant. This coefficient was removed, and the Jack-knife was redone without reliability. This has some impact on the trade-off ratio.

Table 15: Estimated coefficients and t-ratios (in brackets) for MNL model for carriers in air transport

Segment	Air		Air – Jack-knife	
File name	CaAir304.F12		Air1305.j12	
Experiments used	1, 2a and 2b		1, 2a and 2b	
Observations	324		324	
Respondents	18		18	
Final log (L)	-200.1		-205.3	
D.O.F.	3		2	
Rho²(0)	0.109		0.086	
	Value	(T-ratio)	Value	(T-ratio)
BetaCost (relative)	-4.725	(-5.3)	-4.533	(-3.0)
BetaTime (relative)	-2.044	(-2.3)	-2.830	(-3.0)
BetaRel (relative)	-0.535	(-2.8)		
	Derived value		Derived value	
Trade-off ratio time vs cost	0.433	(2.7)	0.624	(3.1)
Trade-off ratio reliability vs cost	0.113	(2.7)		

Note:

- Utility used:

$$U = \beta_C^{rel} \cdot \frac{C}{C_0} + \beta_T^{rel} \cdot \frac{T}{T_0} + \beta_R^{rel} \cdot \frac{\sigma}{\sigma_0}$$

- Relative Cost: impact of a change in cost (relative to base cost) on utility
- Relative Time: impact of a change in time (relative to base time) on utility
- Relative Reliability: impact of a change in reliability (relative to base reliability) on utility

5.3.4 Carriers-inland waterways transport

For inland waterways we had 480 observations for SP experiments in the context of waiting for a quay and 432 in the context of waiting for a lock/bridge. These two experiments give clearly different trade-off ratios and were kept separate in the final models. So, we obtain a VOT for waiting for a quay and a VOT for waiting for a lock/bridge. For CBA for inland waterways, VOTs are required for several vessel types. The SP survey is too small to provide so many VOTs. But we could successfully estimate a relative model (using mean travel time) for all vessel types together, which provides trade-off ratios that can be applied to different transport costs figures for different vessel types.

The time and cost coefficients are clearly significant, but the reliability coefficient was not significant.

The Jack-knife has no impact on the trade-off ratios.

Table 16: Estimated coefficients and t-ratios (in brackets) for MNL model for carriers in inland waterway transport

Segment	Inland waterways - waiting for a quay		Inland waterways - waiting for a quay – Jack-knife	
File name	CaIWW404.F12		IWW1402.j12	
Experiments used	1 and 2b		1 and 2b	
Observations	480		480	
Respondents	40		40	
Final log (L)	-308.7		-308.7	
D.O.F.	3		3	
Rho²(0)	0.072		0.072	
	Value	(T-ratio)	Value	(T-ratio)
BetaCost (relative)	-23.11	(-6.1)	-22.75	(-3.9)
BetaTime (relative)	-2.854	(-4.7)	-2.840	(-4.1)
Scale experiment 1	0.182	(0.7)	0.147	(0.5)
	Derived value		Derived value	
Trade-off ratio time vs cost	0.124	(5.4)	0.125	(4.5)

Note:

- Utility used:

$$U = \beta_C^{rel} \cdot \frac{C}{C_0} + \beta_T^{rel} \cdot \frac{T}{T_0}$$

- Relative Cost: impact of a change in cost (relative to base cost) on utility
- Relative Time: impact of a change in time (relative to base time) on utility

Table 17: Estimated coefficients and t-ratios (in brackets) for MNL model for carriers in inland waterway transport

Segment	Inland waterways - Waiting for lock/bridge		Inland waterways - Waiting for lock/bridge – Jack-knife	
File name	CaIWW405.F12		IWW1412.j12	
Experiments used	1 and 2b		1 and 2b	
Observations	432		432	
Respondents	36		36	
Final log (L)	-251.1		-251.1	
D.O.F.	3		3	
Rho²(0)	0.162		0.162	
	Value	(T-ratio)	Value	(T-ratio)
BetaCost (relative)	-6.298	(-5.4)	-5.952	(-3.4)
BetaTime (relative)	-5.840	(-7.8)	-5.709	(-7.2)
Scale experiment 1	0.204	(1.3)	0.236	(1.4)
	Derived value		Derived value	
Trade-off ratio time vs cost	0.927	(7.0)	0.959	(4.6)

Notes, see previous table.

5.3.5 Carriers-Sea transport

The best results for sea transport carriers were obtained with a relative model (using mean travel time). Here we only use SP experiments for waiting for a quay (there were not enough observations for a model on waiting for a lock/bridge). Time and cost are significant, but reliability was not. Again, the coefficient values do not change much as a result of the Jack-knife and only the t-ratios go down.

Table 18: Estimated coefficients and t-ratios (in brackets) for MNL model for carriers in sea transport

Segment	Sea – waiting for a quay		Sea – waiting for a quay – Jack-knife	
File name	CaSea502.F12		Sea1502.j12	
Experiments used	1 and 2b		1 and 2b	
Observations	336		336	
Respondents	28		28	
Final log (L)	-212.0		-212.0	
D.O.F.	3		3	
Rho²(0)	0.090		0.090	
	Value	(T-ratio)	Value	(T-ratio)
BetaCost (relative)	-4.843	(-4.9)	-4.829	(-3.2)
BetaTime (relative)	-2.859	(-4.0)	-2.716	(-2.8)
Scale experiment 1	0.284	(0.9)	0.340	(1.0)
	Derived value		Derived value	
Trade-off ratio time vs cost	0.590	(4.1)	0.563	(3.1)

Note:
- Utility used:

$$U = \beta_C^{rel} \cdot \frac{C}{C_0} + \beta_T^{rel} \cdot \frac{T}{T_0}$$

- Relative Cost: impact of a change in cost (relative to base cost) on utility
- Relative Time: impact of a change in time (relative to base time) on utility
- T-ratio scale coefficient is with respect to zero

5.3.6 Shippers that contract out

The best model here was a relative model (using mean travel time) with a segmentation by container versus non-container. Other segmentations would also have been possible (modes, base time classes), but not in combination with that by container/non-container. Distinguishing by raw and intermediate goods versus final products, or between long and short distance (with a threshold at 100 km) did not lead to plausible VOT results. So a re-weighting of the VOT and VOR outcomes by the importance of these segments in the population, as suggested by Ton and Tavasszy (2010), is not worthwhile.

The VOTs are much lower than those for carriers. This finding is in line with our assumption that for shippers that contract out the VOT is cargo-related only. The trade-off ratio for reliability versus time for non-containers is close to 1, but for containers it is only 0.3 (the difference between the trade-off ratios with cost is smaller: 0.09 versus 0.06).

Table 19: Estimated coefficients and t-ratios (in brackets) for MNL model for shippers that contract out (non-container)

Segment	Shippers – non-container		Shippers – non-container – Jack-knife	
File name	ShAll610.F12		Shipper1612.j12	
Experiments used	1, 2a and 2b		1, 2a and 2b	
Observations	4482		4482	
Respondents	249		249	
Final log (L)	-2623.7		-2623.7	
D.O.F.	5		5	
Rho²(0)	0.155		0.155	
	Value	(T-ratio)	Value	(T-ratio)
BetaCost (relative)	-7.026	(-15.3)	-6.992	(-13.1)
BetaTime (relative)	-0.709	(-3.2)	-0.706	(-2.7)
BetaRel (relative)	-0.639	(-8.1)	-0.634	(-5.7)
Scale experiment 1	0.558	(8.5)	0.556	(8.5)
Scale experiment 2b	1.293	(11.2)	1.293	(9.2)
	Derived value		Derived value	
Trade-off ratio time vs cost	0.101	(3.3)	0.101	(2.8)
Trade-off ratio reliability vs cost	0.091	(9.6)	0.091	(6.6)
Trade-off ratio reliability vs time	0.901	(2.8)	0.898	(2.2)

Note:

- Utility used:

$$U = \beta_C^{rel} \cdot \frac{C}{C_0} + \beta_T^{rel} \cdot \frac{T}{T_0} + \beta_R^{rel} \cdot \frac{\sigma}{\sigma_0}$$

- Relative Cost: impact of a change in cost (relative to base cost) on utility
- Relative Time: impact of a change in time (relative to base time) on utility
- Relative Reliability: impact of a change in reliability (relative to base reliability) on utility
- T-ratio scale coefficient is with respect to zero

Table 20: Estimated coefficients and t-ratios (in brackets) for MNL model for shippers that contract out (container)

Segment	Shippers – container		Shippers – Container – Jack-knife	
File name	ShAll613.F12		Shipper1602.j12	
Experiments used	1, 2a and 2b		1, 2a and 2b	
Observations	2520		2520	
Respondents	140		140	
Final log (L)	-1380.0		-1379.9	
D.O.F.	4		4	
Rho²(0)	0.210		0.210	
	Value	(T-ratio)	Value	(T-ratio)
BetaCost (relative)	-10.30	(-17.0)	-10.19	(-11.5)
BetaTime (relative)	-2.068	(-4.3)	-2.043	(-3.2)
BetaRel (relative)	-0.635	(-6.1)	-0.629	(-4.2)
Scale experiment 1	0.463	(7.8)	0.464	(7.3)
Scale experiment 2b				
	Derived value		Derived value	
Trade-off ratio time vs cost	0.201	(4.5)	0.200	(3.4)
Trade-off ratio reliability vs cost	0.062	(6.3)	0.062	(4.5)
Trade-off ratio reliability vs time	0.307	(3.1)	0.307	(3.1)

Notes, see previous table.

5.4 Final VOT results

5.4.1 The new VOTs

For the VOT (and VOR) to be used in CBA, we select the results from the Jack-knife models, since these include proper corrections for repeated measurements. An overview of the results is given in Table 21. When calculating the VOT in freight transport it is important to distinguish between shippers who contract out and carriers/own account shippers. The VOT for shippers who contract out is only cargo-related. That is to say, it takes into account interest, deterioration, disruption of production and out of stock costs. Costs related to vehicles and staff are not part of these factors. In order to compare the values to those found in 2003/2004, VOTs for carrier and shipper (own account) should be added together (see Section 3.1 and also TNO-Inro and MuConsult, 2002). This has been done in the 'Total' rows in Table 21.

For many segments in which we estimated a relative model, the final estimation result was a trade-off ratio. These need to be multiplied by the transport costs per hour (factor costs). This implies that in order to obtain monetary values for the VOT and VOR in freight transport for use in CBA, recent factor cost calculations for all modes are required. The derivation of factor cost was not part of this project. The new survey data gathered in this project on the observed typical transport can give some indication, but the sample sizes by mode are generally not sufficient to derive proper factor costs.

Table 21: Partial value of time (in 2010 euro/shipment/hour) and trade-off ratios (TR) for time vs cost

		Road	Rail	Air	Inland waterways	Sea
Container	**Carrier / own account shipper**	*[2-40t truck]:* 45.97	*[full train]:* TR=0.32	*Not applicable*	*[ship waiting for a quay]:* TR=0.13 *[ship waiting for a lock/bridge]:* TR=0.96	*[ship waiting for a quay]:* TR=0.56
	Shipper*	*[All]:* TR=0.20	*[All]:* TR=0.20	*Not applicable*	*[All]:* TR=0.20	*[All]:* TR=0.20
	Total	*[2-40t truck]:* 45.97 + 0.20*factor cost	*[full train]:* TR=0.52	*Not applicable*	*[ship waiting for a quay]:* TR=0.33 *[ship waiting for a lock/bridge]:* TR=1.16	*[ship waiting for a quay]:* TR=0.76
Non-container	**Carrier / own account shipper**	*[2-15t truck]:* 18.49 *[15-40t truck]:* 36.87	*[full train]:* TR=0.32	*[full freighter aircraft]:* TR=0.62	*[ship waiting for a quay]:* TR=0.13 *[ship waiting for a lock/bridge]:* TR=0.96	*[ship waiting for a quay]:* TR=0.56
	Shipper*	*[All]:* TR=0.10	*[All]:* TR=0.10	*[All]:* TR=0.10	*[All]:* TR=0.10	*[All]:* TR=0.10
	Total	*[2-15t truck]:* 18.49 + 0.10* factor cost *[15-40t truck]:* 36.87 + 0.10*factor cost	*[full train]:* TR=0.42	*full freighter aircraft]:* TR=0.72	*[ship waiting for a quay]:* TR=0.23 *[ship waiting for a lock/bridge]:* TR=1.06	*[ship waiting for a quay]:* TR=0.66
All		*[2-15t truck]:* 19.98 + 0.11* factor cost *[15-40t truck]:* 37.42 + 0.11*factor cost *:*	*[full train]:* TR=0.46	*full freighter aircraft]:* TR=0.72	*[ship waiting for a quay]:* TR=0.24 *[ship waiting for a lock/bridge]:* TR=1.07	*[ship waiting for a quay]:* TR=0.68

* Shipper = shipper that contracts out

In the 2003/2004 freight VOT study, the factor costs were supplied by a study carried out for DVS (NEA et al., 2003). For all modes, except air transport, there are recent updates or revised calculations available from the Kostenbarometer (NEA, 2011) carried out for DVS. These factor costs are in Table 22. Please note that for inland waterways and sea transport the method of factor cost calculation changed considerably, so that the values for 2002 are also different now from NEA et al. (2003). For air transport, we used the value from (NEA et al. (2003) and used the same percentage change as we use for road transport to obtain the 2009 value. To go from 2009 to 2010 values (the trade-off ratios from the SP refer to 2010) we used the general price index 2009/2010 (+1.3%). Note that the factor costs do not include VAT; so, the VOTs we calculate in this report also do not include VAT.

Combining Table 21 and Table 22, we obtain the VOTs in Table 23:

Table 22: Factor cost (in 2009 euro) per vehicle or vessel per hour (taken from NEA 2003, 2011)

	Road	Rail	Air	Inland waterways	Sea
Container	2002: 50 *(NEA, 2003)* 2002: 51 *(NEA, 2011)* 2009: 62 *(NEA, 2011)*	2002: 941 *(NEA, 2003)* 2002: 963 *(NEA, 2011)* 2009: 1668 *(NEA, 2011)*	*Not applicable*	2002: 212 *(NEA, 2003)* 2002: 261 *(NEA, 2011)* 2009: 293 *(NEA, 2011)*	2002: 588 *(NEA, 2003)* 2002: 733 *(NEA, 2011)* 2009: 984 *(NEA, 2011)*
Non-container	*[2-15t truck]:* 2002: 37 *(NEA, 2003)* 2002: 38 *(NEA, 2011)* 2009: 46 *(NEA, 2011)* *[15-40t truck]:* 2002: 52 *(NEA, 2003)* 2002: 53 *(NEA, 2011)* 2009: 66 *(NEA, 2011)* *[all non-container]* 2002: 47 *(NEA, 2003)* 2002: 48 *(NEA, 2011)* 2009: 59 *(NEA, 2011)*	*[bulk]* 2002: 1460 *(NEA, 2003)* 2002: 1485 *(NEA, 2011)* 2009: 2874 *(NEA, 2011)* *[wagonload train]* 2002: 1360 *(NEA, 2003)* 2002: 1376 *(NEA, 2011)* 2009: 2617 *(NEA, 2011)* *[all non-container]* 2002: 1425 *(NEA, 2003)* 2002: 1444 *(NEA, 2011)* 2009: 2777 *(NEA, 2011)*	2002: 14132 *(NEA, 2003)* 2009: 17740 *(NEA, 2003)*	2002: 201 *(NEA, 2003)* 2002: 228 *(NEA, 2011)* 2009: 278 *(NEA, 2011)*	2002: 766 *(NEA, 2003)* 2002: 933 *(NEA, 2011)* 2009: 1245 *(NEA, 2011)*

Notes:

- The values for rail are for a train (not a wagon).

Table 23: Values of time for freight transport (Euro/hour per vehicle or vessel, price level 2010)

	Road	Rail	Air	Inland waterways	Sea
Container	*[2-40t truck]:* 59	*[full train]:* 880	*Not applicable*	*[ship waiting for a quay]:* 98 *[ship waiting for a lock/bridge]:* 340	*[ship waiting for a quay]:* 760
Non-container	*[2-15t truck]:* 23 *[15-40t truck]:* 44 *[all non-container]:* 37	*[bulk]:* 1200 *[wagonload train]:* 1100 *[all non-container]:* 1200	*[full freighter aircraft]:* 13000	*[ship waiting for a quay]:* 65 *[ship waiting for a lock/bridge]:* 300	*[ship waiting for a quay]:* 830
All	*[2-40t truck]:* 38	*[full train]:* 1100	*[full freighter aircraft]:* 13000	*[ship waiting for a quay]:* 69 *[ship waiting for a lock/bridge]:* 300	*[ship waiting for a quay]:* 820

Notes:

- All these values are combined values from shippers and carriers and were obtained after rounding off.
- The values for rail are for a train (not a wagon).
- The values for inland waterways and sea refer to a ship.
- The values for “all” types of transport are the average of the values for container and non-container weighted by the number of tonnes transported. The weight factors container / non-container are 7% - 93% for road, 35% - 65% for rail, 13% - 87% for inland waterways and 19% - 81% for sea transport.
- These values do not include VAT.

For road-non-container, we have VOTs for two weight categories (which broadly correspond to solo truck and road transport combination). If in a CBA, a single road transport VOT was needed, these two values can be weighted by the number of vehicle km by each type of road vehicle, and similarly to combine the values for container and non-container road transport. The VOTs we obtain for road transport, which have not been derived using factor costs, are quite close to the factor costs (as in Table 22). Not all transport costs are transport time-related (the carrier component of the VOT is somewhat smaller than the factor cost), but when one adds the cargo-related component from the shipper, the sum for the VOT comes close to the full factor cost.

In Table 21 we see that the contribution of the shippers to the VOT (which is related to the cargo, see Table 1) is usually much smaller than the contribution of the carrier (which is related to the costs, such as those of staff and vehicles, of providing transport services). We get a cargo component share in the VOT of 15-22% for road transport and between 10% and 60% for the other modes. This is in line with prior expectations (e.g. Bruzelius, 2001; de Jong, 2008).

In Norway, a national freight VOT and VOR study was carried out (Halse et al., 2010), which used SP experiments very similar to ours (partly based on Significance et al., 2007). The distinction between the cargo component in the VOT and VOR, which can be obtained from the shippers, and the transport vehicles and staff component, which can be asked from the carriers, was not made so explicitly as in the Dutch survey, but the Norwegian results seem to be in line with this hypothesis. We interpret the Norwegian values for shippers that contract out as representing the "goods" component and those for carriers as representing the '"vehicles and staff" component. The Norwegian study found a shipper's share of 12% of the shipper plus carrier model-estimated VOT for road transport and 14% for all modes. The ratio of the shippers VOT to the recommended VOT for Norway (which is a factor cost value, and somewhat higher than the model-estimated VOT) is around 10%.

For inland waterway transport we find that the VOT for waiting at a lock (or bridge) is almost equal to the total transport costs per hour (even slightly higher if we include the VOT related to the cargo), which is considerably higher than the trade-off ratio we found in the 2003/2004 survey (0.37). Given that the time spent at a lock or bridge should be regarded as unproductive, it makes sense that changes in waiting time at such locations would be valued at the full factor cost, as was our hypothesis in Table 1). The VOTs for waiting at a quay to be loaded/unloaded are much lower; presumably this time can easily be factored into the schedule for inland shipping, and is not regarded as fully unproductive. For rail, air and sea transport we get a total trade-off ratio between costs and time between 0.46 and 0.68.

5.4.2 Discussion on the use of factor costs and trade-off ratios

Some of the trade-off ratios found are substantially lower than 1. This means that the value of a time gain is considerably lower than the factor cost of an hour. A discussion on the use of the factor costs method versus (SP and/or RP) models for obtaining values of time for use in CBA can be found in the reports of the first national Dutch freight value of time study (Hague Consulting Group, 1991, 1992). In these reports it is argued that value of time research in freight transport needs to find the “time-marginal transport cost”: the transport costs that will change when the amount of transport time changes. This is the derivative of the total logistics cost function with respect to transport time (the standard marginal cost approach is about the derivative with respect to a unit of transport services, say measured in tonne-kilometres). The total logistics costs consists of transport staff cost (e.g. truck drivers), energy costs (e.g. diesel), vehicle costs, overhead costs (e.g. office space and administrative staff of the carrier firm), which are all costs that carriers incur, but also of the deterioration of the goods, the interest costs on the value of the goods during transport and the costs of having a reserve stock for safety (the last three items then relate to the cargo component of the VOT). The factor cost used in the previous tables and the transport cost in the SP only refer to the costs of the carriers (the transport costs). Therefore, when including the cargo component in the value of time, the trade-off ratio taken relative to the transport cost may in principle exceed 1. For most commodities however, deterioration, interest and safety stocks will be very limited.

Before 1992, freight values of time in The Netherlands had to be based on the factor cost and various assumptions were used regarding the costs that should be included in the value of time: only the transport staff cost and the fuel costs (e.g. McKinsey, 1986) or all transport costs minus overheads (NEA, 1990).

An argument for not including energy (fuel) costs savings in the VOT is that most transport projects nowadays are carried out to reduce congestion, not to reduce transport distances: there are time gains, but the project does not change the fuel costs (and even if a project leads to shorter routes, it may be better to evaluate these fuel costs benefits separately, as is done in the UK, and not include these through the time gains).

It can happen in practice that the trade-off ratio for transport time versus transport costs can be smaller than 1, because it may be difficult for firms to convert the time gains fully into cost reductions or additional revenues. The time gain for instance could be too small to use for other transport activities, or additional work for a transport firm can only be realised against high costs (marketing, discounts), taking into account that the volume of transport services is not very price elastic (because the demand for transport largely depends on product markets). Furthermore there are regulations in the opening times of firms at the origin and destination, on driving and on sailing times and on labour contracts, that prevent full flexibility in using time gains productively for other transports or for reducing costs. In the longer run, which is the proper perspective for CBA of transport infrastructure, there will be more possibilities for reorganising logistics and therefore to reduce costs or increase output to benefit from time savings.

The imperfect flexibility (or kinked production function or cost function) argument could be more relevant for train, inland waterways and sea transport, since these modes have much larger indivisibilities (large vehicle and vessels that are used for trips that take a long time, possibly also with slot allocation). Also for the products transported using these modes, which generally have a lower value per tonne than products transported by road and air transport, the cargo component in the VOT will be relatively small.

Therefore in the long run we expect that the trade-off ratio for road transport will not be far below 1. Those for other modes may be somewhat smaller, but in the long run these too should not be very far from 1 (in the 2004 SP for instance, the trade-off ratio for sea transport was 0.16, which clearly is too low).

From the current survey we now obtain a trade-off ratio (after weighting for the shares of container and non-container by mode) for road transport of 0.65, rail 0.46 and air 0.72 (the latter value is not significantly different from 1; the others are). Given the above, these seem plausible values, though the value for rail is rather at the low end. The TRs for inland waterways and sea transport come from very different SP experiments, and refer to a somewhat different setting: that of waiting at a lock, bridge or quay. For these comparisons we have found TRs of 0.24 and 1.07 (inland waterways: quays and locks/bridges respectively) and 0.68 (maritime, quays). Here it seems prudent to take the average for the values for inland waterways, so that we obtain a value of 0.66 (for quays, locks and bridges), close to the quay value for sea transport.

A related question is whether Stated Preference is capable of providing the long run cost savings in freight transport that arise in case of time gains. In general, SP is more oriented to the short and medium run, because respondents may find it hard to imagine circumstances very different from the current situation, which is used to customise the SP experiment. In our freight VOT survey, we include as explanation just before the choice tasks that the changes in time, costs and reliability are generic: these apply to all carriers using the same infrastructure, and are not competitive advantages for their specific firm. This should also make clear that the time savings do not only relate to the shipment that is studied, but that these occur on a much wider scale. Carriers were told that a shorter transport time might be used for other transports: the staff and vehicles/vessels can be released for other productive activities and that a higher reliability entails that the carrier can be more certain about such re-planning/re-scheduling. Shippers were asked to take into account what would happen (deterioration, disruption of production process, running out of stock, etc.) to the goods if the delivery were late. Nevertheless, respondents may still have difficulty in including other logistics structures in their valuations of time and reliability and not take a long run view. The TRs that we obtain should therefore be regarded as a lower boundary for the value of time in the longer run. The upper boundary will be around 1. How this should be used in CBA will be discussed in Chapter 9.

5.4.3 Comparison against the 2003/2004 VOTs

In 2003/2004 the VOTs per shipment per hour (in 2010 euros) are displayed in Table 24:

Table 24: Value of time (in 2010 euro/shipment/hour) and trade-off ratios (TR) for time versus cost from 2003/2004 survey

	Road	Rail^2	Air	Inland waterways	Sea
Factor cost per hour (in euro 1-1-02)	*[container]:* 50 *[non-container]:* 46 *[all]:* 47	*[full train]:* 1285	*[full freighter]:* 14132	*[ship]:* 201	*[ship]:* 628*
Factor cost per hour (in euro 1-1-10)	*[container]:* 58 *[non-container]:* 53 *[all]:* 55	*[full train]:* 1491	*[full freighter]:* 16390	*[ship]:* 233	*[ship]:* 728
Trade-off ratio time vs cost	*[container]:* 0.83 *[non-container]:* 0.80 *[all]:* 0.80	*[full train]:* 0.71	*[full freighter]:* 0.56	*[ship]:* 0.37	*[ship]:* 0.16
VOT (in euro 1-1-10 per transport per hour)	*[container]:* 49 *[non-container]:* 43 *[all]:* 44	*[full train]:* 1100	*[full freighter]:* 9200	*[ship]:* 86	*[ship]:* 120

Notes:
- VOTs after rounding off
- The rail values are for a train, not a wagon
- * Originally this was 457 euro, but later corrected to 628 euro

The VOTs in Table 24 are not exactly the values that are used at the moment in CBA. Table 24 gives the results of the previous VOT survey, after correcting for inflation. The current CBA value for road transport is 45.78 euro (in 2010 prices, see Rijkswaterstaat 2011), which is slightly different from the 44 for road transport in Table 24, because the CBA value took the previous VOT survey (RAND Europe et al., 2004) and corrected it for inflation and additional wage increases. The current CBA values for other modes than road transport are not based on SP survey results, but simply assume a trade-off ratio of 1 and then use the factor costs per hour.

Table 25: Value of time (in 2010 euro/shipment/hour) as currently used in CBA

	Road	Rail[2]	Air	Inland waterways	Sea
VOT (in euro 2010)	*[all]:* 45.78	*[full train]:* 1101.27	*[full freighter]:* 9519.14	*[ship]:* 88.77	*[ship]:* *87.57*

In these cases, the trade-off ratios can be directly compared between the 2003/2004 survey and the new 2010 survey. We then find that the new ratios for rail are somewhat lower and for air somewhat higher than those of 2003/2004, whereas those for inland waterways (waiting at locks) and sea are clearly higher than before (for sea transport and for locks/bridges in inland waterway transport, the full 95% confidence interval for the new estimate lies to the right of the old value). This may have to do with the fact that in the current project we carried out completely different SP experiments for these modes focussing on waiting for a lock or quay, where we expect that the VOT is more relevant for the inland waterway and sea transport carriers. The previous values for inland waterway and especially sea transport were regarded as probably too low, so in that respect the new values are an improvement.

For road transport the finding in 2003/2004 that the VOT for containers exceeds that for non-container transport also applies to the new survey. But keeping in mind that in the period 2002-2010 the general price level has gone up by almost 20%, the new road VOT for container is somewhat higher and the new one for non-container somewhat lower than before.

5.4.4 Comparison against the international literature

The new VOTs can also be compared to the international literature. There are two recent overview papers of freight VOTs: de Jong (2008) and Feo-Valero et al. (2011). The latter paper includes a table of road freight transports VOTs in 2005 Euro per tonne. We obtain a value for tonne for road of about 5 euro (similar to the 2003/2004 study). Feo-Valero et al. (2011) show a very large range, and our old and new values are well within that range. De Jong (2008) makes the distinction between the transport cost component of the VOT and the cargo-related component, and finds that many studies that seek for the sum of these find values between 30 and 50 euro (of 2002) per transport. This is consistent with our values for road.

Our rail transport value (about 1 euro per tonne) also falls well inside the range in Feo-Valero et al. (2011) and de Jong (2008). The most recent Norwegian study (Halse and Killi, 2012) found 1.7 euro per tonne per hour for rail (over all goods). For other modes, there is only very limited literature on the VOT.

5.5 Final VOR results

5.5.1 The new VORs

Table 26 shows the VORs from our selected models (after Jack-knife). The values '0' in this table do not denote cells that are zero by definition, but values that are not significantly different from zero. In the CBA, we recommend to use zero here. We observe that for the VOR there are clearly more cells with a non-significant value than for the VOT. We regard it as an indication that for various segments (especially for carriers) reliability, unlike transport time, is not such an important attribute. The costs of variability (such as early and late deliveries) are largely borne by the producers, not the transporters.

With the exception of a segment within road transport, we only find significant VORs for shippers that contract out. The shippers' VORs are higher for non-containers than for containers, possibly because the timely delivery of goods transported without containers is more important at the destination for further processing than for containerised goods. A 10% increase in reliability is worth as much as a 0.6-0.9% decrease in transport costs. These trade-off ratios refer to a relative change in the standard deviation of transport time and need to be multiplied by the factor cost to obtain the VOR per transport. Alternatively they can be expressed as a fraction of the VOT: 0.3 of the VOT for containers and 0.9 for non-containers (these ratios are similar to reliability ratios, but they refer to a relative change in reliability and time, not an absolute change as in a reliability ratio).

Table 26: Partial value of reliability (in 2010 euro/shipment/hour) and trade-off ratios (TR) for reliability versus cost

		Road	Rail	Air	Inland waterways	Sea
Container	**Carrier / own account shipper**	*[2-40t truck]:* 0	*[full train]:* TR=0	*Not applicable*	*[ship waiting for a quay]:* TR=0 *[ship waiting for a lock/bridge]:* TR=0	*[ship waiting for a quay]:* TR=0
	Shipper*	*[All]:* TR=0.06	*[All]:* TR=0.06	*Not applicable*	*[All]:* TR=0.06	*[All]:* TR=0.06
	Total	*[2-40t truck]:* TR=0.06	*[full train]:* TR=0.06	*Not applicable*	*[ship waiting for a quay]:* TR=0.06 *[ship waiting for a lock/bridge]:* TR=0.06	*[ship waiting for a quay]:* TR=0.06
Non-container	**Carrier / own account shipper**	*[2-15t truck]:* 29.62 *[15-40t truck]:* 0	*[full train]:* TR=0	*[full freighter aircraft]:* TR=0	*[ship waiting for a quay]:* TR=0 *[ship waiting for a lock/bridge]:* TR=0	*[ship waiting for a quay]:* TR=0
	Shipper*	*[All]:* TR=0.09	*[All]:* TR=0.09	*[All]:* TR=0.09	*[All]:* TR=0.09	*[All]:* TR=0.09
	Total	*[2-15t truck]:* 29.62 + 0.09* factor cost *[15-40t truck]:* 0.09*factor cost	*[full train]:* TR=0.09	*[full freighter aircraft]:* TR=0.09	*[ship waiting for a quay]:* TR=0.09 *[ship waiting for a lock/bridge]:* TR=0.09	*[ship waiting for a quay]:* TR=0.09
All		*[2-15t truck]:* 27.84 + 0.09* factor cost *[15-40t truck]:* 0.09*factor cost	*[full train]:* TR=0.08	*[full freighter aircraft]:* TR=0.09	*[ship waiting for a quay]:* TR=0.09 *[ship waiting for a lock/bridge]:* TR=0.09	*[ship waiting for a quay]:* TR=0.09

Notes:

- * Shipper = shipper that contracts out
- The values for shipper (both for container and non-container) do not differ between modes since the final models (Subsection 5.3.6) do not have terms depending on mode.

Table 27 gives the VORs, when combining the SP results of Table 26 with the factor costs in Table 22.

Table 27: Values of reliability for freight transport (Euro/hour per vehicle or vessel, price level 2010)

	Road	Rail	Air	Inland waterways	Sea
Container	*[2-40t truck]:* 4	*[full train]:* 100	*Not applicable*	*[ship waiting for a quay]:* 18 *[ship waiting for a lock/bridge]:* 27	*[ship waiting for a quay]:* 45
Non-container	*[2-15t truck]:* 34 *[15-40t truck]:* 6 *[all non-container]:* 15	*[bulk]:* 260 *[wagonload train]:* 240 *[all non-container]:* 250	*[full freighter aircraft]:* 1600	*[ship waiting for a quay]:* 25 *[ship waiting for a lock/bridge]:* 25	*[ship waiting for a quay]:* 110
All	*[2-40t truck]:* 14	*[full train]:* 200	*[full freighter aircraft]:* 1600	*[ship waiting for a quay]:* 24 *[ship waiting for a lock/bridge]:* 26	*[ship waiting for a quay]:* 100

Notes:

- All these values are combined values from shippers and carriers and were obtained after rounding off.
- The values for rail are for a train (not a wagon).
- The values for inland waterways and sea refer to a ship.
- The values for "all" types of transport are the average of the values for container and non-container weighted by the number of tonnes transported. The weight factors container / non-container are 7% - 93% for road, 35% - 65% for rail, 13% - 87% for inland waterways and 19% - 81% for sea transport.
- These values do not include VAT.

5.5.2 Comparison against previous values

Concerning the value of reliability, we cannot compare the outcomes of our study directly to the 2003/2004 results, but in De Jong et al. (2009) reliability ratios for freight transport on the basis of the standard deviation of transport time were calculated based on the 2003/2004 study. For road transport, a reliability ratio of 1.24 was found. The provisional value for CBA of 1.2 is based on this. When using the 2003/2004 VOTs in combination with this 1.24, we get a VOR per shipment per hour of 52 euro (of 2002) for container and 46 for non-container. Now we clearly find much lower money values (4-35 euro of 2010) for reliability of transport time in road transport (taking account of the contributions of both shippers and carriers. For the other modes, the previous study had a wide range of reliability ratios, and our new values match the low end of these. The previous values were only provisional, and were based on many assumptions that could not be tested. In the current survey, unreliability, its context and its consequences were made much more explicit and the presentation format is much more suitable for measuring unreliability in terms of the standard deviation of transport time (or scheduling terms). So whereas our values of time are comparable to the previous study, or higher for some modes (inland waterways, sea), the new values of reliability are lower than the earlier provisional values.

For shippers using road transport, Halse et al. (2010) obtain a reliability ratio of 1.2 (where we found 0.3 – 0.9). For carriers, they find the same as we do: the estimated coefficient for reliability is not significantly different from zero. Given that in Norway, the cargo component is about 10% of the VOT, setting the carrier VOR to 0 results in an overall (shippers and carriers) RR in road transport of 0.11. We find an overall RR for road transport of 0.37 (using Table 23 and Table 27), which exceeds the Norwegian finding. In Fowkes (2006) values of 0.19 and 0.34 for the overall RR are reported. We conclude that the overall RRs that we now find are substantially lower than the earlier value of 1.24, but that this value was based on many assumptions and that the few available empirical values in the literature are also much lower than 1.24.

The impact of just-in-time deliveries and perishable commodities on the VOR should be reflected in the shipper's component of the VOR. This component is significant in estimation, but usually not very large in money terms. One might have expected higher values for this component to reflect the popularity of just-in-time in modern logistics thinking, but the results that we obtain should also take into account that time-critical segments are still a relatively minor part of all freight transport (unless we would measure transport in terms of the value of the cargo shipped).[15]

The carrier component of the VOR has to do with the impact of reliability on being able to use vehicles and services for other transports. For this effect we find a coefficient that is not significantly different from zero (except for road non-container, 2-15 ton). This could be due to the small samples that we had to use in estimation and therefore we have to be careful in interpreting and using these results. In principle carriers could take into account that they could lose customers if their transport reliability became worse, but in our freight SP experiments, the changes in reliability are presented explicitly as things that happen to all carriers, so there are no competitive advantages or disadvantages here.

We expect that for carriers the sum of VOT and VOR will not exceed the total transport costs per hour. The reason for the benefits is that in the presence of time and reliability gains they can use their vehicles and staff for other transports. Their total benefits can never be higher than the transport costs per hour, otherwise they should have been carrying out those other transports instead of the current ones. It may happen that a carrier cannot reap the full benefits from a transport time gain, because there is uncertainty about the transport times.[16] As soon as this uncertainty is reduced (transport times become more certain, as do transport time gains), there could be benefits for the carrier, but according to our estimation results for most segments this is a small and not-significant effect. The

---

[15] The relatively small monetary values that we find for reliability seem to contradict surveys among shippers that found that reliability is the most important non-cost factor in mode choice (e.g. NERA et al., 1997). These studies however usually compare reliability to scheduled time, not to expected time (as we do), which will be more relevant if this often deviates from scheduled time (and then some of the value of unreliability will transfer to the value of expected transport time). More generally, a ranking study that finds reliability at the top of the list of non-costs attributes provides considerably less information than a stated preference study that gives a value of unreliability in money or transport time equivalents.

[16] A way this could be investigated would be to do case studies with qualitative in-depth interviews with firms involved in freight transport: how do they cope with the current uncertainties in transport time, and what might happen if there were shorter transport times, more reliable transport times or both?

VOR for road-container is lower than for non-container. This may have to do with the fact that the VOT for road-container is relatively high, leaving less room for a high VOR. The only segment for which we find a significant carrier VOR is road transport 2-15 ton. These vehicles are often used for urban distribution, where the uncertainty of travel times is large due to the heavy congestion in cities. Moreover, the planning patterns for these smaller freight vehicles are often quite complex, unlike those for larger road vehicles and other modes that are not so easily disturbed.

5.6 Validation of freight transport VOT/VOR outcomes against practice

We have assessed the new freight values of time and reliability in the eyes of the freight and logistics sector validity. This assessment is based on knowledge of logistics process and supporting calculations in firms, but not on any new surveys among firms.[17]

The main conclusion of the assessment is that the factor cost part of time losses (the carrier part of the VOT) will be recognised by those working in the field. The shipper part of the VOT is difficult to judge because there are hardly any values on this used in practice. Since on average the factor cost will be dominant, we can expect that the full VOTs presented in this report will be at levels where many firms would expect them to be. For some time-critical market segments, the shipper part can be relatively large, and the VOTs in this report may be seen as an underestimation of the true values for those segments. The new VORs are even more difficult to assess, in practice the value of reliability depends very much on the market segment in question and is difficult to separate from the value of time.

---

[17] This section is based on Ploos van Amstel & Tavasszy (2011)

PASSENGER TRANSPORT

CHAPTER 6 Passenger transport questionnaire

6.1 Set-up of the questionnaire

The passenger transport questionnaire consisted of the following parts:

1. questions regarding the attributes of the trip recently carried out[18], e.g. travel time and costs. These values are used as the base levels for the attributes presented in the SP experiments;
2. questions regarding the availability of another mode for this trip and what the attribute levels would be for that mode (to allow an RP model to be estimated).
3. SP experiment 1 (travel time versus travel cost)
4. Introduction of variable travel times and SP experiment 2a (usual travel time, variation in travel times, most likely arrival time, and travel costs). Note that respondents in the recreational navigation segment did not participate in this experiment.
5. SP experiment 2b (same as 2a without the variation in most likely arrival time).
6. questions in which respondents were asked to evaluate the choices they made in the experiments.
7. questions about the person (age, gender, etc.) and household (composition, income, etc.)

The full questionnaires are available on request.

---

[18] For the 2009 survey, respondents were asked to think back of the most recent trip they made for a certain (preselected) purpose. For the 2011 survey, respondents were asked to think back of the trip they made when they were recruited for this survey.

6.2 SP experiments for passenger transport (excluding recreational navigation)

6.2.1 Description of the experiments

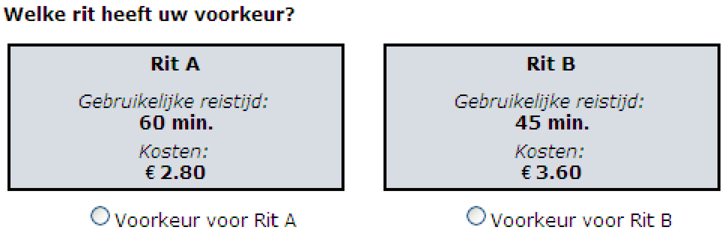
The choice situations in all SP experiments are identified by within-mode choices. Given a certain mode, each choice set consists of two generic alternatives and the respondent was asked to choose the most preferred one.

For passenger private car SP, the choice alternatives are identified as two available routes for a respondent's particular car journey, while for passenger public transport SP (including air), the two alternatives were introduced as two possible services that differ in terms of cost, (timetable) travel time, and unreliability. The recreational navigation SP was slightly different, since it involved a waiting situation before a bridge or lock (see Section 6.3).

The presentation formats and the number of choice situations per experiment were the same as for freight transport (see Chapter 3). Experiment 1 and 2a consisted of six choice pairs and experiment 2b consisted of seven choice pairs of which the last-but-one choice pair was a dominant choice. This choice is discussed in more detail in Section 6.4.

The statistical design and the choice of presentation format for reliability are described in Significance et al. (2007) and Tseng et al. (2009). The most important tables are repeated in Appendix A. An overview of the attributes for all respondents except recreational navigation is similar to Table 2, whereas an overview of the attributes for recreational navigation respondents is similar to Table 3. Note that the order of the SP questions within each experiment was randomized.

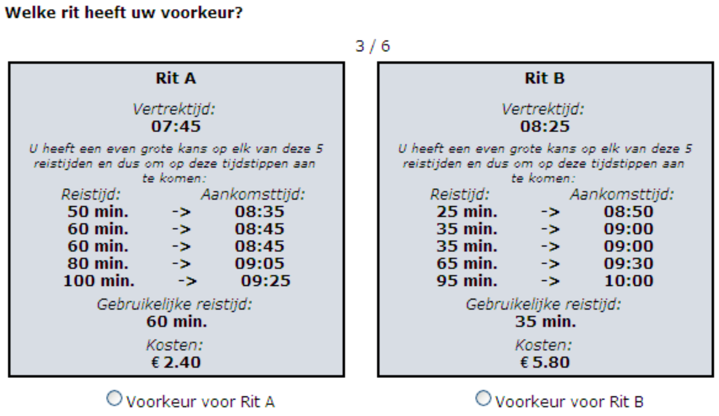
Figure 6 to Figure 8 show actual screen-shots of the web-based SP questions for a car respondent. As can be seen, they are very similar to the freight SP questions (see Figure 2 and Figure 3). Note also that the most likely arrival times for both alternatives in experiment 2b (08:45, see Figure 8) are the same, whereas they are different in experiment 2a (08:45 versus 09:00, see Figure 7).



Welke rit heeft uw voorkeur?

Rit A	**Rit B**
Gebruikelijke reistijd: **60 min.**	*Gebruikelijke reistijd:* **45 min.**
Kosten: **€ 2.80**	*Kosten:* **€ 3.60**
○ Voorkeur voor Rit A	○ Voorkeur voor Rit B

Figure 6: Example of SP question of experiment 1 for car respondents



Welke rit heeft uw voorkeur?

3 / 6

Rit A

Vertrektijd:
07:45

U heeft een even grote kans op elk van deze 5 reistijden en dus om op deze tijdstippen aan te komen:

Reistijd:		*Aankomsttijd:*
50 min.	->	**08:35**
60 min.	->	**08:45**
60 min.	->	**08:45**
80 min.	->	**09:05**
100 min.	->	**09:25**

Gebruikelijke reistijd:
60 min.

Kosten:
€ 2.40

○ Voorkeur voor Rit A

Rit B

Vertrektijd:
08:25

U heeft een even grote kans op elk van deze 5 reistijden en dus om op deze tijdstippen aan te komen:

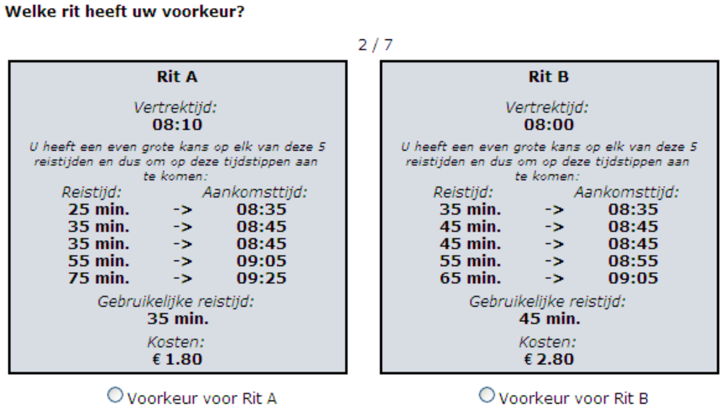
Reistijd:		*Aankomsttijd:*
25 min.	->	**08:50**
35 min.	->	**09:00**
35 min.	->	**09:00**
65 min.	->	**09:30**
95 min.	->	**10:00**

Gebruikelijke reistijd:
35 min.

Kosten:
€ 5.80

○ Voorkeur voor Rit B

Figure 7: Example of SP question of experiment 2a for car respondents



Welke rit heeft uw voorkeur?

2 / 7

Rit A

Vertrektijd:
08:10

U heeft een even grote kans op elk van deze 5 reistijden en dus om op deze tijdstippen aan te komen:

Reistijd:		*Aankomsttijd:*
25 min.	->	**08:35**
35 min.	->	**08:45**
35 min.	->	**08:45**
55 min.	->	**09:05**
75 min.	->	**09:25**

Gebruikelijke reistijd:
35 min.

Kosten:
€ 1.80

○ Voorkeur voor Rit A

Rit B

Vertrektijd:
08:00

U heeft een even grote kans op elk van deze 5 reistijden en dus om op deze tijdstippen aan te komen:

Reistijd:		*Aankomsttijd:*
35 min.	->	**08:35**
45 min.	->	**08:45**
45 min.	->	**08:45**
55 min.	->	**08:55**
65 min.	->	**09:05**

Gebruikelijke reistijd:
45 min.

Kosten:
€ 2.80

○ Voorkeur voor Rit B

Figure 8: Example of SP question of experiment 2b for car respondents

6.2.2 Description of the attributes in private car transport

Usual travel time

Travel time refers to the amount of door-to-door journey time for a one-way trip. It is based on the expected travel time at the moment of departure of the recent trip described by the respondent.

Travel cost

Travel cost refers to the total cost that a respondent has to pay for his one-way car journey. For the running cost, it may not be easy for a car driver to answer the exact running (or fuel) cost of a particular trip: therefore, we choose to collect the information on trip distance (in kilometres), the fuel type and perceived fuel efficiency, and use it as a reference to estimate the cost approximately. Hence, the base value of a respondent's cost attribute is based on kilometres travelled for the trip under consideration. Note that this calculated cost was always presented to the respondent for verification. If the respondent did not agree with this cost, he was able to change it. However, a large majority of respondents accepted our cost estimate.

Reliability

This is the reliability of travel time, which is presented by a series of five possible travel times. Since this is the most important element in the SP experiment, in the survey design phase the research team carried out in-depth face-to-face interviews to determine the best concept and format for presenting reliability to respondents. The verbal presentation of five possible travel times turned out to work best in many respects (Significance et al., 2007; Tseng et al., 2009). This format was used in the main survey.

Departure / arrival time

For a scheduled trip, the timing (departure and arrival times) of the trip is an important decision factor for the traveller. Different departure times and travel times of course result in different arrival times, and this is presented consistently in experiments 2a and 2b. The instruction makes clear that the respondent is supposed to indicate the *most preferred* among the alternatives shown, which may mean that in reality (s)he would further fine tune departure time. We show mutually consistent combinations of departure times, travel times, and arrival times. It is the most likely arrival time that is an attribute to the design of the experiment, the departure times being implied.

We help the respondent by showing explicitly the usual travel time (=the most common travel time = the median travel time = the second and third travel time from the series of five; the mean travel time will be greater than this value). To limit the amount of information we decided not to show the average travel time for each alternative.

6.2.3 Description of the attributes in public transport

Fare

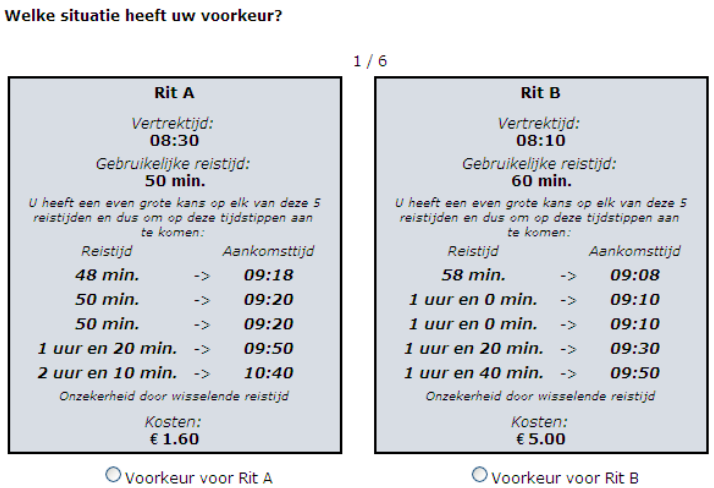
This is the full price that a respondent has to pay for his one-way public transport journey. The fare level should ideally be based on a respondent's actual payment for his one-way public transport journey. However, for those cases in which the respondent does not know his actual fare, it is difficult to compute the fares for different types of tickets used by the respondents. For example, some people use a yearly public transport card (paid by employers), or a free student card, and there are all sorts of discounted tickets, etc. We resolved this problem by relating the fare to the amount of travel time (or trip length) and public transport mode. Based on these inputs, we could derive a plausible price level that is close to official fare set by public transport companies, and ask respondents to imagine that this is the price to be considered. Note that the respondents again had the option to change the fare level that was calculated by us, if they thought this was not realistic.

Timetable travel time

An important feature for public transport is adherence to the timetable. This leads to a rather specific type of distribution of travel times. Therefore, it is important to emphasize that the time attribute shown in the SP refers to the publicly advertised *timetable* travel time.

Reliability

The presentation of travel time reliability is identical to that for car, though the shape of the time distribution is different. Since the travel times are timetable times, it is highly unlikely that the travel time will be much less than published in the timetable. Therefore, the fastest of the five possible travel times does not differ much from the most likely (second and third) travel time. As can be seen from Figure 9, the fastest trip is only 2 minutes quicker than the usual trip, whereas for car travel (Figure 7) the quickest trip can be up to 10 minutes quicker.



Welke situatie heeft uw voorkeur?

1 / 6

Rit A

Vertrektijd:
08:30
Gebruikelijke reistijd:
50 min.
U heeft een even grote kans op elk van deze 5 reistijden en dus om op deze tijdstippen aan te komen:

Reistijd		*Aankomsttijd*
48 min.	->	***09:18***
50 min.	->	***09:20***
50 min.	->	***09:20***
1 uur en 20 min.	->	***09:50***
2 uur en 10 min.	->	***10:40***

Onzekerheid door wisselende reistijd
Kosten:
€ 1.60

○ Voorkeur voor Rit A

Rit B

Vertrektijd:
08:10
Gebruikelijke reistijd:
60 min.
U heeft een even grote kans op elk van deze 5 reistijden en dus om op deze tijdstippen aan te komen:

Reistijd		*Aankomsttijd*
58 min.	->	***09:08***
1 uur en 0 min.	->	***09:10***
1 uur en 0 min.	->	***09:10***
1 uur en 20 min.	->	***09:30***
1 uur en 40 min.	->	***09:50***

Onzekerheid door wisselende reistijd
Kosten:
€ 5.00

○ Voorkeur voor Rit B

Figure 9: Example of SP question of experiment 2a for public transport respondents

In the case of public transport, two different types of reliability were considered: unreliable waiting time at the platform and unreliable in-vehicle travel time. To distinguish these two types of values, without burdening the respondents with too complicated choices:

- 1/2 of the PT respondents were told that the variability of travel time was due to variable waiting times;
- 1/2 of the PT respondents were told that the variability of travel time was due to variable in-vehicle travel times.

As regards the issue of 'transfers', we proposed to allow for trips that have transfers, but to keep the number of transfers fixed for a given respondent; i.e., there was no attempt to provide a 'value of transfers' (or rather the disutility associated with having to interchange). However, we are able to determine whatever effect the number of transfers has on the valuation of travel time and reliability (over the full trip) in public transport.

Departure / arrival time

These are presented in the same way as for car transport.

6.2.4 Description of the attributes in air transport

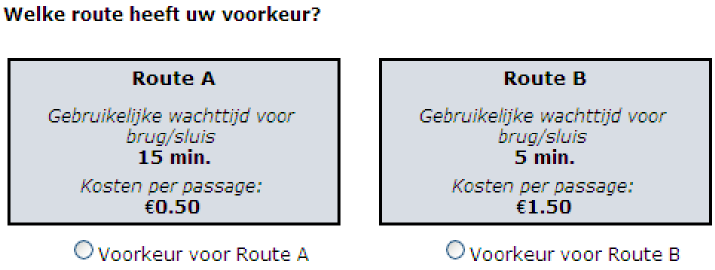
These attributes are very similar to the attributes used in public transport. However, the travel time and travel costs only refer to the air-part of the trip, i.e. from departure airport to the final destination airport. The access and egress trips to and from the airports are not included, either in the travel time attribute, or in the travel cost attribute.

6.3 SP experiments for recreational navigation

6.3.1 Description of the SP experiments

Since the purpose of a recreational navigation trip is usually not so much to travel from A to B, but to have an enjoyable trip, we expect the travel time to be valued positively instead of negatively (the longer the trip, the more preferred). However, in the Netherlands, the VOT for recreational navigation is not so much used to evaluate possible new canals or other boat routes, but to evaluate new (or lower) bridges and locks.

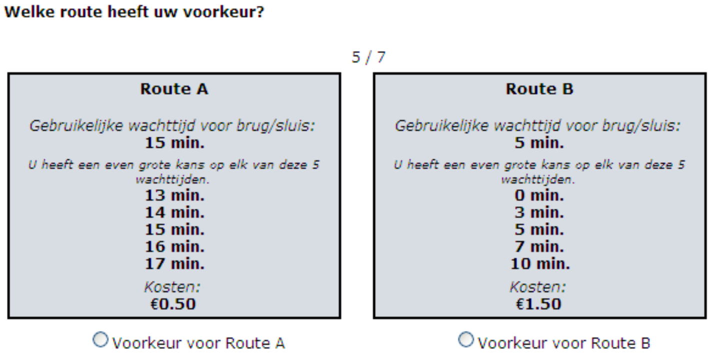
Therefore, we have used a different setting for the recreational navigation experiments: respondents are asked to think of a situation where they have to wait for a bridge or a lock. Since we do not believe they have to depart or arrive at a certain time, we do not include departure and arrival times in the experiment. Hence, experiment 2a and 2b become identical and only experiment 2b is presented to them. Screen shots from actual choice pairs are presented in Figure 10 (experiment 1) and Figure 11 (experiment 2b). The complete design can be found in Appendix A.



Welke route heeft uw voorkeur?

Route A	Route B
Gebruikelijke wachttijd voor brug/sluis	*Gebruikelijke wachttijd voor brug/sluis*
15 min.	**5 min.**
Kosten per passage:	*Kosten per passage:*
€0.50	**€1.50**
○ Voorkeur voor Route A	○ Voorkeur voor Route B

Figure 10: Example of SP question of experiment 1 for recreational navigation respondents



Welke route heeft uw voorkeur?

5 / 7

Route A	Route B
Gebruikelijke wachttijd voor brug/sluis:	*Gebruikelijke wachttijd voor brug/sluis:*
15 min.	**5 min.**
U heeft een even grote kans op elk van deze 5 wachttijden.	*U heeft een even grote kans op elk van deze 5 wachttijden.*
13 min. **14 min.** **15 min.** **16 min.** **17 min.**	**0 min.** **3 min.** **5 min.** **7 min.** **10 min.**
Kosten:	*Kosten:*
€0.50	**€1.50**
○ Voorkeur voor Route A	○ Voorkeur voor Route B

Figure 11: Example of SP question of experiment 2a for recreational navigation respondents

6.3.2 Description of the attributes in recreational navigation

Usual waiting time for a bridge/lock

This is the time that one usually has to wait before a bridge or lock can be passed. In experiment 2b it is equal to the average waiting time.

Cost

This is the cost that has to pay to pass the bridge or lock.

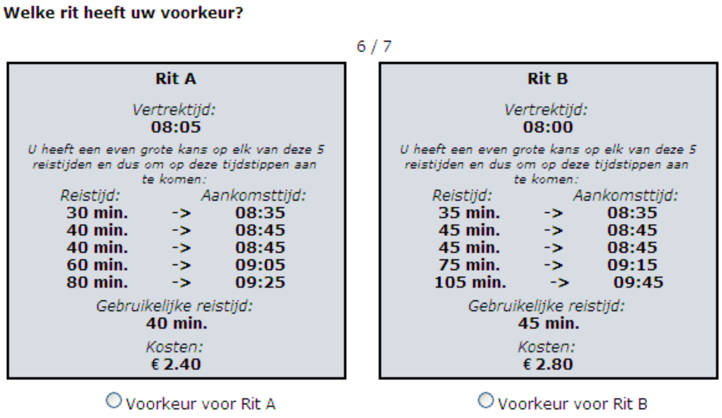
Reliability

This is presented as five possible and equally likely waiting times.

6.4 Dominant questions

As has been mentioned before, the last-but-one question in experiment 2b for all respondents (both in freight and passenger transport, also including respondents in inland waterways/sea and respondents in recreational navigation) was a so-called dominant question. This is a question for which the attributes of one alternative are all better or equal compared to the attributes of the other alternative.

An example for respondents in the car segment is shown in Figure 12. As can be seen from this example, the trip costs for Trip A (left) are lower than for Trip B (right), the usual travel time is shorter and the travel time is more reliable (less variation in travel time). The most likely arrival time is the same for both alternatives (i.e. 08:45). Since the most likely travel time for Trip A is less, the departure time is later compared to Trip B (08:05 versus 08:00).



Welke rit heeft uw voorkeur?

6 / 7

Rit A

Vertrektijd:
08:05

U heeft een even grote kans op elk van deze 5 reistijden en dus om op deze tijdstippen aan te komen:

Reistijd:		*Aankomsttijd:*
30 min.	->	**08:35**
40 min.	->	**08:45**
40 min.	->	**08:45**
60 min.	->	**09:05**
80 min.	->	**09:25**

Gebruikelijke reistijd:
40 min.

Kosten:
€ 2.40

○ Voorkeur voor Rit A

Rit B

Vertrektijd:
08:00

U heeft een even grote kans op elk van deze 5 reistijden en dus om op deze tijdstippen aan te komen:

Reistijd:		*Aankomsttijd:*
35 min.	->	**08:35**
45 min.	->	**08:45**
45 min.	->	**08:45**
75 min.	->	**09:15**
105 min.	->	**09:45**

Gebruikelijke reistijd:
45 min.

Kosten:
€ 2.80

○ Voorkeur voor Rit B

Figure 12: Example of a dominant SP question

CHAPTER 7 Passenger transport data

For passenger transport, two data sets were collected. The original plan was to use only data obtained from internet interviews with members of the PanelClix panel, but initial data analysis on this sample alone yielded questionable results, that are probably to a considerable degree due to biased response (see Section 8.5.2). Therefore, the decision was taken to gather an additional data set, using recruitment of respondents in line with previous research, i.e. at petrol stations/service areas along motorways, parking garages, train stations, bus stops, airports, recreational harbours and locks. These were then followed with interviews by means of the internet. Below we describe both data sets in turn.

7.1 The 2009 data (Panelclix panel)

7.1.1 Recruitment

Initially, all respondents for passenger transport were drawn from the internet panel of PanelClix. This is the largest on-line panel of The Netherlands (240.000 participants), and with an expected response of around 40% large enough to cover all segments.

The aim for the main survey was to obtain minimally 5,200 successfully completed interviews in passenger transport, with a distribution over segments or sampling strata as given in Table 28. Within each segment, the sample was drawn to be representative for the Dutch population in terms of income, job participation, age, gender and region (different access to internet is taken into account here).

The distribution over the segments does not have to be representative for the Dutch population or the distances travelled in The Netherlands. The target number of interviews per segment is based on the number that is minimally required for good and reliable model estimation and for the different VOTs and VORs that this research should deliver at the end. As a result of this, while outcomes per stratum can be used directly, an expansion/re-weighting is required for outcomes over several or all segments (e.g. on the basis of the Dutch national travel survey, MON/OViN).

Table 28: Targets for the numbers of successfully completed passenger interviews

Experimental segments	Target segments		Targets
Private car transport	Trip purpose and time of day	Commuting Business Others Peak, off-peak	2500 for car drivers[19] in total. with 1250 for commuting 500 for business 750 for others 50% peak / 50% off peak
Public transport	Mode, trip purpose and time of day	Train/metro, bus/tram, high speed rail, air Commuting Business Others Peak, off-peak	2500 for passenger public transport in total. 1250 for commuting 500 for business 750 for others Among these trip purposes, 40% is for train/metro/lightrail/"sneltram", 40% is for bus/tram, 20% is for high-speed rail/air. 50% peak / 50% off-peak 50% with transfers within public transport (e.g. bus-train, train-train)/50% without
Recreational navigation	-	-	200

PanelClix invited members of its own on-line panel (in various survey waves) to participate in the survey, which could be started by clicking on a weblink. The members received a reward for successfully completing the interview (100 Clix-points, which is equivalent to €1.50).

All respondents were asked which modes they had used in the past three months, for which travel purpose, whether peak or off-peak, and how often they had made an interchange (for public transport). This was used to allocate respondents to questionnaires for specific segments, in particular where they had not yet reached their interview target.

We used different questionnaires for car driver, public transport, air transport and recreational navigation. A pilot survey amongst 284 respondents was carried out early October 2009 to test whether the web questionnaire was functioning correctly. After some final changes had been incorporated, the main survey using the PanelClix panel was carried out in the last three weeks of November 2009. The data from the pilot survey was not included in the final analysis.

---

[19] In the 1997-1998 passenger VOT survey there were no interviews among car passengers (meaning here non-drivers). Instead for CBA, the recommendation was to use 0.8 times the car driver VOT for car passengers. For the current VOT/VOR survey we decided to stick to this approach.

7.1.2 Number of respondents

Between 11 and 27 November 2009 44,320 members were invited. 17,845 persons made an initial response (40.3%) and started with the introductory questions. 7,052 (39.5% of 17,845) respondents were forwarded to the main part of the questionnaire. The others had made trips for segments that were already fully covered. 5,788 respondents fully completed the questionnaire (82% of 7,052). Of these, 28 persons were discarded because essential information from the opening questions was missing. This leaves 5,760 useable questionnaires, which were distributed between modes as follows.

Table 29: Successfully completed interviews from the PanelClix panel, by mode

	Number of respondents at start of questionnaire	Number of respondents at end of questionnaire	Effective number of respondents
Car	3143	2661	2654
Public transport	3075	2379	2361
Airplane	601	534	531
Recreational navigation	233	214	214
Total	**7052**	**5788**	**5760**

On average, a respondent who completed the survey needed about 15 minutes to complete the questionnaire (also see Table 30).

Table 30: Time to complete the interview (in minutes)

	Car	Public transport	Airplane	Recreational navigation
Minimum (min.)	3.1	4.1	4.5	2.7
Maximum (min.)	171.3	159.3	107.5	127.2
Mean (min.)	15.9	18.5	16.1	12.8
Median (min.)	13.9	16.0	13.4	10.6

In total 2,654 respondents answered the car questionnaire, which exceeds the target of 2,500. Also the targets for the three purposes and peak/off-peak were met.

2,361 persons completed the public transport questionnaire. Additionally, there are 531 respondents that completed the air questionnaire. This adds up to 2,892 interviews, clearly above the target of 2,500. The targets for travel purposes, peak/off-peak, trips with and without transfers and the subdivision over modes have all been met.

214 persons responded to the recreational navigation questionnaire, which also exceeds the target number of 200.

7.2 The 2011 data (en-route recruitment)

In December 2010 it was decided that additional respondents for the passenger survey should be recruited using a recruitment method that is similar to the method used for the 1997 survey. This decision was taken by the Ministry of Infrastructure and the Environment, following the recommendation of the research team, when the initial VOT results from models estimated on the 2009 survey were found to be implausibly low (see Section 8.5.2), after which various external experts were consulted to advice on further steps (Prof. dr. Peter Bonsall of ITS Leeds, Prof. dr. Jonas Eliasson of KTH Stockholm and Dr. Eric Molin of Delft University of Technology with Prof. dr. Harry Timmermans of the Technical University of Eindhoven).

Both the external experts and the research team emphasised the possibility that the sample of respondents obtained from PanelClix were biased with respect to their value of time. Within each segment (socio-economic, trip purpose, trip length, mode), the respondents that participate in such an online panel (which takes time, for a rather low monetary reward) might have a lower VOT than non-participants. Even after expansion, the resulting VOT would then be lower than the true VOT.

7.2.1 Recruitment

The recruitment was done as follows.

- A recruiter was stationed at a location where a lot of travellers can be contacted, e.g. a petrol station/service area, parking garage, a bus stop or a train station. See Table 31 for a complete list of recruitment locations.
- He explains the traveller briefly about our survey and asks whether he/she is willing to participate. A few travel cheques of € 500 were awarded to (randomly selected) participants of the survey.
- If a person agrees to participate, the recruiter hands him/her a small card with a web address where the survey can be found and a login-code for this website. Furthermore, the recruiter writes down the e-mail address of the participant.
- Within a few days, we send the participant an e-mail reminder of the survey. This e-mail again informs the reader about the research, the reward and the way to participate (web address).
- If the participant has not responded, a second e-mail reminder is sent after a week.

In total, 3,650 respondents were recruited over a period of about 5 months in 2011. Of these 1,757 persons (48%) visited the website, 1,734 persons used a login-code and 23 persons accessed the website without a login-code, which was also possible. In 1,671 cases, we were able to link a web-site visitor to a person that was recruited at a certain location. For the remaining cases, this was not possible because of one of the following reasons:

- the person had accessed the website without a login-code;
- the person had received a randomly distributed card: on a few days a limited number of cards were given out to travellers without writing down their personal details;

- the person made a mistake when typing in his login-code;
- the recruiter had made a mistake when noting the login-code that he had given to the respondent.

Table 31: Recruitment locations for the 2011 survey

Mode	Location
Car	A1 petrol station / service area De Hackelaar / Honswijk (Muiden)
	A2 petrol station / service area Ooijendonk (Liempde)
	A2 petrol station / service area Velder (Liempde)
	A4 petrol station / service area Den Ruygenhoek Oost (Nieuw-Vennep)
	A4 petrol station / service area Den Ruygenhoek West (Nieuw-Vennep)
	Amsterdam parking garage P1 Centrum
	Amsterdam parking garage Muziektheater
	Arnhem parking garage Musisgarage
	The Hague parking garage Centraal Station
	The Hague parking garage Plein
	The Hague parking garage Rijnstraat / VROM
	Rotterdam parking garage Groothandelsgebouw
Public transport	Train station Den Bosch - Stationsplein
	Train station The Hague Central Station - Stationsplein
	Train station Rotterdam Central Station - Stationsplein
	Bus/tram stop Amsterdam - Muntplein
	Bus/tram stop Amersfoort NS Station - Streekbusstation
	Bus/tram stop Arnhem NS Station - Breng
	Tram stop The Hague - Oostinje
	Bus/tram stop The Hague - Laan van NOI
	Bus/tram stop Utrecht Central Station - Streekbusstation
	Bus/tram stop Eindhoven NS Station - Streekbusstation
	Bus/tram stop Leiden Central Station - Streekbusstation
Air	Schiphol airport
	Eindhoven airport
Recreational Navigation	Naarden - harbour
	Muiden - lock
	Bruinisse - harbour
	Bruinisse - lock

Nonetheless, the answers of these respondents can still be used in the survey, since all relevant information is asked in the questionnaire and knowing the recruitment location and date is not essential for the analysis.

Of the 1,757 persons that accessed the website, 1,431 completed the questionnaire (i.e. completed the first series of stated preference questions). Most respondents who aborted the questionnaire stopped already after the first (or first few) question(s).

The response rates (i.e. visits to our website) per recruitment day and per location varied from 15% to 73% (the actual response rates were slightly higher, because some respondents could not be linked to their recruitment location). We believe this reflects differences between recruiters: some will have been more accurate and more able to motivate people to really participate in our survey.

7.2.2 Number of respondents

The survey was closed on the 24th of August 2011. At that moment, 1430 respondents had completed the survey, which was 95% of the targeted number of 1501 interviews (see Table 32).

Table 32: Number of completed surveys

	Commute	Business	Other	Total*
Car	184 *(400)*[#]	306 *(111)*	125 *(175)*	615 *(686)*
Public transport	256 *(385)*	69 *(88)*	194 *(128)*	519 *(611)*
Train[^]	131	52	103	286
BTM[^]	125	17	91	233
Air	9	29 *(41)*	163 *(102)*	201 *(143)*
Recr. navigation	0	0	95 *(61)*	95 *(61)*
Total	**449** *(785)*	**404** *(250)*	**577** *(466)*	**1430** *(1501)*

* Note that we included each respondent that finished the first SP experiment as "Completed". A few more respondents dropped out at a later point in the survey, but their answers can still be used.

[#] The numbers between brackets indicate the targeted number of respondents.

[^] The number of respondents in public transport is also split by mode: train and BTM (= bus, tram and metro).

We conclude that:

1. the total number of respondents is somewhat lower than planned, especially for car and public transport;
2. the distribution over the purposes is different than planned.

These two issues will be discussed in more detail below.

Ad 1. Lower number of respondents

We had anticipated that about 2/3 of all travellers that agreed to participate and that had left us their e-mail address would actually complete the survey. In reality, this percentage turned out to be lower (about 40%). Many e-mails bounced due to errors in the e-mail address. Some problems with recruiters recruiting the wrong type of travellers were also reported. But the biggest problem was that many recruited persons did not even open the web-page that was on their card.

We therefore recruited 62% more respondents than initially planned (at our own cost). Table 33 shows the number of recruited persons and the response rates:

Table 33: More detail on the number of completed surveys

	Planned	Realised		Started survey	Completed*
Car	1029	1509	+47%	745	615
Train	291	361	+24%	618	286[#]
BTM	626	1047	+67%		233[#]
Air	215	529	+146%	262	201
Recr. navigation	92	204	+122%	131	95
Total	**2253**	**3650**	**+62%**	**1756**	**1430**

Notes:

* These numbers differ slightly from the numbers in Table 32, since some respondents completed for instance a public transport questionnaire though they were recruited at a car location.

[#] Note that respondents were categorized as train or BTM based on the public transport mode that they used for the greatest distance and not based on the location where they were recruited. This explains why there are about the same number of completed interviews for train and BTM whereas many more respondents were recruited at BTM stops.

Based on a partial analysis of the data, we concluded that the quality of the data was higher than expected. This means that even though the final number of respondents was a little less than planned, we were able to estimate excellent models. It was therefore decided that recruitment could be stopped and the survey could be closed.

Ad 2. Distribution over the purposes.

The distribution of the target number of respondents over the purposes was originally made on the basis of the purpose distribution in the 2009 survey. In that survey each respondent was asked to think about a recent trip. In the 2011 survey, each respondent is asked to think about the trip that they were making when they were recruited. If the probability of recruitment is independent of trip purpose (which it was: the recruiter did not ask about the trip purpose before he asked whether a person was willing to participate), a different distribution over the purposes is to be expected. In fact, the purpose distribution in the 2011 survey is more like the purpose distribution in the 1997 and 1988 survey which used a similar recruiting procedure. We conclude that it would have been better to use the results from the 1997 and 1988 surveys for the distribution of the target number of respondents over the purposes.

However, since we will estimate separate models by purpose, we do not expect any major problems as a result of the slightly different distribution over the purposes.

7.2.3 Additional questions

In the interviews with the respondents recruited en-route, we used basically the same questionnaire as for the survey in 2009. Some small changes were necessary, since there was no need any more to select a recent trip. We added a question on whether the respondent would have made a different choice in the SP experiments in November 2009 (at the time of the PanelClix survey). There might be a difference between the periods because of changes in the economic situation and outlook (persistence of the economic slowdown/recession) or between concerns about the imminent introduction of road pricing (no longer an issue in 2011), which could affect the VOT and/or VOR. We also added a question as to whether the respondent is a member of an internet panel. We present the results of these questions in the tables below.

Table 34: Answers to additional questions on choice behaviour

A: If you had completed this questionnaire in November 2009, would you then have made different choices in the repeated comparisons of two alternatives?

	Yes	No	Number of respondents who answered this question	Total number of respondents
Car	12.4%	87.6%	532	615
Public transport	17.1%	82.9%	461	519
Air	5.1%	94.9%	175	201
Recr. navigation	10.9%	89.1%	92	95
Total	**13.0%**	**87.0%**	**1260**	**1430**

B: Are you a member of an internet panel that sometimes asks you to complete questionnaires?

	Yes	No	Number of respondents who answered this question	Total number of respondents
Car	19.1%	80.9%	523	615
Public transport	23.9%	76.1%	452	519
Air	16.4%	83.6%	171	201
Recr. navigation	13.0%	87.0%	92	95
Total	**20.0%**	**80.0%**	**1238**	**1430**

C: [If Yes to the previous question] Are you a member of the PanelClix panel?

	Yes, just the PanelClix panel	Yes and also another panel	No, another panel (s)	Number of respondents who answered this question	Total number of respondents
Car	4.0%	9.0%	87.0%	100	615
Public transport	4.6%	5.6%	89.8%	108	519
Air	0.0%	3.6%	96.4%	28	201
Recr. navigation	8.3%	8.3%	83.3%	12	95
Total	**4.0%**	**6.9%**	**89.1%**	**248**	**1430**

In Table 34 we find that only 13% of the respondents would have answered differently in 2009, which is a low fraction. The 20% internet panel members (Table 34B) will be investigated separately from the other 80% (see Section 8.5.2). Preferably we would also have done that for the 4% members of the PanelClix members in the 2011 data (Table 34C), but this share is just too small for separate analysis.

7.3 Data selection and data quality

7.3.1 Treatment of the dominant alternative

Table 35 shows the distribution of the 5760 respondents of the 2009 survey and the 1429[20] respondents of the 2011 survey over the mode/purpose segments.

Table 35: Data base used for estimation: number of respondents per mode-purpose segment

			Purpose			
Year			Commute	Business	Other	Total
2009	Mode	Car	1341	523	790	2654
		Train	908	284	329	1521
		BTM	586	80	174	840
		Plane	0	157	374	531
		Recr.Nav.	0	0	214	214
	Total		**2835**	**1044**	**1881**	**5760**
2011	Mode	Car	184	305	125	614
		Train	131	52	103	286
		BTM	125	17	91	233
		Plane	9	29	163	201
		Recr.Nav.	0	0	95	95
	Total		**449**	**403**	**577**	**1429**
Total	Mode	Car	1523	828	915	3266
		Train	1038	336	432	1806
		BTM	711	97	265	1073
		Plane	9	186	537	732
		Recr.Nav.	0	0	309	309
	Total		**3281**	**1447**	**2458**	**7186**

Of the 2009 respondents, 8.3% made a non-intuitive choice in the dominant choice pair (see Section 6.4). This choice could either be a valid choice, or it could be an indication that these respondents did not understand the choice experiment, or were not paying attention to the attribute levels of the alternatives. In the latter situation, including these respondents in the final analysis would lead to additional noise and a worse estimate of the model coefficients. Indeed, a test revealed that excluding these respondents led to an improved t-ratio of the coefficient, or the t-ratio did not deteriorate as much as was to be expected by excluding this number of respondents.

Lancsar and Louviere (2006) have argued that analysts need to be very careful with the exclusion of respondents that gave unexpected answers. However, due to the careful design

---

[20] Due to unclear answers, one respondent in the 2011 could not be assigned to a mode/purpose segment. This respondent was excluded from further analysis.

of the dominant question, their arguments do not apply to our situation. All attributes in which the two alternatives are different are quantitative attributes, with a clear expectation of what is to be valued more. The only attribute of which we do not have a clear expectation as to what a respondent might like better (i.e. the most likely arrival time) was kept constant between the two alternatives. So, a respondent making a non-intuitive choice chose the slower, more expensive and less reliable trip over a faster, cheaper, more reliable trip. Hence, it is very likely that these respondents did not understand the experiment or did not pay attention to the attribute values.

After the final models were estimated, we compared the final results with an estimation that included these respondents. From these estimations, we concluded that data from these respondents contained indeed more noise than average. By excluding them, we did not introduce any bias, we just removed noise (see Section 8.5.6).

7.3.2 Outliers

In this section we describe the further checks that we made on the passenger transport data. We also give the number of respondents that were excluded from further analysis and the criteria used for exclusion. Some of these criteria refer to all four passenger questionnaires, others are specific for one or a few of the questionnaire types.

We start with the numbers per questionnaire from Sections 6.1 (2009) and 6.2 (2011). These are in the top row of the Table 36 and Table 37 respectively. Then follow all the exclusions and at the bottom is the number of interviews that remains for analysis.

Table 36: Number of interviews per questionnaire before and after checks 2009 data (PanelClix internet panel)

2009 data	Car	Public transport	Air	Recreational navigation
Data collected	**2654**	**2361**	**531**	**214**
Error in base cost or time calculation	2	1	-	-
Origin or destination municipality cannot be identified	102	63	n.a.	n.a.
Use of other modes than just car	93	n.a.	n.a.	n.a.
Travel time without delay exceeds real travel time	26	18	n.a.	n.a.
Travel time without delay exceeds expected travel time	48	60	n.a.	n.a.
Big difference between real and expected travel time (> 60 min.)	39	17	15	n.a.
Inconsistent times at stations/stops	n.a.	26	3	n.a.
Inconsistency between reported car distance and origin-destination pair	98	n.a.	n.a.	n.a.
Big difference between delay and time in queue (or wait time exceeds travel time) (> 30 min.)	25	2	n.a.	n.a.
Big difference between reported and computed PAT (> 30 min.)	10	5	n.a.	n.a.
Big difference between PAT and expected arrival time (> 90 min.)	18	2	2	n.a.
Implausible speed	54	45	n.a.	n.a.
Implausible car time	31	n.a.	n.a.	n.a.
Implausible car cost	14	n.a.	n.a.	n.a.
Implausible public transport fare or six or more transfers	n.a.	24	n.a.	n.a.
Implausible air fare	n.a.	n.a.	51	n.a.
Implausible number of locks or waiting time at lock	n.a.	n.a.	n.a.	19
No adults in household	3	11	3	-
Interview done too fast	56	100	22	1
Always same answer in SP	5	1	-	2
Chose non-intuitive alternative of dominant choice pair	135	137	42	14
Total excluded	**759**	**512**	**138**	**36**
Total remaining	**1895**	**1849**	**393**	**178**
% remaining	**71.4%**	**78.3%**	**74.0%**	**83.2%**

n.a. = not applicable

Table 37: Number of interviews per questionnaire before and after checks 2011 data (from en-route recruitment)

2009 data	Car	Public transport	Air	Recreational navigation
Data collected	**614**	**519**	**201**	**95**
Error in base cost or time calculation	-	-	-	-
Origin or destination municipality cannot be identified	12	5	n.a.	n.a.
Use of other modes than just car	18	n.a.	n.a.	n.a.
Travel time without delay exceeds real travel time	1	7	n.a.	n.a.
Travel time without delay exceeds expected travel time	5	18	n.a.	n.a.
Big difference between real and expected travel time	4	3	3	n.a.
Inconsistent times at stations/stops	n.a.	8	1	n.a.
Inconsistency between reported car distance and origin-destination pair	12	n.a.	n.a.	n.a.
Big difference between delay and time in queue (or wait time exceeds travel time)	4	-	n.a.	n.a.
Big difference between reported and computed PAT	3	2	n.a.	n.a.
Big difference between PAT and expected arrival time	2	-	-	-
Implausible speed	21	21	n.a.	n.a.
Implausible car time	17	n.a.	n.a.	n.a.
Implausible car cost	5	n.a.	n.a.	n.a.
Implausible public transport fare or six or more transfers	n.a.	9	n.a.	n.a.
Implausible air fare	n.a.	n.a.	8	n.a.
Implausible number of locks or waiting time at lock	n.a.	n.a.	n.a.	7
No adults in household	1	3	1	2
Interview done too fast	-	6	-	-
Always same answer in SP	1	-	-	1
Chose non-intuitive alternative of dominant choice pair	30	39	6	4
Total excluded	**136**	**116**	**19**	**14**
Total that remains	**478**	**403**	**182**	**81**
% that remains	**77.9%**	**77.6%**	**90.5%**	**85.3%**

n.a. = not applicable

When we compare the number of exclusions between the 2009 and the 2011 surveys, we see that in 2011 the fraction of respondents that was excluded is substantially lower than in 2009 for car and plane. For public transport and recreational navigation, the share of exclusions is rather similar between 2009 and 2011.

7.3.3 Trip characteristics

Table 38 shows the distribution of the remaining 5459 respondents after the selection procedure over the mode/purpose segments.

Table 38: Data base used for estimation: number of respondents per mode-purpose segment

			Purpose			
Year			Commute	Business	Other	Total
2009	Mode	Car	1008	349	538	1895
		Train	699	235	249	1183
		BTM	469	61	136	666
		Plane	0	96	297	393
		Recr.Nav.	0	0	178	178
	Total		**2176**	**741**	**1398**	**4315**
2011	Mode	Car	150	235	93	478
		Train	105	41	79	225
		BTM	97	11	70	178
		Plane	7	23	152	182
		Recr.Nav.	0	0	81	81
	Total		**359**	**310**	**475**	**1144**
Total	Mode	Car	1158	584	631	2373
		Train	804	276	328	1408
		BTM	566	72	206	844
		Plane	7	119	449	575
		Recr.Nav.	0	0	259	259
	Total		**2535**	**1051**	**1873**	**5459**

Table 39 displays statistics on travel time, cost and distance for the resulting samples in 2009 and 2011. For car, we find higher mean travel time and cost for 2011 than for 2009. This was expected, given that the recruitment in 2011 took place at petrol stations/service areas along motorways and in parking garages, whereas in 2009 recruitment took place by means of an existing internet panel. Recruiting along motorways gives a relative high inclusion probability for persons making long (and frequent) trips. This is not the case for parking garages, but the effect of recruiting along motorways appears to dominate,

For public transport, the mean travel time for the 2011 data on the contrary is smaller than for 2009 (though higher than in 1997). Also for air travel the mean transport time in the 2011 data is smaller than in 2009.

Table 39: Travel time, travel cost and distance in the 2009 and 2011 survey

		Car		Public transport				Plane	
				Train		BTM			
		2009	2011	2009	2011	2009	2011	2009	2011
Travel time (in minutes)	minimum	5	10	28	22	15	9	30	50
	maximum	240	285	335	245	230	150	1430	1270
	mean	49.4	60.9	102.6	91.9	54.4	53.2	375	275
	st. dev.	39.2	40.6	53.1	46.7	25.0	23.7	315	263
Travel cost (in euro)	minimum	0.10	0.10	0.5	1	0.5	1	30	30
	maximum	60	38	50	45	50	42.6	2000	1750
	mean	4.68	6.16	9.7	8.1	3.0	3.2	254.8	184.1
	st. dev.	4.12	6.04	7.6	7.2	3.1	4.6	224.6	215.8
Distance (in kilometer)	minimum	1	1	0	0	0	0	n.a.	
	maximum	300	290	220	202	123	60		
	mean	49.7	65.3	55.2	48.1	10.2	9.4		
	st. dev.	56.3	57.0	42.2	38.3	13.1	11.1		

7.3.4 Trading

For each of the experiments, we checked whether respondents were trading the attributes when making their choices or whether they always chose the cheapest of the fastest trip. As can be seen from Table 40 the fraction of non-traders in the 2011 data is quite low (only 5.7%) and is much lower than in the 2009 data. Note that these non-traders have been kept in the analysis.

Table 40: Trading behaviour

Year	Behaviour	Exp. 1	Exp. 2a	Exp. 2b	All exp.
2009	Always chose cheapest trip	30.4%	26.3%	26.0%	12.9%
	Always chose fastest trip	2.9%	5.6%	3.0%	0.7%
	Trading time/cost attributes	66.7%	68.1%	71.0%	86.4%
2011	Always chose cheapest trip	14.6%	14.7%	15.4%	4.0%
	Always chose fastest trip	8.0%	9.8%	7.0%	1.7%
	Trading time/cost attributes	77.4%	75.5%	77.6%	94.3%

7.4 RP data

During the interview, the respondents were asked whether they had any alternative route (in the case of car and recreational navigation respondents) or any alternative route or mode (public transport and air respondents). If so, they were asked about the mode (in the case of public transport and air respondents) and the time that route would have taken them. Also questions regarding the costs were asked.

Of the car respondents in the final sample, 38% indicated that they had an alternative route available, but only 16% were able to estimate the time that that route would have taken them. 43% of the public transport respondents had an alternative, but only 13% could indicate a travel time and even fewer respondents were able to estimate the costs. In air transport, the availability of alternative routes was even lower (9% of respondents, and only 6% with travel time information). In recreational navigation, the response was even lower.

CHAPTER 8 Passenger transport data analysis

8.1 Results for MNL mean-dispersion models

We start from a simple MNL utility function in WTP space without interaction terms for socio-economic influences, similar to equation [13] in section 2.3. In these models, we combine data from both surveys (2009 and 2011) and from all three SP experiments. The seven respondents in the Airplane – commute segment are most likely not conventional commuters and are combined with the Airplane – business segment.

Since we have data from two different recruitment methods, we estimate separate VOTs and reliability ratios for each survey. Scale factors are introduced to capture possible differences in error levels between the two surveys (where the scale factor for the 2009 survey is constrained to 1) and between the experiments (where the scale factor for experiment 2a data is constrained to 1). So, the utility function used is:

$$U = \left(Sc^{09} \cdot \delta^{09} + Sc^{11} \cdot \delta^{11}\right) \cdot \left(Sc^{\exp 1} \cdot \delta^{\exp 1} + Sc^{\exp 2a} \cdot \delta^{\exp 2a} + Sc^{\exp 2b} \cdot \delta^{\exp 2b}\right) \cdot$$
$$\beta_C \cdot \left[C + \left(VOT^{09} \cdot \delta^{09} + VOT^{11} \cdot \delta^{11}\right) \cdot \left(T + \left(RR^{09} \cdot \delta^{09} + RR^{11} \cdot \delta^{11}\right) \cdot \sigma\right)\right] \tag{26}$$

where Sc^{09} is the scale factor for the 2009 survey and δ^{09} is a dummy (0 or 1) that indicates whether an observation belongs to the 2009 survey, etc. Similar to our analysis of the freight survey, the travel time T is taken to be the mean travel time, rather than the median travel time (see Section 5.1.3).

All cost levels are corrected for inflation to 2010 levels: The 2009 cost levels are increased by 1.08% and the 2011 cost levels are decreased by 2.41% (consumer price index values taken from CBS 2012). Hence, the final VOTs will be in 2010 euros.

From the estimated coefficients (Table 41) we conclude:

- The VOT^{11} is generally about a factor 2 higher than the VOT^{09}, except for the Airplane – other and the Recreational navigation – other segments. This could indicate that we have recruited respondents with a higher intrinsic VOT in the 2011 survey compared to the 2009 survey. However, it is also possible that the difference is caused by other reasons (e.g. recruited from different income categories, or with trips of different lengths). This issue is discussed in more detail in Section 8.5.1.

Table 41: Estimated coefficients and t-ratios for MNL mean-dispersion models

		Commute		Business		Other	
Car	File name	Aut_ww_F001.F12		Aut_za_F001.F12		Aut_ov_F001.F12	
	Observ. (resp.)	20712 (1158)		10248 (584)		11238 (631)	
	Final log (L)	-12212.4		-5834.4		-6356.9	
	Rho²(0)	0.149		0.179		0.184	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-0.7711	(-22.7)	-0.5649	(-15.8)	-0.6984	(-17.0)
	VOT09	3.87	(31.0)	4.727	(20.1)	2.978	(22.0)
	VOT11	9.115	(18.4)	11.63	(19.6)	5.996	(12.3)
	RR09	0.5131	(8.5)	0.7706	(7.1)	0.6347	(6.9)
	RR11	0.1735	(2.1)	0.6001	(6.5)	0.3534	(2.4)
	Sc11	0.7	(12.4)	0.4867	(12.2)	0.631	(9.1)
	ScEx1	1.888	(14.4)	2.04	(10.2)	2.35	(10.4)
	ScEx2b	1.062	(15.7)	0.9812	(11.9)	1.177	(12.1)
Train	File name	Tre_ww_F001.F12		Tre_za_F001.F12		Tre_ov_F001.F12	
	Observ. (resp.)	14382 (804)		4932 (276)		5838 (328)	
	Final log (L)	-8382		-2872		-3163.7	
	Rho²(0)	0.159		0.159		0.218	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-0.3825	(-20.9)	-0.1697	(-12.3)	-0.3494	(-12.9)
	VOT09	5.817	(35.0)	10.39	(19.5)	3.982	(20.9)
	VOT11	8.796	(15.8)	22.22	(9.8)	5.02	(13.8)
	RR09	0.2945	(6.8)	0.4073	(5.0)	0.489	(5.3)
	RR11	0.1234	(1.5)	0.00256	(0.0)	0.0862	(0.8)
	Sc11	0.936	(11.6)	0.8025	(6.6)	1.246	(10.1)
	ScEx1	2.167	(14.7)	2.333	(9.3)	2.934	(8.8)
	ScEx2b	1.162	(15.0)	1.083	(8.6)	1.264	(10.1)
BTM	File name	BTM_ww_F001.F12		BTM_za_F001.F12		BTM_ov_F001.F12	
	Observ. (resp.)	10092 (566)		1296 (72)		3582 (206)	
	Final log (L)	-5534		-675.2		-1891.7	
	Rho²(0)	0.209		0.248		0.238	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-0.832	(-19.0)	-0.4309	(-7.0)	-0.9467	(-10.3)
	VOT09	2.669	(20.8)	3.485	(7.0)	2.055	(9.4)
	VOT11	5.276	(11.4)	6.566	(5.8)	3.658	(7.6)
	RR09	0.7517	(7.6)	2.588	(4.2)	1.281	(5.1)
	RR11	-0.00464	(-0.0)	0.4566	(1.9)	0.3743	(1.5)
	Sc11	0.5624	(9.9)	1.677	(3.9)	0.6292	(7.2)
	ScEx1	2.304	(11.0)	2.976	(3.0)	2.357	(6.6)
	ScEx2b	1.169	(13.0)	0.8927	(4.7)	1.221	(7.6)
Plane	File name			Vli_za_F001.F12		Vli_ov_F001.F12	
	Observ. (resp.)			2226 (126)		7524 (449)	
	Final log (L)			-1294		-4560.2	
	Rho²(0)			0.161		0.126	
				Value	(T-ratio)	Value	(T-ratio)
	BetaCost			-0.0163	(-6.1)	-0.03039	(-10.9)
	VOT09			33.54	(13.6)	24.48	(25.7)
	VOT11			79.93	(4.1)	29.72	(13.7)
	RR09			-0.03472	(-0.3)	0.05896	(1.0)
	RR11			0.4087	(1.6)	0.2933	(2.6)
	Sc11			0.6911	(3.0)	0.9097	(7.5)
	ScEx1			2.2	(4.2)	2.031	(8.2)
	ScEx2b			1.572	(4.7)	1.092	(8.6)
Recr. nav.	File name					Ple_ov_F001.F12	
	Observ. (resp.)					3102 (259)	
	Final log (L)					-1748.9	
	Rho²(0)					0.186	
						Value	(T-ratio)
	BetaCost					-0.3541	(-10.6)
	VOT09					5.561	(-24.4)
	VOT11					5.319	(-12.6)
	RR09					0.1585	(-0.5)
	RR11					-1.494	(-2.1)
	Sc11					0.8065	(-8.5)
	ScEx1					3.207	(-8.8)

- The RR^{11} is generally much lower than the RR^{09}. It might be consistent with a VOR^{11} that is not significantly different from a VOR^{09}. This is not investigated in more detail at this point.
- The Sc^{11} is usually less than 1, indicating that the unexplained variance in the 2011 model is greater than for the 2009 model.
- The Sc^{Ex1} is usually around 2, indicating that the unexplained variance in the experiment 1 responses is smaller than in experiment 2a.
- The Sc^{Ex2b} is usually not significantly different from one[21], indicating that the amount of unobserved error in experiment 2b is the same as in experiment 2a. This was expected, since the presentation of both experiments is identical.

8.2 Results for advanced MNL mean-dispersion models

In the next step we have estimated advanced MNL models (Section 2.6) that include sensitivity for higher base levels and subsequently diminishing values for smaller changes (such as lower VOT for small time savings). For this, we used utility function [23] with the scale factors, the VOT and the RR similar to the previous MNL models (see equation [26]).

For the reference points in equation [23], we chose to set $C_{ref} = 3$ euro, which is close to both the mean and median values for the BaseCost of the 2009 and 2011 car commute respondents, and we chose to set $T_{ref} = 40$ minutes, again close to the mean and median values for BaseTime of the 2009 and 2011 car commute respondents. Based on the mean and median values of ΔC and ΔT in our dataset, we chose $\Delta C_{ref} = 1$ euro and $\Delta T_{ref} = 5$ minutes.

As the base level for the standard deviation σ_0, we used the third (and middle) level of the five possible standard deviation levels as described in the design. As the reference level for the standard deviation σ_{ref} we used the base level for the standard deviation that corresponds to a base time level of 40 minutes.

The VOT and the RR are estimated separately for both surveys. However, the γ and λ coefficients (see equation [23]) are jointly estimated on both surveys, since we believe that these are more intrinsic to the population and early estimates show that these coefficients are indeed similar in both data sets. By doing so, we can still make optimal use of the large 2009 data set to estimate some coefficients (such as these γ and λ factors, and also in the next section the socio-economic interaction coefficients), without having a bias in the VOT and RR.

We estimated this model for all modes and all purposes. The results can be found in Table 42. A few remarks concerning these estimation results:

- Since all respondents had the same base values in the recreational navigation experiments, no λ terms could be estimated.

---

[21] Note that the t-ratios in Table 41 are with respect to zero.

- $\lambda < 0$ (which is what we find) means that as the base time (cost) increases people will be less sensitive to changes in these variables.
- $0 < \gamma < 1$ (which is what we find) means a sensitivity between a strong damping effect as the offered changes in time and cost increase (γ close to 0) and a linear sensitivity to time and cost changes (γ=1).
- The values for γ_C in car, train and BTM models are very consistently about 0.45. For plane and especially for recreational navigation, they are higher.
- The values for γ_T are very consistently about 0.9 (close to a linear effect). For recreational navigation, γ_T seems to be slightly higher, but it is probably not significantly different from the other models
- The value for λ_C ranges from -0.16 to -0.55 with most estimates approximately equal to -0.35.
- The value for λ_T ranges from -0.19 to -0.63 with most estimates approximately equal to -0.4.
- The value for λ_R ranges from -0.88 to -1.85.
- We do not find significant RR^{09} and RR^{11} values for airplane, nor a significant RR^{09} for recreational navigation. The RR^{11} for recreational navigation is negative and on the edge of being significant (t-ratio 2.0 without correcting for the panel effect). This is an unlikely result.
- The VOT_{ref}^{09} is between 50% and 80% of VOT_{ref}^{11}, which is a clear indication of the internet panel bias. Note that we have now corrected for a possible distance effect compared to the models estimated in Section 8.1, so this is a firmer result. The only exception is recreational navigation where we did not find a significant difference between VOT_{ref}^{09} and VOT_{ref}^{11}.

Please note that the VOT_{ref} should not be seen as the resulting mean VOT from these models, since the final VOT depends not only on the VOT_{ref}, but also on other factors, see equation [21]. The VOT_{ref} values in Table 42 cannot be directly compared to the VOT values in Table 41. The final mean VOT should be calculated by means of a weighting procedure (on the basis of the changes offered in the SP and the base levels in the trips at the population level). This weighting procedure is discussed in Section 8.9.

Table 42: Estimated coefficients and t-ratios for advanced MNL mean-dispersion models

		Commute		Business		Other	
Car	File name	Aut_ww_F002.F12		Aut_za_F002.F12		Aut_ov_F002.F12	
	Observ. (resp.)	20712 (1158)		10248 (584)		11238 (631)	
	Final log (L)	-10964.5		-5337.6		-5615.7	
	Rho²(0)	0.236		0.249		0.279	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-1.011	(-32.1)	-1.058	(-23.0)	-1.255	(-23.9)
	VOTref09	4.596	(28.7)	4.33	(18.3)	3.434	(16.7)
	VOTref11	7.827	(26.2)	7.418	(24.3)	5.498	(15.9)
	RR09	0.4902	(10.3)	0.8316	(8.0)	0.4725	(6.3)
	RR11	0.3495	(3.8)	0.9909	(8.1)	0.4816	(2.7)
	gammaC	0.4236	(26.2)	0.4971	(19.2)	0.4583	(20.7)
	gammaT	0.9098	(35.3)	0.9081	(24.8)	0.8841	(23.2)
	lambdaC	-0.3496	(-18.0)	-0.3564	(-13.4)	-0.3446	(-13.8)
	lambdaT	-0.3806	(-10.0)	-0.3912	(-6.7)	-0.4114	(-7.3)
	lambdaR	-1.126	(-13.7)	-1.148	(-12.9)	-0.9321	(-6.1)
	Sc11	0.7554	(18.4)	0.6158	(18.9)	0.6466	(14.4)
	ScEx1	2.975	(19.8)	2.839	(13.8)	3.034	(14.7)
	ScEx2b	1.324	(24.0)	1.207	(17.1)	1.38	(18.8)
Train	File name	Tre_ww_F002.F12		Tre_za_F002.F12		Tre_ov_F002.F12	
	Observ. (resp.)	14382 (804)		4932 (276)		5838 (328)	
	Final log (L)	-7808.9		-2691.7		-2974.8	
	Rho²(0)	0.217		0.212		0.265	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-0.8324	(-25.7)	-0.7066	(-12.9)	-0.792	(-14.9)
	VOTref09	4.609	(26.4)	5.456	(14.8)	3.878	(12.3)
	VOTref11	6.612	(20.3)	9.501	(11.1)	4.996	(14.3)
	RR09	0.7626	(5.8)	1.593	(4.9)	1.947	(3.7)
	RR11	0.5938	(2.8)	0.5719	(1.5)	0.4296	(1.3)
	gammaC	0.4564	(23.0)	0.463	(13.8)	0.4856	(16.0)
	gammaT	0.9577	(36.4)	0.8685	(18.8)	0.8867	(21.0)
	lambdaC	-0.303	(-15.1)	-0.3575	(-9.6)	-0.1613	(-6.4)
	lambdaT	-0.4108	(-10.0)	-0.3818	(-5.2)	-0.2221	(-3.5)
	lambdaR	-1.722	(-8.3)	-1.634	(-9.5)	-1.956	(-7.3)
	Sc11	0.7581	(14.7)	0.7255	(7.3)	0.9805	(12.3)
	ScEx1	3.78	(17.9)	4.024	(10.0)	4.129	(10.9)
	ScEx2b	1.41	(19.5)	1.387	(10.8)	1.391	(12.2)
BTM	File name	BTM_ww_F002.F12		BTM_za_F002.F12		BTM_ov_F002.F12	
	Observ. (resp.)	10092 (566)		1296 (72)		3582 (206)	
	Final log (L)	-5169.8		-569.5		-1730.9	
	Rho²(0)	0.261		0.366		0.303	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-1.046	(-23.0)	-0.9017	(-9.0)	-1.253	(-13.4)
	VOTref09	4.372	(17.4)	4.673	(7.2)	4.312	(9.5)
	VOTref11	6.29	(16.3)	7.858	(7.6)	5.819	(10.3)
	RR09	0.5778	(7.1)	1.61	(3.7)	0.6416	(4.6)
	RR11	0.03573	(0.2)	0.7097	(2.2)	0.2602	(1.7)
	gammaC	0.4418	(17.9)	0.45	(8.7)	0.4141	(11.8)
	gammaT	0.9387	(23.7)	0.9007	(9.9)	0.8104	(13.2)
	lambdaC	-0.2457	(-8.2)	-0.4792	(-5.2)	-0.2324	(-5.3)
	lambdaT	-0.585	(-7.9)	-0.1943	(-1.1)	-0.4482	(-4.4)
	lambdaR	-1.451	(-8.1)	-1.01	(-3.6)	-0.7018	(-3.1)
	Sc11	0.6141	(13.2)	1.114	(6.4)	0.6634	(11.6)
	ScEx1	3.179	(13.2)	3.792	(5.4)	2.893	(7.8)
	ScEx2b	1.348	(16.6)	1.118	(6.4)	1.431	(10.4)

		Commute	Business		Other	
Plane	File name		Vli_za_F002.F12		Vli_ov_F002.F12	
	Observ. (resp.)		2226 (126)		7524 (449)	
	Final log (L)		-1173.4		-4202.2	
	Rho²(0)		0.239		0.194	
			Value	(T-ratio)	Value	(T-ratio)
	BetaCost		-1.386	(-4.2)	-2.212	(-6.0)
	VOTref09		2.984	(4.8)	2.222	(6.8)
	VOTref11		6.098	(5.1)	2.922	(7.4)
	RR09		-0.00994	(-0.6)	1.41	(1.2)
	RR11		0.0871	(0.4)	0.7095	(1.9)
	gammaC		0.609	(9.9)	0.5776	(18.1)
	gammaT		0.9903	(14.1)	0.9263	(28.5)
	lambdaC		-0.5688	(-8.0)	-0.6149	(-12.4)
	lambdaT		-0.7132	(-9.0)	-0.6537	(-13.7)
	lambdaR		-0.1289	(-0.1)	-1.851	(-3.4)
	Sc11		0.5524	(5.1)	0.7507	(14.4)
	ScEx1		2.976	(6.5)	2.784	(12.5)
	ScEx2b		1.511	(7.1)	1.257	(13.9)
Recr. nav.	File name				Ple_ov_F002.F12	
	Observ. (resp.)				3102 (259)	
	Final log (L)				-1739.3	
	Rho²(0)				0.191	
					Value	(T-ratio)
	BetaCost				-0.2759	(-9.0)
	VOTref09				6.714	(-7.4)
	VOTref11				6.379	(-7.0)
	RR09				0.1819	(-0.5)
	RR11				-1.487	(-2.0)
	gammaC				1.273	(-20.4)
	gammaT				1.04	(-15.3)
	Sc11				0.8375	(-9.3)
	ScEx1				3.329	(-9.2)

Given the similarities of the estimates between modes, we also estimated models per purpose, combining car, train and BTM, as in the national VOT studies in 1988-1990 and 1997-1998. To allow for possible differences between the three modes, we replaced the VOT term

$$\left(VOT_{ref}^{09}\cdot\delta^{09}+VOT_{ref}^{11}\cdot\delta^{11}\right)$$

by:

$$\left(VOT_{ref}^{09}\cdot\delta^{09}+VOT_{ref}^{11}\cdot\delta^{11}\right)\cdot\left(1+facTrein\cdot\delta^{Trein}+facBTM\cdot\delta^{BTM}\right) \qquad [27]$$

and estimated values for facTrain and facBTM. The results are shown in Table 43.

Given the low number of respondents, we have estimated a combined model for both purposes with a facBus (similar to facTrain and facBTM in equation [27]) to allow the VOT_{ref} for Business to differ between from the VOT_{ref} for other purposes. The result is also shown in Table 43.

Table 43: Estimated coefficients and t-ratios for combined advanced MNL mean-dispersion models

		Commute		Business		Other	
Car / Train / BTM	File name	All_ww_F002A.F12		All_za_F002A.F12		All_ov_F002A.F12	
	Observ. (resp.)	45186 (2528)		16476 (932)		20658 (1165)	
	Final log (L)	-24002.8		-8645.6		-10388.2	
	Rho²(0)	0.234		0.243		0.275	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-0.9615	(-48.3)	-0.9217	(-28.6)	-1.132	(-31.5)
	VOTref09	4.663	(41.6)	4.469	(24.0)	3.657	(22.9)
	VOTref11	7.265	(35.2)	7.583	(30.4)	5.345	(24.2)
	RR09	0.4948	(15.1)	0.9364	(11.8)	0.5266	(9.3)
	RR11	0.2922	(4.2)	0.8827	(8.5)	0.2895	(3.2)
	facTrein	-0.01231	(-0.6)	0.192	(4.5)	0.06915	(2.0)
	facBTM	-0.1034	(-4.5)	0.2434	(4.0)	0.1325	(3.0)
	gammaC	0.4312	(39.2)	0.4585	(25.2)	0.4425	(28.3)
	gammaT	0.9278	(56.4)	0.8833	(32.5)	0.8591	(33.9)
	lambdaC	-0.3266	(-26.3)	-0.386	(-19.2)	-0.2892	(-17.4)
	lambdaT	-0.4161	(-16.7)	-0.4376	(-10.5)	-0.3617	(-9.7)
	lambdaR	-1.2	(-21.1)	-1.252	(-21.4)	-0.9658	(-11.6)
	Sc11	0.7119	(26.9)	0.6882	(22.2)	0.6942	(22.3)
	ScEx1	3.28	(30.3)	3.308	(18.3)	3.372	(19.9)
	ScEx2b	1.356	(35.3)	1.262	(21.2)	1.398	(24.8)
Plane	File name			Vli_al_F002.F12			
	Observ. (resp.)			9750 (575)			
	Final log (L)			-5389.8			
	Rho²(0)			0.202			
				Value	(T-ratio)		
	BetaCost			-1.972	(-7.5)		
	VOTref09			2.392	(8.5)		
	VOTref11			3.343	(9.4)		
	RR09			0.9708	(1.5)		
	RR11			0.6016	(2.4)		
	facBus			0.06705	(1.5)		
	gammaC			0.5837	(21.2)		
	gammaT			0.9339	(32.2)		
	lambdaC			-0.6013	(-15.3)		
	lambdaT			-0.6686	(-16.6)		
	lambdaR			-1.674	(-4.8)		
	Sc11			0.722	(15.8)		
	ScEx1			2.858	(14.1)		
	ScEx2b			1.307	(15.8)		

8.3 Results for advanced MNL scheduling models

So far, we have only estimated mean-dispersion models (see Section 2.1.2). As described in Section 2.1.3 unreliability can also be modelled with a schedule delay early and a schedule delay late term. In this section, we will compare the results of the mean-dispersion models, of the scheduling models and of the combined mean-dispersion/scheduling models (Section 2.1.4). In this Section, we also distinguish between scheduling models in WTP space and in log WTP space (see Section 2.4).

Note that all scheduling terms are mean early/late arrival terms (and not median, see Section 5.1.3) with respect to the preferred arrival time (which is not always equal to the expected arrival time).

8.3.1 Scheduling models in WTP space

We focussed this analysis on the commute segment (car/train/BTM). The mean-dispersion model is repeated in first column. The second column presents the result of a scheduling model. The last column shows a combined mean-dispersion / scheduling model. Note that the λ_R coefficient in scheduling model is not relevant, since no RR term is included.

The scheduling model (model F013) has a strong coefficient for Late arrival, while the coefficient for Early is not significant, where one would expect a significant negative influence of being too early. We tried various other scheduling model specifications, but this problem did not go away.

The mean-dispersion model outperforms the scheduling model on the basis of the log-likelihood test. The combined mean-dispersion / scheduling model is slightly better than the mean-dispersion model. However, in that model the t-ratios for both the reliability ratios and the Late coefficient go down. In a pure mean-dispersion model, the reliability ratios pick up more of the unreliability effect.

Since the difference in log-likelihood is small and since we might get into interpretation problems when using a model with both dispersion and scheduling terms (see Section 2.1.4), we prefer the mean-dispersion model.

Table 44: Estimated coefficients and t-ratios for mean-dispersion and scheduling models in WTP space

		Mean-dispersion		Scheduling		Combined mean-dispersion / scheduling	
Car / Train / BTM	File name	All_ww_F002A.F12		All_ww_F013.F12		All_ww_F012.F12	
	Observ. (resp.)	45186 (2528)		45186 (2528)		45186 (2528)	
	Final log (L)	-24002.8		-24058.4		-23993.8	
	Rho²(0)	0.234		0.232		0.234	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-0.9615	(-48.3)	-0.9582	(-47.7)	-0.9694	(-48.0)
	VOT09	4.663	(41.6)	4.837	(44.1)	4.71	(41.7)
	VOT11	7.265	(35.2)	7.422	(36.5)	7.315	(35.4)
	RR09	0.4948	(15.1)	-		0.3972	(9.7)
	RR11	0.2922	(4.2)	-		0.246	(3.4)
	ValArrEarl	-		0.09733	(0.3)	-0.3861	(-1.2)
	ValArrLate	-		0.9878	(12.6)	0.3731	(3.5)
	facTrein	-0.01231	(-0.6)	-0.01175	(-0.5)	-0.01267	(-0.6)
	facBTM	-0.1034	(-4.5)	-0.1042	(-4.4)	-0.1026	(-4.4)
	gammaC	0.4312	(39.2)	0.4263	(39.0)	0.4297	(39.1)
	gammaT	0.9278	(56.4)	0.924	(55.7)	0.9179	(55.6)
	lambdaC	-0.3266	(-26.3)	-0.3188	(-26.0)	-0.3245	(-26.1)
	lambdaT	-0.4161	(-16.7)	-0.4505	(-18.1)	-0.4145	(-16.6)
	lambdaR	-1.2	(-21.1)	-		-1.353	(-20.5)
	Sc11	0.7119	(26.9)	0.7126	(27.1)	0.7108	(26.9)
	ScEx1	3.28	(30.3)	3.363	(30.7)	3.276	(30.2)
	ScEx2b	1.356	(35.3)	1.341	(34.9)	1.343	(35.2)

8.3.2 Scheduling models in logWTP space

We repeated these three models in logWTP space. To improve the convergence we have constrained the γ_T coefficient to 1. The results are shown in Table 45.

In earlier analyses for passenger transport with mean-dispersion models (i.e. before the γ- and λ-coefficients were added) models estimated in logWTP space usually performed better than models in WTP space. Now, with extra terms for the influence of the base levels and the size of the changes offered, the models in WTP space both perform better than their counterparts in logWTP space.

We conclude that it is preferable to estimate these advanced MNL models in WTP space rather than in logWTP space, because they perform better and because we do not have to worry about situations when the argument of the logarithm becomes zero or negative.

Note that in logWTP space the scheduling model has a slightly better loglikelihood than the mean-dispersion model. Again, the Early coefficient is not significant and the reliability ratios are much smaller in a combined mean-dispersion / scheduling model compared to a pure mean-dispersion model.

Table 45: Estimated coefficients and t-ratios for mean-dispersion and scheduling models in logWTP space

		Mean-dispersion		Scheduling		Combined mean-dispersion / scheduling	
Car / Train / BTM	File name	All_ww_F009F.F12		All_ww_F009H.F12		All_ww_F009I.F12	
	Observ. (resp.)	45186 (2528)		45186 (2528)		45186 (2528)	
	Final log (L)	-24059.9		-24053		-24031.4	
	Rho²(0)	0.232		0.232		0.233	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-184.1	(-43.6)	-186.6	(-43.8)	-185.7	(-43.7)
	VOT09	4.442	(63.4)	4.484	(64.7)	4.424	(62.4)
	VOT11	7.105	(39.2)	7.083	(39.6)	7.08	(38.9)
	RR09	0.4428	(14.2)	-		0.2301	(5.7)
	RR11	0.2175	(3.7)	-		0.09374	(1.4)
	ValArrEarl	-		0.3619	(1.0)	0.05134	(0.1)
	ValArrLate	-		1.574	(16.2)	1.046	(7.9)
	facTrein	-0.03258	(-1.6)	-0.03918	(-1.9)	-0.0372	(-1.8)
	facBTM	-0.1005	(-5.7)	-0.1091	(-6.1)	-0.1054	(-5.9)
	gammaC	0.4796	(43.4)	0.4844	(43.0)	0.4843	(43.3)
	gammaT	1	(*)	1	(*)	1	(*)
	lambdaC	-0.2501	(-17.8)	-0.2348	(-16.8)	-0.2448	(-17.3)
	lambdaT	-0.5	(-24.1)	-0.4831	(-22.8)	-0.4899	(-23.2)
	lambdaR	-0.6773	(-9.8)	-		-1.018	(-10.6)
	Sc11	1.159	(23.9)	1.138	(25.2)	1.155	(24.0)
	ScEx1	3.105	(30.8)	3.099	(30.8)	3.051	(30.4)
	ScEx2b	1.315	(34.3)	1.298	(34.1)	1.293	(34.2)

8.4 Results for advanced MNL mean-dispersion models with socio-economic interactions

It is likely that the VOT will differ for different segments in population. To capture this form of heterogeneity, we have estimated models with socio-economic interaction terms for age, gender, education, income, household composition and peak/off-peak travelers. When we find significant differences between these groups, it will be important to take the correct weights into account when calculating the average VOT for the whole population. Therefore, we tested for interaction terms that were available in both this survey and in the national travel survey OViN (see Section 8.9). Furthermore, we tested for similar interaction factors to those used in the 1988 and 1997 surveys.

The interaction terms were introduced as multiplicative factors with the reference VOT. For instance for gender, the VOT_{ref} term is multiplied by a factor:

$$\left(1+ facFem\cdot\delta^{Fem}\right) \qquad [28]$$

where facFem is the interaction coefficient to be estimated and δ^{Fem} a dummy variable that equals 1 for female respondents and 0 for male respondents. An estimated value for facFem of 0.1 means that the VOTref for female respondents is 10% higher than for male respondents.

BetaCost and the lambda and gamma terms were not interacted. Since we estimate a reliability ratio (RR), the VOR will go up and down together with the VOT and so the VOR is interacted with the same factors as the VOT. For segments with a higher VOT, there will also be a higher VOR. The results are presented in Table 46.

For airplane, and for recreational navigation we did not estimate socio-economic interaction terms, since in the absence of a national travel survey for these modes, no re-weighting can be done anyway. For recreational navigation we have set both the RR09 and the RR11 to zero.

Many socio-economic interaction terms were tested. When a certain term had a t-ratio of less than 2, we re-estimated the model and checked whether the loglikelihood was significantly better given the reduction in the degrees of freedom. Also, when two similar terms had similar values, we tested whether combining the two terms improved the model. In one case, the facEdu1 interaction factor (for those respondents in the lowest education class) was less than -1 (though not significantly less than -1). Given equation [28], this might lead to negative VOTs, therefore this term was combined with the facEdu2 term (for the second education class), though this combination did not pass the loglikelihood tests. However, we felt that the resulting model was more plausible. Similarly, we removed three terms (two about household composition and one facBTM) that had low t-ratios (between 2 and 3) and which had counterintuitive signs. Given the fact that these t-ratios were not yet corrected for the panel-effect and were therefore likely to be an overestimation of the real t-ratios, we felt that the model would be more plausible by removing them.

Table 46: Estimated coefficients and t-ratios for combined advanced MNL mean-dispersion models with socio-economic interaction terms

		Commute		Business		Other	
Car / Train / BTM	File name	All_ww_F003.F12		All_za_F003.F12		All_ov_F003.F12	
	Observ. (resp.)	45186 (2528)		16476 (932)		20658 (1165)	
	Final log (L)	-23846		-8540.9		-10260.7	
	Rho²(0)	0.239		0.252		0.283	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-0.9741	(-48.5)	-0.9244	(-28.8)	-1.147	(-31.5)
	VOTref09	4.991	(34.0)	4.565	(22.9)	4.811	(20.7)
	VOTref11	7.452	(29.9)	6.74	(25.4)	6.505	(21.1)
	RR09	0.4973	(15.1)	0.9615	(12.1)	0.5194	(9.5)
	RR11	0.28	(4.0)	0.8924	(8.7)	0.2673	(3.0)
	facTrain	-		-		-0.1037	(-3.6)
	facBTM	-0.07422	(-3.4)	-		-	
	fac3650	-0.08343	(-4.5)	-		-0.1098	(-3.6)
	fac51pl	-0.1721	(-9.2)	-0.1397	(-5.4)	-0.2788	(-11.1)
	facEdu1	-0.4724	(-4.7)	-		-	
	facEdu2	-0.1111	(-3.4)	-		-	
	facEdu12	-		-0.3021	(-6.3)	-0.2223	(-6.8)
	facEdu34	-0.08009	(-4.6)	-		-0.1258	(-5.1)
	facFem	-		0.1064	(3.3)	0.135	(4.3)
	facHH1	0.1064	(4.1)	-		-	
	facHH12	-		0.2036	(4.5)	-	
	facInc	0.06598	(10.7)	0.1002	(9.9)	0.03094	(3.5)
	facPeak	0.06608	(3.3)	0.1536	(4.7)	0.1008	(3.1)
	gammaC	0.4357	(40.0)	0.4631	(26.0)	0.4448	(29.2)
	gammaT	0.936	(57.5)	0.8854	(33.3)	0.8642	(34.8)
	lambdaC	-0.3246	(-26.1)	-0.369	(-18.5)	-0.2848	(-17.6)
	lambdaT	-0.4204	(-17.6)	-0.3546	(-9.2)	-0.3046	(-8.4)
	lambdaR	-1.239	(-21.8)	-1.217	(-22.0)	-0.9299	(-11.0)
	Sc11	0.7181	(26.9)	0.6976	(22.4)	0.677	(22.2)
	ScEx1	3.235	(30.6)	3.292	(18.8)	3.406	(20.3)
	ScEx2b	1.337	(35.3)	1.237	(21.1)	1.385	(24.7)
Plane	File name			Vli_al_F003.F12			
	Observ. (resp.)			9750 (575)			
	Final log (L)			-5402.4			
	Rho²(0)			0.201			
				Value	(T-ratio)		
	BetaCost			-1.972	(-7.5)		
	VOTref09			2.392	(8.5)		
	VOTref11			3.343	(9.4)		
	RR09			0.9708	(1.5)		
	RR11			0.6016	(2.4)		
	facBus			0.06705	(1.5)		
	gammaC			0.5837	(21.2)		
	gammaT			0.9339	(32.2)		
	lambdaC			-0.6013	(-15.3)		
	lambdaT			-0.6686	(-16.6)		
	Sc11			-1.674	(-4.8)		
	ScEx1			0.722	(15.8)		
	ScEx2b			2.858	(14.1)		
Recr. nav.	File name					Ple_ov_F003.F12	
	Observ. (resp.)					3102 (259)	
	Final log (L)					-1742.4	
	Rho²(0)					0.189	
						Value	(T-ratio)
	BetaCost					-0.2784	(-9.4)
	VOTref					6.637	(7.6)
	RR					-	
	gammaC					1.275	(20.5)
	gammaT					1.041	(15.4)
	Sc11					0.8509	(9.2)
	ScEx1					3.269	(9.5)

In terms of socio-economic interaction terms, we find the following significant influences on the VOT (and, through the RR, also on the VOR):

- fac3650: travellers in the age class 36-50 have a lower commuting and other VOT (than the younger age classes, which form the reference category).
- fac51pl: travellers in the age class 51 and older have a lower VOT than those younger than 36 (and also than those in the age class 36-50), for all three purposes.
- facEdu1: travellers with primary school as highest education have a lower commuting VOT than those with high education levels (College University)
- facEdu2: travellers with lower secondary school as highest education have a lower commuting VOT than those with high education levels (College/University)
- facEdu34: travellers with medium/higher secondary school as highest education have a lower commuting and other VOT than those with high education levels (College/University)
- facEdu12: travellers with primary school or lower secondary school as highest education have a lower other VOT than those with high education levels (College/University)
- facFem: females have a higher business and other VOT (possibly since they are often involved more in multi-tasking).
- facHH1: households with only one member have a higher commuting VOT than other households (we expect that these households cannot share other tasks with other members, so they have higher opportunity costs for travel).
- facHH12: households with one adult with or without children have a higher business VOT than other households (we expect that these households have more difficulty in sharing other tasks with other members than other households, so they have higher opportunity costs for travel).
- facInc: we tested several dummies for income classes, and these revealed an almost linear pattern for the income interactions with the VOT. Subsequently we used linear income: higher incomes have a higher VOT for all purposes.
- facPeak: trips in the peak (midpoint of trip falls within 7-9 hours or 16-18 hours) have a higher VOT for al purposes; this may have to do with the additional nuisance of travelling in congested/crowded conditions.
- facBTM: we estimated the same model for all three modes and tested whether mode-specific dummies were still significant (reference alternative is car). Only for bus/tram/metro for commuting (and train for other, see below) did we find a significant (negative) estimate. For all other purposes, a distinction between modes in the estimated coefficients is no longer needed: other coefficients (e.g. income) pick up differences in behaviour between modes. However, for producing the recommended VOTs, we will be using sample enumeration and expansion to national mobility figures for each mode and purpose, and for instance differences

in trip length between modes (e.g. longer trips by train than for other modes) can then lead to differences in the final VOTs between modes.

- facTrain: train travellers have a lower other VOT (possibly because they can use their time in the train more pleasantly and productively than in other modes, using information technology).

8.5 Additional tests

8.5.1 Differences between experiments

For the three surface mode segments (commute, business, other), we have estimated a VOT[11] for each of the three experiments (1, 2a and 2b) separately using a simple MNL that also allowed for a reliability ratio. For each purpose, these VOTs were averaged (using the inverse variance of the estimated VOT as weight factor). Figure 13 shows the difference between the VOT of each experiment and the average VOT over the three experiments. Error bars indicated one standard deviation.

From this figure it is clear that there is no significant difference between the three experiments (within 2 standard errors). It has been hypothesized that the VOT of experiment 1 would be larger than of experiment 2a and 2b since in the latter two the value of reliability is estimated explicitly and in the former one it may be part of the VOT. However, as can be seen from Figure 13, there is no indication whatsoever that this hypothesis is true. There is no evidence that the VOT from experiment 1 contains any value of reliability.

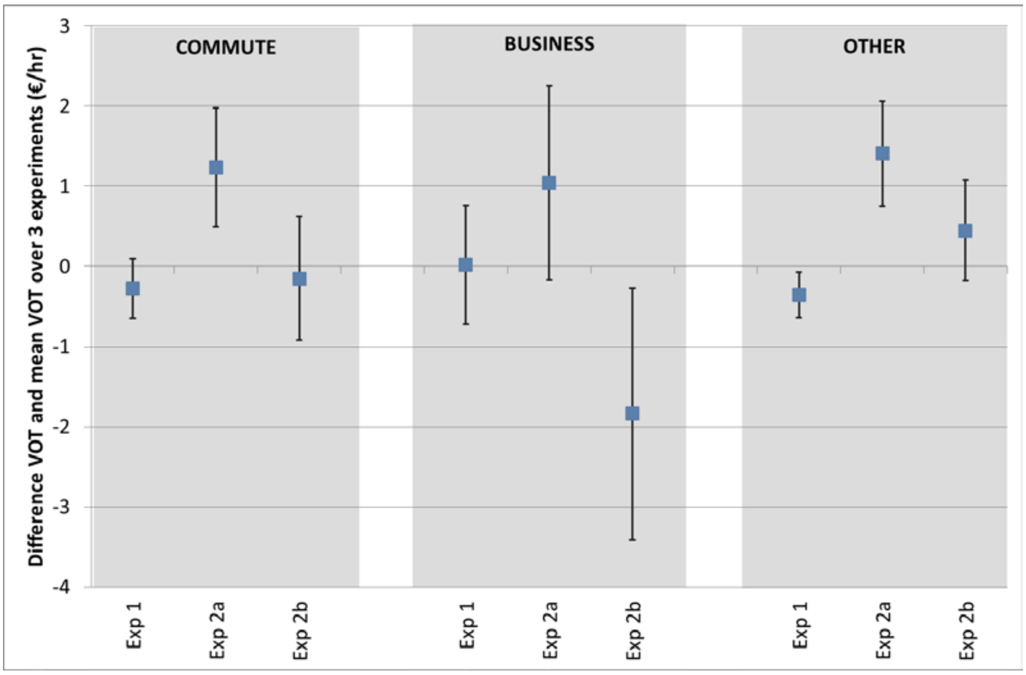


Figure 13: Differences in the VOT between the experiments

8.5.2 Member of internet panel

All respondents of the 2009 survey were members of the PanelClix internet panel. In the 2011 survey, we asked respondents whether they were a member of an internet panel, and if so, whether they were a member of the PanelClix panel (see Section 7.2.3). We have estimated a simple MNL model (similar to the one used for Table 41), an advanced MNL model (see Table 43) and an advanced MNL with socio-economic interaction terms (see Table 44) with the VOT11 factor split between panel members and non-panel members. The number of respondents that were specifically member of the PanelClix panel was too small to allow a separate VOT for this group. Table 47 shows the results for the VOT^{11}.

Table 47: VOT for respondents that are / are not a member of an internet panel

Mode	Mode	Purpose	VOT09 MEMBER OF PANELCLIX PANEL		VOT11 INTERNET MEMBER PANEL		VOT 11 NO MEMBER OF PANEL	
			Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
Simple MNL	Car/Train/BTM	Commute	4.33	(37.2)	5.36	(9.2)	8.79	(22.9)
	Car/Train/BTM	Business	5.89	(21.9)	11.56	(11.3)	12.38	(20.6)
	Car/Train/BTM	Other	3.20	(26.8)	4.65	(8.6)	5.14	(16.3)
	Airplane	All	24.90	(24.9)	34.71	(6.3)	32.55	(12.8)
	Recr. nav.	Other	5.57	(24.3)	5.09	(3.9)	5.32	(12.0)
Advanced MNL	Car/Train/BTM	Commute	4.66	(41.6)	6.08	(17.1)	7.57	(33.5)
	Car/Train/BTM	Business	4.49	(24.1)	6.84	(16.8)	7.79	(28.6)
	Car/Train/BTM	Other	3.66	(22.8)	5.02	(13.5)	5.42	(23.2)
	Airplane	All	2.39	(8.6)	2.80	(6.9)	3.42	(9.4)
	Recr. nav.	Other	6.74	(7.4)	6.24	(3.8)	6.38	(6.8)
Advanced MNL with socio-econ. interaction factors	Car/Train/BTM	Commute	4.98	(34.0)	6.21	(16.1)	7.77	(28.7)
	Car/Train/BTM	Business	4.59	(23.0)	6.05	(15.9)	6.95	(24.2)
	Car/Train/BTM	Other	4.81	(20.7)	6.54	(12.5)	6.50	(20.4)
	Airplane	All	2.39	(8.6)	2.80	(6.9)	3.42	(9.4)
	Recr. nav.	Other	6.74	(7.4)	6.24	(3.8)	6.38	(6.8)

Especially for the car/train/BTM – commute segment, we see that with a simple MNL model the VOT^{11} for the panel members is much lower than for the non-members. The VOT^{11} for the panel members is close to the VOT^{09} for the PanelClix panel members. For the advanced MNL model, we see that the VOT_{ref}^{11} for panel members is exactly between the VOT_{ref}^{09} and the VOT_{ref}^{11} for non-members. This indicates that a part of the difference between the VOT_{ref}^{11} for panel members and non-members as seen from the simple MNL model is due to a different basetime / basecost level, but not all. Adding socio-economic factors does not change this. So, there remains an intrinsic difference between members and non-members.

Also for the car/train/BTM – business segment and the airplane segment we see a clear difference between members and non-members in the advanced MNL models. For the car/train/BTM – other segment and the recreational navigation – other segment, there is no significant difference.

This means that in the 2009 survey there was a bias towards low-VOT people that have the time to be a member of an internet panel, and an additional bias towards those internet panel members that were willing to give up the time required to actually participate in this survey. Probably, people who cannot afford much (after correction for income) are more likely to participate in an internet panel and to fill in the web questionnaire for a reward. These people have a relatively low VOT.

We conclude that the 2009 SP survey leads to substantially lower VOTs than the 2011 SP survey. Those for 2011 are much more in line with the values found on the basis of the surveys in 1988 and 1997, which have always been regarded as very plausible values by the various transport sectors, and are not considered to be particularly high in an international perspective. The meta-analysis of Shires and de Jong (2009) found comparable or higher values for many other Western countries compared to the Dutch values from 1997. Our conclusion is that the most likely explanation is that the 2011 values are correct and that the 2009 values are biased downwards, mainly because persons with a lower value of time (in every socio-economic segment) have a higher probability of becoming a member of an internet panel (and within those panel members the ones that are most likely to participate in this relatively long survey are the ones with even lower values of time). The shorter distances of the trips sampled in 2009 (and the corresponding smaller time savings offered in the SP) will also have played a role (see below) in the downward bias in the 2009 VOTs.

More sophisticated models that were estimated on the separate data sets for 2009 and 2011 showed smaller differences than the above basic models: part of the difference can be explained by, for instance, differences in the trip lengths between the 2009 and 2011 data sets. However, the differences between the VOTs from both data sets remain substantial and significant: there is clear evidence for a downward bias in the VOTs in the 2009 data.

8.5.3 Different choices in November 2009

In the 2011 survey, we asked respondents whether they would have made different choices in the SP experiments if they had completed the questionnaire in November 2009. 13% answered positively. If we estimate a multiplicative factor (analogous to the factors for modes and gender in equations [27] and [28]) for those respondents we find that their VOT is not significantly different from the respondents who answered negatively (i.e. who indicated that they would have made the same choices in November 2009).

Table 48: Factor on VOT for 2011 respondents who said they would have given different answers in November 2009

Mode	Purpose	facDiff09	
		Value	(T-ratio)
Car / Train / BTM	Commute	-0.0533	(-0.9)
	Business	-0.02989	(-0.4)
	Other	-0.03943	(-0.6)
Airplane	All purposes	0.2268	(1.2)
Recr.nav.	Other	0.3067	(1.4)

8.5.4 Type of unreliability

As explained in Section 6.2.3 the cause of unreliability was sometimes presented as due to variable in-vehicle times and sometimes due to variable waiting times. We estimated whether public transport respondents (and airplane respondents) had a different VOT for both types of unreliability. As can be seen from the middle column in Table 49, none of these factors is significant.

Next we tried a similar factor on the reliability ratio. As can be seen from the right column in Table 49, none of these factors is significant either.

Table 49: Factor on VOT for public transport respondents who were presented with unreliability due to variable in-vehicle times compared to variable waiting times

Mode	Purpose	facTypeReliab on VOT		facTypeReliab on RR	
		Value	(T-ratio)	Value	
Car / Train / BTM	Commute	0.00324	(0.2)	-0.08065	(-0.6)
	Business	0.03924	(1.1)	-0.04288	(-0.3)
	Other	-0.0524	(-1.8)	0.2364	(0.9)

8.5.5 Number of transfers

We also tested, for train and BTM, whether the number of transfers or a dummy for one or more transfers) during the observed trip had an impact on the VOT, but did not get acceptable results (we obtained wrong signs and/or insignificant estimates for this variable, and inconsistencies between purposes).

8.5.6 Impact of exclusions

As discussed in Section 7.3, we excluded 1445 respondents from the 2009 survey (25.1%) and 285 respondents from the 2011 survey (19.9%) of which some were excluded based on their answer to the dominant choice pair. In this section, we investigate the impact of these exclusions to the final coefficient estimates.

For each segment, we have three model runs: the final run from Table 46, the same run with the inclusion of all respondents that were excluded based on the dominant question, and finally the same run with the inclusion of all respondents that were excluded on the basis of the other criteria. Table 50 presents the results for the VOT^{11} and RR^{11} for all runs and for all segments.

If the choices from the excluded respondents contained valid information, the t-ratios should increase with the square root of the increase in the number of respondents. For instance: the number of respondents in the car/train/BTM commute segment increases by 7.8% when the exclusion criterion on the dominant question is left out. So, one would expect an increase in the t-ratio of 3.8% for the VOT^{11} and RR^{11}. The VOT^{11} increases slightly (but not significantly) and its t-ratio increases as well, but by less than 3.8%. This indicates that the choices from the added respondents contain more noise. This is what we see in other segments as well, both for the VOT^{11} and the RR^{11}. For the car/train/BTM commute segment and the airplane segment, we also see that the RR^{11}increases by about 50%, however, this is not significant. We conclude that the exclusion of the respondents who gave a non-intuitive answer to the dominant question indeed removed some noise from the model estimates.

When we include all respondents in the analysis (i.e. also the outliers), we see that this has a bigger impact on the VOT^{11} and RR^{11}. This was to be expected, since these outliers were left out for the reason that they had a possibly large impact on the estimates. Especially in the airplane segment, the impact is large. Again, we conclude that these excluded respondents were not random exclusions (containing the same level of information per respondent as the remaining data), but were valid exclusions.

Table 50: VOT$_{ref}^{11}$ and RR with the inclusion of previously excluded respondents

		Final run		+ dominant			+ outliers		
Car/Train/BTM - commute	Respondents	45186		48714			58494		
				+7.8%		Expect: +3.8%	+29.5%		Expect: +13.8%
		Coeff.	(t-ra)	Coeff.	(t-ra)		Coeff.	(t-ra)	
	VOTref11	7.45	(29.9)	7.75	(30.6)	+2.3%	8.21	(33.9)	+13.4%
	RR11	0.28	(4.0)	0.16	(2.4)	-40.0%	0.14	(2.5)	-37.5%
Car/Train/BTM - business	Respondents	16476		17556			22296		
				+6.6%		Expect: +3.2%	+35.3%		Expect: +16.3%
		Coeff.	(t-ra)	Coeff.	(t-ra)		Coeff.	(t-ra)	
	VOTref11	6.74	(25.4)	6.72	(25.6)	+0.8%	7.15	(28.5)	+12.2%
	RR11	0.89	(8.7)	0.88	(8.5)	-2.3%	0.78	(9.0)	+3.4%
Car/Train/BTM - other	Respondents	20658		22062			28596		
				+6.8%		Expect: +3.3%	+38.4%		Expect: +17.7%
		Coeff.	(t-ra)	Coeff.	(t-ra)		Coeff.	(t-ra)	
	VOTref11	6.51	(21.1)	6.74	(22.0)	+4.3%	6.82	(24.8)	+17.5%
	RR11	0.27	(3.0)	0.27	(3.1)	+3.3%	0.23	(3.2)	+6.7%
Airplane - all	Respondents	9750		10560			12966		
				+8.3%		Expect: +4.1%	+33.0%		Expect: +15.3%
		Coeff.	(t-ra)	Coeff.	(t-ra)		Coeff.	(t-ra)	
	VOTref11	3.34	(9.4)	3.61	(9.7)	+3.2%	8.26	(14.3)	+52.1%
	RR11	0.60	(2.4)	0.44	(1.7)	-29.2%	0.28	(0.6)	-75.0%
Recr. nav. - other	Respondents	3102		3318			3696		
				+7.0%		Expect: +3.4%	+19.1%		Expect: +9.2%
		Coeff.	(t-ra)	Coeff.	(t-ra)		Coeff.	(t-ra)	
	VOT	6.63	(7.6)	6.50	(7.6)	0.0%	6.55	(7.4)	-2.6%

8.5.7 Gains and losses

We have performed additional tests to check whether gains and losses are valued differently. First, we tested whether the models improved if the γ_C and γ_T exponents were split into separate exponents for gains and losses. The resulting estimated values were not significantly different from each other, i.e. the γ_C and γ_T are not different for gains and losses.

Next, we estimated separate values for the VOT for the four quadrants of ΔC and ΔT (see Figure 1). These are displayed in Table 51. Note that the theoretical prediction (WTA: WTP < EG, EL < WTA, see Section 2.5) indeed holds for all segments. It is predicted that $\sqrt{WTP \cdot WTA} = \sqrt{EG \cdot EL}$ (de Borger & Fosgerau, 2008). Within the margins of error, this equality holds as well and they are both equal (again, within the margins of error) to the value of VOT11 as found before (Table 46). Therefore, we keep that model specification.

Table 51: WTP, EG, EL, WTA for each segment

	Car/train/BTM - commute	Car/train/BTM - business	Car/train/BTM - other	Airplane - all
VOT11	7.5	6.7	6.5	3.3
WTP11	6.8	6.4	8.9	3.1
EG11	7.5	6.6	6.3	3.2
EL11	7.3	6.9	6.9	3.4
WTA11	8.7	7.4	7.5	3.6
Sqrt(WTP11·WTA11)	7.7	6.9	6.7	3.3
Sqrt(EG11·EL11)	7.4	6.7	6.6	3.3

8.6 Mixed logit results

8.6.1 Introduction

In this section, we extend the analysis of the preceding sections by performing panel mixed logit estimations where some of the parameters are distributed in the population. The previous MNL estimations took into account that respondents do not all have the same VOT by introducing covariates. However, it may be that some of the heterogeneity is not observed, meaning that respondents with the same covariates have different VOTs. Cross-sectional mixed logit estimations allow for the estimation of a distribution of VOTs rather than one number. Panel mixed logit are able to control for the panel structure of the data (i.e. respondents answer a series of SP questions). We will not discuss the mathematics and details of the mixed models and refer to Train (2009) for more technical details. All models were estimated in WTP space.

We investigated three types of mixed logit models. The first approach models unobserved heterogeneity with continuous distributions, where we define the shapes of the distributions beforehand. We tried normal, lognormal, uniform, triangular, and Johnson's S_B distributions. Continuous distributions led to implausible results as we will discuss later on. Therefore, we proceeded with two other methods to account for unobserved heterogeneity. First, we tried to predefine a grid of the parameters and estimate the probability of each combination of the parameters.[22] The problem with this approach is that the grid is predefined by the analyst, and the results proved to be rather sensitive to the choice of the grid. Therefore, we turned to a more flexible approach: a panel latent-class model where we estimate both the (discrete mixture) distribution and the parameters. This mixed logit model is very flexible and does not impose any pre-defined shape of the distribution.[23, 24] Before we present our main results, we will briefly discuss the other approaches.

8.6.2 Continuous distribution panel mixed logit

We started with panel mixed logit models with continuous distributions for the unobserved heterogeneity of the parameters. For example, we tried a model with a normally distributed VOT. The mean and the standard deviation were estimated.[25] Other distribution shapes we tried were lognormal, normal, uniform triangular, and Johnson's S_B.

---

[22] For example, if the VOT and VOR are heterogeneous on a 7 by 7 grid, then there are 49 probabilities to be estimated. See Bajari et al. (2007), Train (2008, Sections 6 and 7) and Train (2009, Section 14.3.1.) for further technical details on this model (although, differently from these authors, we estimate the model by maximum likelihood, and not by an inequality constrained regression or expectation maximization algorithm).

[23] See, for instance, Swait (1994), Bhat (1997), Train (2008, Sections 4 and 7), and Train (2009, Section 14.3.2), for further discussion of the latent class mixed logit.

[24] The continuous distribution mixed logits were estimated using Biogeme 2.1/2.2. The discrete heterogeneity models by Pythonbiogeme 2.2. See Bierlaire (2003) and http:/ biogeme.epfl.ch for further information.

[25] We looked at the 3 types of set-up for the attributes, where all of them had cost and expected travel time as attributes. The first also included the standard deviation of the travel time, the second the schedule delays, and the thirds the standard deviation and schedule delays.

This continuous distribution approach proved problematic, probably because the distribution is misspecified, and therefore no results are reported in this section. The mean estimates of the random parameters were very sensitive to the chosen shape of the distribution. Moreover, when we used distributions that are bounded to be positive for the VOT (which prevents the unreasonable outcome that some people would prefer longer travel times and higher unreliability) the mean VOT was really high. When we did not impose this constraint, we found that for an implausibly large share of the respondents the VOTs would be negative. For instance, when we used normal distributions for the VOT and the VOR for the mixed logit model for car commuters we found that 37% had a negative VOR, for the 2009 VOT this was 35% and for the 2011 VOT 16%.

Part of the problem for the continuous distributions might to be due to how the real VOTs are distributed since there may be a fraction of the respondents that only take into account the costs; while there may be also a large group that is not really interested in the costs but is very sensitive to the expected travel time and/or standard deviation of the travel time. This can be investigated using a more flexible approach where no predefined shape of the distribution is assumed.

8.6.3 Discrete distribution panel mixed logit – fixed points of the distribution, but estimated class probabilities

Because of the problems with continuous distributions, we turned to mixed logit models with discrete distributions (see also Hess et al. (2011) and Hensher et al. (2011c)). We analysed two cases. First, we chose the VOTs of the distribution, and estimated the probabilities of each combination of VOTs. Second, we estimated both the VOTs and the probabilities. The second method proved most useful, and is discussed in more detail in the next section.

The advantage of assuming the VOTs beforehand is that there are fewer parameters to estimate, meaning that a more detailed distribution can be estimated (i.e. one can estimate a distribution with more classes). The disadvantage is that for this dataset the mean VOTs are very sensitive to the choice of the points. Train (2008) encountered similar problems for his dataset. This is why we prefer the method of the next section, which also estimates the points.

Still, it is worth noting that in these estimations we found that a large fraction of the respondents had a very low, or even zero, VOT or VOR. There is also a large group with a substantial VOT, but very low, or even zero, RR; a smaller group with a substantial VOR and small VOT; and a group with high VOT and VOR.

8.6.4 Discrete distribution panel mixed logit – estimated points and class probabilities of the distribution

We proceed with the estimation of discrete distributions for the VOT and estimate the reliability ratio for 2009 and 2011.[26] To take advantage of the sample size of the 2009 data, it is assumed that the distribution probabilities are the same for the 2009 and 2011 data, whereas the values for the VOT for a given (class) probability may be different.

---

[26] We also tested models in logWTP space and models with a mixing distribution on the reliability ratio. This led to unstable results.

Furthermore, it is assumed that covariates have the same proportional effect on the VOT and VOR in 2009 and 2011. We estimate these models for commuting, business and other travel separately where we also include the covariates that were found to be significant in the MNL estimation. For each of the estimations we optimize the number of classes using a statistical criterion.[27] The optimal number of classes is 5 for commuting, 4 for business, other and air travel and 2 for recreational navigation. The results are given in Table 52. The t-ratios in this table are so-called 'robust' t-ratios, which allow for non-severe misspecification errors (Bierlaire, 2008). Note that for recreational navigation no stable estimation of the gammas could be obtained, therefore they were constrained to one.

---

[27] We used the Bayesian Information Criterion for this. In order to avoid local optima, we also used multiple runs with different starting values and report the estimated model with the highest loglikelihood.

The loglikelihood is calculated as follows: For each class g (of the G classes), we first calculate the probability of the chosen alternative j (of J alternatives), of choice situation k (of K choices), of individual i (of N persons):

$$P_{kjig} = \frac{\exp(\beta_{ig} \cdot X_{kji})}{\sum_{l=1}^{l=J} \exp(\beta_{ig} \cdot X_{kli})}.$$

Then, we calculate the probability of the K choices made by individual i, for class g, as

$$P_{ig} = \prod_{k=1}^{k=K} P_{kjig} = \prod_{k=1}^{k=K} \frac{\exp(\beta_{ig} \cdot X_{kji})}{\sum_{l=1}^{l=J} \exp(\beta_{ig} \cdot X_{kli})}.$$

The likelihood of the sequence of choices for individual i over all classes G is

$$L_i = \sum_{g=1}^{g=G} \rho_g \cdot P_{ig} ;$$

where ρ_g is the estimated latent probability of being in class g. Note that the only difference with standard panel mixed logit is in this step: in the standard case, one calculates the expected L_i by integrating over one or more distributions (usually the integration is done by simulation); with latent class, one calculates the expected likelihood by summation.

This makes the log-likelihood for our panel latent class:

$$LL = \sum_{i=1}^{i=N} \ln(L_i) = \sum_{i=1}^{i=N} \ln\left(\sum_{g=1}^{g=G} \rho_g \cdot \left(\prod_{k=1}^{k=K} \frac{\exp(\beta_{ig} \cdot X_{kji})}{\sum_{l=1}^{l=J} \exp(\beta_{ig} \cdot X_{kli})} \right) \right);$$

where *ln(x)* is the natural log of x.

Table 52: Estimated coefficients and t-ratios for combined advanced LC mean-dispersion models with socio-economic interaction terms

		Commute		Business		Other	
Car / Train / BTM	File name	All_ww_F003-LC.F12		All_za_F003-LC.F12		All_ov_F003-LC.F12	
	Observ. (resp.)	45186 (2528)		16476 (932)		20658 (1165)	
	Final log (L)	-22103.134		-7925.771		-9471.395	
	Rho²(0)	0.33		0.362		0.156	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-1.16	(-35.2)	-1.05	(-20.6)	-1.45	(-22.6)
	VOTref09_0	4.06	(19.8)	38.1	(6.4)	0.168	(0.8)
	VOTref09_1	69.3	(5.1)	9.37	(12.6)	10.5	(14.0)
	VOTref09_2	0	(0.0)	4.18	(10.9)	5.01	(15.1)
	VOTref09_3	15.3	(9.2)	0	(0.0)	163	(3.1)
	VOTref09_4	7.75	(19.8)	-		-	
	VOTref11_0	5.57	(19.3)	44.1	(5.9)	6.82	(10.3)
	VOTref11_1	57.3	(2.0)	1.31	(1.3)	39.1	(4.5)
	VOTref11_2	11.9	(10.7)	5.19	(12.3)	1.68	(1.3)
	VOTref11_3	46.3	(2.6)	12.2	(7.1)	12.4	(2.3)
	VOTref11_4	0	(0.0)	-		-	
	Group_0	0	(*)	0	(*)	0	(*)
	Group_1	-2.55	(-21.6)	1.17	(4.5)	-1.5	(-5.4)
	Group_2	-0.33	(-3.8)	2.28	(14.1)	-0.13	(-0.8)
	Group_3	-2.42	(-8.8)	1.58	(9.5)	-3.06	(-9.0)
	Group_4	-0.775	(-6.5)	-		-	
	RR09	1.17	(11.5)	1.51	(10.4)	1.2	(3.4)
	RR11	0.408	(2.2)	1.15	(6.8)	0.624	(1.3)
	facTrain					-0.106	(-1.9)
	facBTM	-0.0891	(-2.5)				
	fac3650	-0.107	(-4.1)			-0.0396	(-0.7)
	fac51pl	-0.186	(-6.5)	-0.104	(-1.9)	-0.233	(-4.8)
	facEdu1	-0.331	(-1.8)				
	facEdu2	-0.11	(-2.5)				
	facEdu12			-0.284	(-3.8)	-0.0826	(-1.0)
	facEdu34	-0.0409	(-1.5)			-0.0911	(-2.5)
	facFem			0.0062	(0.1)	0.0465	(0.8)
	facHH1	0.138	(3.4)				
	facHH12			0.165	(2.0)		
	facInc	0.0761	(8.1)	0.109	(6.3)	0.0273	(1.8)
	facPeak	0.0789	(2.8)	0.188	(3.0)	0.123	(2.4)
	gammaC	0.523	(40.9)	0.548	(28.4)	0.537	(28.6)
	gammaT	1.06	(48.8)	1.01	(31.2)	1.05	(18.8)
	lambdaC	-0.386	(-22.1)	-0.473	(-16.7)	-0.382	(-16.0)
	lambdaT	-0.526	(-14.8)	-0.515	(-10.5)	-0.559	(-7.3)
	lambdaR	-1.05	(-7.2)	-1.19	(-15.4)	-0.86	(-4.1)
	Sc11	0.723	(20.6)	0.683	(16.8)	0.636	(16.8)
	ScEx1	4.02	(24.7)	4.34	(15.1)	3.52	(14.1)
	ScEx2b	1.33	(25.1)	1.35	(15.3)	1.37	(18.6)

		Commute	Business	Other
Plane	File name		Vli_al_F003-LC.F12	
	Observ. (resp.)		9750 (575)	
	Final log (L)		-4952.262	
	Rho²(0)		0.164	
			Value (T-ratio)	
	BetaCost		-3.32 (-5.1)	
	VOTref09_0		3.74 (5.9)	
	VOTref09_1		0 (0.0)	
	VOTref09_2		1.79 (5.7)	
	VOTref09_3		9.83 (4.9)	
	VOTref11_0		0.593 (1.0)	
	VOTref11_1		6.68 (5.7)	
	VOTref11_2		2.36 (4.8)	
	VOTref11_3		35 (0.6)	
	Group_0		0 (*)	
	Group_1		-0.0667 (-0.2)	
	Group_2		0.762 (1.9)	
	Group_3		-1.66 (-4.9)	
	RR09		1.35 (1.1)	
	RR11		0.653 (1.4)	
	facBus		0.0295 (0.4)	
	gammaC		0.64 (22.6)	
	gammaT		1.01 (24.9)	
	lambdaC		-0.722 (-13.2)	
	lambdaT		-0.796 (-11.6)	
	lambdaR		-1.62 (-3.3)	
	Sc11		0.72 (12.8)	
	ScEx1		4.16 (10.7)	
	ScEx2b		1.28 (12.1)	
Recr. nav.	File name			Ple_ov_F003-LC.F12
	Observ. (resp.)			3102 (259)
	Final log (L)			-1660.0
	Rho²(0)			0.25
				Value (T-ratio)
	BetaCost			-0.336 (-9.5)
	VOTref09_0			15.9 (8.1)
	VOTref09_1			5.15 (19.1)
	VOTref09_2			1.39 (4.2)
	VOTref11_0			0.291 (0.4)
	VOTref11_1			6.39 (13.0)
	VOTref11_2			40.8 (4.5)
	Group_0			0 (*)
	Group_1			1.54 (6.2)
	Group_2			-0.632 (-1.7)
	Sc11			1.01 (6.6)
	ScEx1			5.75 (6.9)

From these results we can conclude that the parameters for reference dependence are still significant, except for the γ_T variable which is not significantly different from 1 in the models for business and other travel. All the covariates have the intuitive signs although some of the covariates are not significant any more. The reliability ratios for the 2009 and 2011 surveys are significantly different from each other and also significantly different from 1. In Table 53 the mean values and the distributions of the $\text{VOT}_{\text{ref}}{}^{11}$ are calculated. This is done using the Group variable. The probability of a person belonging to group *i* is given by:

$$Prob(i) = \frac{e^{Group(i)}}{\sum_i e^{Group(i)}} \qquad [29]$$

All these average values are significantly higher than the estimated values for the MNL models. How this carries over to the final sample VOT and VOR depends on the sample enumeration and will be discussed in Section 8.9.

Table 53: Mean values for the reference value of time and for reliability ratio from the panel-latent class mixed logit model

	Car/Train/BTM Commute		Car/Train/BTM Business		Car/Train/BTM Other		Airplane All purposes		Recr. nav. Other	
Class	VOT$_{ref}^{11}$	*Prob.*	VOT$_{ref}^{11}$	*Prob.*	VOT$_{ref}^{11}$	*Prob.*	VOT$_{ref}^{11}$	*Prob.*	VOT$_{ref}^{11}$	*Prob.*
1	5.57	*42.6%*	44.1	*5.3%*	6.82	*46.6%*	0.593	*23.4%*	0.291	*16.2%*
2	57.3	*3.3%*	1.31	*17.1%*	39.1	*10.4%*	6.68	*21.9%*	6.39	*75.3%*
3	11.9	*30.6%*	5.19	*51.9%*	1.68	*40.9%*	2.36	*50.2%*	40.8	*8.6%*
4	46.3	*3.8%*	12.2	*25.8%*	12.4	*2.2%*	35	*4.5%*		
5	0	*19.6%*								
Average	9.680		8.396		8.194		4.347		8.360	
RR11	0.408		1.15		0.624		0.653		0	

8.7 Discussion on the use of MNL or panel-LC mixed logit models

From the literature it is well known that MNL models have two important limitations which might bias the estimation results.

1. **Unobserved heterogeneity**
   Different decision-makers are likely to have different values of time, as has been confirmed by many value of time studies. To some degree, modellers can account for this by including an influence of observed factors (such as income and education of the respondents) on the time or cost coefficients (or directly on the value of time coefficient, if the model is specified in WTP-space). This is called 'observed heterogeneity'. But there usually are more differences between decision-makers that affect the VOT than can be explained by available exogenous variables. This additional variation is called 'unobserved heterogeneity'. The influence of this on the VOT can be included by:

   - regarding these coefficients as random variables and making assumptions on the distribution function of these (e.g. normal, lognormal, triangular, etc.) in a continuous distribution mixed logit model (see Section 2.7), or by

   - estimating a number of points (classes) of the distribution and the class membership probabilities in a latent class model (see Section 2.8).

2. **Panel effects**
   Every respondent made 18 different choices in our experiment. MNL models assume that these choices are independent, i.e. they assume that every choice was made by a different individual. However, this is not the case. It is plausible that a respondent who makes large mistakes in the first choice will also do this in the

> remaining choices. This has two effects. First, the standard errors of the estimated VOTs are underestimated by the MNL models (t-values are overestimated). It is possible to correct for this, for instance with a jack-knife technique. Second, the VOTs could be biased, when the estimated MNL coefficients are higher or lower than the true values. This could occur if the size of the margin of error is related to other variables in the model (e.g. if persons with a lower education make more mistakes). The MNL models cannot be corrected for such an estimation bias. There is some evidence that the effect of ignoring the panel structure is usually larger for the standard errors than for the coefficient values (de Jong et al., 2007).

In the last 15 years various techniques have been developed to solve these two limitations. Mixed Logit models estimate a distribution of the VOT for all respondents instead of a single VOT. However, Mixed Logit models also have limitations. An important limitation is that the researcher has to assume the shape of the distribution (e.g. normal or lognormal distribution) and this assumption could strongly affect the final result. In addition, this technique does not always yield a stable estimation, which happened to be case in our study as well.

A possible solution is the use of semi- and non-parametric estimation techniques. For example, Fosgerau (2006a) estimates a non-parametric distribution of the value of time using data from the Danish value of time experiment. However, he assumes a cross-sectional model. As discussed in his paper, it would be more appropriate, but difficult, to account for the repeated nature of the choices. This would allow for a better separation of errors and taste heterogeneity.

Although under certain theoretical assumptions cross-sectional models may give the same results as models accounting for the panel structure, the difference between cross-sectional and panel models is mainly an empirical issue. We compared our panel-LC estimation with a cross-sectional LC estimation and found significant differences in the $\mathrm{VOT}_{\mathrm{ref}}^{11}$ and the RR^{11}. The panel-LC model resulted in a substantially better model fit and is more plausible. Therefore panel estimation is preferred.[28]

Furthermore, the state-of-the-art methods applied to the Danish, Swedish and Norwegian valuation studies can only be applied when there is only one variable to estimate (i.e. the VOT). Since we are also interested in estimating the VOR (or RR) it is not possible to use these methods.

Panel-LC models assume that different classes of travellers exist in the data, each with their own VOT. The MNL model can be viewed as a latent class model with a single class. The panel-LC model estimates for every class the share of respondents belonging to this particular class. The mean VOT for the whole sample is then the weighted average of the VOTs of the classes, where the estimated class values are weighted by the respective share of the classes. This weighted average VOT usually differs from the VOT estimated by MNL models, because the panel-LC models lack the aforementioned limitations: respondents with equal (observed) characteristics could very well belong to different classes

---

[28] For commuters, we found a final log-likelihood of the cross-sectional model with 5 classes of -23703.7, an estimated VOT11ref of 11.52 and an estimated VOR11ref of 2.59. This suggests that the cross-sectional model overestimates the VOT and underestimates the VOR.

(and thus have a different VOT), and the panel-LC model estimates the probability that a series of 18 choices is made instead of analysing every choice independently (like the MNL model). The optimal number of classes is estimated by different panel-LC models by estimating a different number of classes every time and then determining which is the optimal model (and thus which number of classes).

The panel-LC models and the MNL models use the same specification of utility. Both the MNL and the panel-LC models are logit models and based on the Random Utility Model (RUM) framework (for the necessary 'mild' assumptions, see McFadden and Train, 2000). The panel-LC models are a bit more advanced, but their methods do not fundamentally differ.

The best panel-LC models for commuters have 12 parameters more than the MNL models[29], but the loglikelihood of the panel-LC models is much better than the loglikelihood of the MNL models and far better than you would expect from adding 12 parameters. For instance, the improvement in the loglikelihood in the commute model is more than 1700 points. This is a significant improvement according to both the standard loglikelihood test and the (stricter) Bayesian Information Criterion, which uses an additional penalty for extra parameters and is often used for this type of models.

In Hess et al. (2011), LC models are compared against mixed logit models with a continuous distribution, with a positive outcome for LC models. Similar conclusions on LC models were reached in Greene and Hensher (2012).

A limitation of the panel-LC models is that it is more difficult to estimate the optimal model. The loglikelihood function is quite flat close to the optimum and has multiple local optima, at which (depending on the chosen initial values) the algorithm might terminate. Furthermore, empirical identification may sometimes be hard due to non-traders. Identification is enhanced because of covariates and the non-linearity of the utility function. At the lower end of the distribution we restrict the VOT and VOR to be larger or equal than 0. For commuters we find a share of approximately 20% that have a VOTref11 of 0. At the higher end the values are still identified but not very significant. The reported values are the best estimates of dozens of runs, each with different initial values, so we are convinced that these values are very close to the global optimum.

Since

- it is known that MNL models might be biased
- it is known that panel-LC models correct for this bias
- the underlying assumptions on the respondents' utilities for the panel-LC models and the MNL models are the same and therefore little extra methodological uncertainty is introduced
- earlier studies in Denmark, Norway and Sweden abandoned the use of MNL models

---

[29] The number of estimated values for VOT-2011 increases from 1 to 5, idem for VOT-2009. Also, in order to estimate the share of every class 4 values have been estimated. So, 3 x 4 = 12 extra parameters in total. The panel-LC models for business and other purpose have four classes, so 3 x (4 – 1) = 9 extra parameters.

we recommend using the estimation results from the panel-LC.

8.8 Results for the RP data

In the car questionnaires there was a question on an alternative route (and its travel time and distance), in the train and BTM questionnaires there was a question about an alternative route or mode (and its travel time and cost). Such information can in principle be used to estimate an RP model. However, only a limited fraction of the respondents gave useable answers to these questions (also we do not have several RP choices for the same respondent, as we have for the SP). MNL models estimated on these RP observations did not yield acceptable results (also after removing outliers in terms of time and cost): in many cases, the estimated coefficients were not statistically significant, or the values of time was clearly below the range of plausible values. Therefore, we did not proceed with RP models or combined SP/RP models.

8.9 Expansion of the survey outcomes for passengers using OViN

The expansion is done using a weighted sample enumeration procedure. This involves the determination of a value of time (first step) and a weight factor (second step) for each respondent. For step 1 we use equation [21] with interactions as in [28]. For each respondent his BaseTime (T_0), BaseCost (C_0) and his socio-economic status are input for this equation. Since it is believed that the dependency on the ΔT and ΔC are (partly) an SP artefact and may also lead to difficulties in a CBA (since a 2 € change then would no longer be twice a 1 € change), we apply equation [21] with $\gamma_C = \gamma_T = 1$. Equation [21] now no longer explicitly depends on ΔT and ΔC. However, the value of VOT_{ref} does depend on the chosen values of ΔT_{ref} and ΔC_{ref}, so a sensible choice for these reference values is required. We solve this by calculating for each respondent the mean time and cost differences between the two alternatives over the 18 choices that he has been asked to make in SP experiments. These means are used as ΔT_{ref} and ΔC_{ref} in equation [22] to derive a sensible VOT_{ref} per respondent.

We validated this method by also estimating advanced MNL models with γ_C and γ_T constrained to 1. Now, equation [21] can be used without having to bother about the choice for ΔT_{ref} and ΔC_{ref}. The average VOT over the sample remained similar (within 10%) to the average VOT calculated using the method described above.

In the second step, we make our survey representative for the mobility of the Dutch population. For this, we divided all trips in the OViN (Onderzoek Verplaatsingen in Nederland 2010 by Statistics Netherlands) based on five population variables (gender, age, income, household composition and education) and two trip variables (period of the day combined with travel mode and travel time category combined with travel mode). The OViN survey contains approximately 136,000 records. Every record in the OViN survey also has a weight factor in order to make the OViN survey representative for all trips of the Dutch population in one year. For this, we only considered people older than 16 and used car, train or bus/tram/metro as method of transportation for their trip. We created three different datasets, one for each of the following purposes of travel: commuting, business and other.

The distribution of the trips in our survey over the seven variables is different from the whole Dutch population (e.g. many more commuters are present in the VOTVOR survey than one would expect in a typical sample). An Iterative Proportional Fitting method was used to calculate new weights for our VOTVOR survey such that the weighted distributions for the seven variables match the weighted distributions of the OViN survey.

In calculating the final values of time, the value of time for each respondent is weighted with this weight factor as determined by this expansion procedure, and with the travel time.

8.10 The employers' component in the value of time for business travel

The new values for business travel presented earlier only refer to the employee's component of the business VOT. The employer's part of the business VOT is calculated on the basis of information from the VOT surveys, as was done for the 1988 and 1997 surveys. We have no information to calculate an employer's VOR; the business VORs in this report will only be based on the valuation by the employee, we have no data to estimate business VORs, see also Section 8.11.

In Section 8.6, we report means from the 2011 sample. For the calculation of the recommended VOTs, both for the employee and the employer components, we use sample enumeration with expansion to national totals from the national travel survey 2010 (OViN), so the means given here are indicative only.

In Table 54, we compare new outcomes from the 2011 survey with those of the 1997 survey (Hague Consulting Group, 1998).

Table 54: Fraction of journey time spent working by mode (%TW)

	1997			2011		
	mean	stdev.	cases	mean	stdev.	cases
Car	0.0351	0.1057	866	0.0359	0.1332	246
Train	0.1613	0.2243	226	0.1569	0.2338	41
BTM	0.0259	0.1035	69	0.0597	0.1121	11
Airplane				0.1356	0.2704	26
Total	**0.0591**	**0.1458**	**1161**	**0.0600**	**0.1703**	**324**

Note that the outcomes for the 2009 survey are not realistic: they are much lower than in 1997, or even 1988, for all modes. We therefore decided to use 2011 for the business VOT inputs.

In 2011, almost the same part of journey time of business travellers using car or train is spent working as in 1997. For BTM this fraction has increased a lot, but it remains much smaller than for train. In planes it is higher than in cars and slightly lower than in trains.

Table 55: Relative productivity of work during travel by mode (%PTW)

	1997			2011		
	mean	stdev.	cases	mean	stdev.	cases
Car	0.9311	0.2128	145	0.9053	0.1930	22
Train	0.9025	0.1774	96	0.9405	0.1548	14
BTM	0.8889	0.1721	6	0.8333	0.2887	3
Airplane				1.0000	-	6
Total	**0.9189**	**0.1987**	**247**	**0.9149**	**0.1739**	**45**

For train the productivity of time spent working during travel has increased and now approaches the productivity of working at the workplace. For plane we have equivalence between the productivity at work and travelling, but here we have very few observations in 2011. This is after truncating the mean to be not higher than 1.

In 1988 and 1997, the share of time that would be used for all activities for the employer (which was defined as Work plus Study) was used in the calculation of the employer VOT. This had dropped somewhat for all three modes in 2011. For plane it is much lower than for the other modes.

We then derive the percentage of saved time that would be spent working using the following assumptions (as in 1997):

1. If the respondent stated that he/she would work or study but selected no other tasks it was assumed that 100% of the saved time would be spent working
2. If the respondent stated that he/she would work or study and selected other tasks as well it was assumed that 50% of the saved time would be spent working
3. If the respondent stated that he/she would not work or study it was assumed that 0% of the saved time would be spent working

Table 56: Percentage of saved time that would be spent working by mode (%W)

	1997		2011	
	factor	cases	factor	cases
Car	0.5445	866	0.5589	246
Train	0.3695	226	0.3780	41
BTM	0.3406	69	0.5454	11
Airplane			0.2115	26
Total	**0.4983**	**1161**	**0.5077**	**324**

The productive values of an hour worked (see Table 57), calculated on the basis of the income and number of workers per household information of 2011 is for car drivers slightly lower than the 1997 outcome (after conversion to euros using a factor 0.454 and inflation correction for 1997-2011 using a factor 1.375). For train and BTM, the 2011 values are clearly higher than in 1997 and more in line with those for car. The productive value for air travellers is the highest of all modes.

Table 57: Productive value of a unit of work time by mode (PVWT)

	1997 in 2010 euros			2011 in 2010 euros		
	mean	stdev.	cases	mean	stdev.	Cases
Car	34.84	23.22	866	31.17	15.49	22
Train	27.26	22.23	226	37.78	13.46	14
BTM	21.12	17.73	69	33.14	5.97	3
Airplane				39.12	16.66	6
Total	**32.55**	**22.70**	**1161**	**34.42**	**14.60**	**45**

We now use the formula (derived from the often used "Hensher formula"; see Gunn, 2008):

$$EPRVOT = PVWT \cdot \left(\%W - \%TW \cdot \%PTW\right) \qquad [30]$$

where:

EPRVOT: employers' component of the business VOT

PVWT: productive value of a unit of work time to the employer

%W: proportion of time savings returned to work in the work place

%TW: proportion of travel time spent working

%PTW: relative productivity of work undertaken while travelling.

This all leads to the following mean employer's VOT for 2011 (Table 58, last but one column).

Table 58: Mean employer's VOT in 2010 euros per hour*

	PVWT (in € 2010)	%W	%TW	%PTW	Mean VOT 2011 (in € 2010)	Mean VOT 1997 (in € 2010)
Car	31.17	0.5589	0.0359	0.9053	16.41	18.31
Train	37.78	0.3780	0.1569	0.9405	8.71	5.14
BTM	33.14	0.5454	0.0597	0.8333	16.43	8.16
Airplane	39.12	0.2115	0.1356	1.0000	2.97	
Total	**34.42**	**0.5077**	**0.06**	**0.9149**	**15.59**	**15.14**

* in the recommended VOTs we do not use these means, but a mean employers' VOT based on sample enumeration and expansion to OViN

In the last column of Table 58 are the mean employer's values of time of 1997. We see a small decline in values for car travel. For train we see an increase, which is mainly due to a higher productive value per hour. The BTM value is much higher than in 1997, as a result of the higher share that would be spent working and the higher productive value of an hour. For plane the employer's VOT is quite low, mainly because only a fifth of the time savings would be allocated to work time.

8.11 Final VOT and VOR results after expansion using OViN

After applying the OViN expansion weight factors and after weighting with the travel time, we can calculate the final VOTs. These are presented in Table 59 after rounding to the nearest multiple of 25 eurocents. This table also contains the VOTs for air travel and recreational navigation (after expansion, based on our own survey, since no information concerning these modes is available in OViN). All these VOTs include VAT, since for the respondents the VAT is included in the travel costs (fuel costs, public transport fares etc.).

Table 59: VOTs (in 2010 euros per hour) after expansion to OViN and weighted with travel time as found in this study

	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Commute	9.25	11.50	7.75	9.75		
Business employee	*12.75*	*15.50*	*10.50*	*13.50*	*85.75*	
Business employer	*13.50*	*4.25*	*8.50*	*10.50*	-	
Business	26.25	19.75	19.00	24.00	85.75	
Other	7.50	7.00	6.00	7.00	47.00	8.25
All purposes	9.00	9.25	6.75	8.75	51.75	8.25

Notes:

- All values are rounded off to the nearest multiple of € 0.25.
- These values include VAT.
- VOTs for "all purposes" are calculated by weighting the VOTs for each purpose with the total minutes travelled in OViN 2010 (surface modes) or our own survey (air). The VOTs for "all surface modes" are calculated similarly from the VOTs for car, train and bus/tram/metro. The percentages for each of these mode/purpose segments are:

	Car	Train	BTM	All surface modes
Commute	17.81%	10.14%	5.33%	33.28%
Business	2.97%	1.29%	0.15%	4.41%
Other	36.50%	15.77%	10.05%	62.32%
All purposes	57.28%	27.20%	15.53%	100.00%

The VOT for "all purposes" for air travel is calculated by weighting the VOTs for each purpose with the total minutes travelled in our own survey. The weight factors for business and other travel are 12.3% and 87.7% respectively.

In order to apply the valuation of reliability in CBA, one should multiply the RR (Table 60) by the corresponding VOT in euro per hour. For commuting and other travel this can be done quite straightforwardly. For business travel we have the situation that the RRs come from the individual travellers, but we have a VOT that consists of an employer and an employee component (through both were derived from interviewing the traveller). There is no information for calculating a separate employer component in the business VOR. We think it is best to assume that the business RR applies to the sum of the employer and employee component, i.e. to the total business VOT. All VORs are displayed in Table 61.

Table 60: Summary of reliability ratios as found in this study

	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Commute	0.4	0.4	0.4	0.4		
Business	1.1	1.1	1.1	1.1	0.7	
Other	0.6	0.6	0.6	0.6	0.7	0
All purposes	0.6	0.6	0.6	0.6	0.7	0

- All values are rounded off to the nearest multiple of 0.1.

Table 61: VORs (in 2010 euros per hour) as found in this study

	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Commute	3.75	4.75	3.25	4.00		
Business employee	*14.50*	*18.00*	*12.00*	*15.50*	*56.00*	
Business employer	*15.50*	*4.75*	*9.75*	*12.25*	-	
Business	30.00	22.75	21.75	27.75	56.00	
Other	4.75	4.50	3.75	4.50	30.75	0
All purposes	5.75	5.50	3.75	5.25	33.75	0

Notes:

- All values are rounded off to the nearest multiple of € 0.25.
- These values include VAT.
- VORs for "all purposes" are calculated by weighting the VORs for each purpose with the total minutes travelled in OViN 2010 (surface modes) or our own survey (air). The VORs for "all surface modes" are calculated similarly from the VORs for car, train and bus/tram/metro. See Table 59 for the weight factors.

As was done in the 1997 study, we also publish the final VOT per income category (see Table 62). These were calculated from the sample enumeration with exclusion of all those respondents who did not know their income, or were not willing to provide income information.

Table 62: VOTs (in 2010 euros per hour) after expansion to OViN and weighted with travel time as found in this study

Monthly household income after taxes	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Less than € 1875	8.75	9.50	6.00	8.25	29.25	8.25
€ 1875 to € 3125	9.50	11.00	6.50	9.00	36.50	8.25
€ 3125 to € 4325	8.25	13.80	7.00	9.25	36.50	8.25
More than €4325	10.50	14.25	11.75	10.75	47.75	8.25

Notes:

- All values are rounded off to the nearest multiple of € 0.25

8.12 Comparison with the previous Dutch national VOT survey

8.12.1 Comparison with current CBA values

In Table 63, our recommended values, based on the latent class models and expansion using OViN 2010, also including the employer component for business, are compared with the current values.

Table 63:Values of time (in euro of 2010 per hour) for car driver, train and BTM (bus/tram/metro) from various sources

	Current CBA value for 2010	New value for The Netherlands
Commuting – car driver	9.55	9.25
Commuting – train	9.62	11.50
Commuting - BTM	8.93	7.75
Business – car driver	33.07	26.25
Business – train	20.36	19.75
Business – BTM	15.56	19.00
Other – car driver	6.59	7.50
Other – train	5.93	7.00
Other – BTM	5.65	6.00
Car – all purposes	10.67	9.00
Train – all purposes	7.58	9.25
BTM- all purposes	6.63	6.75

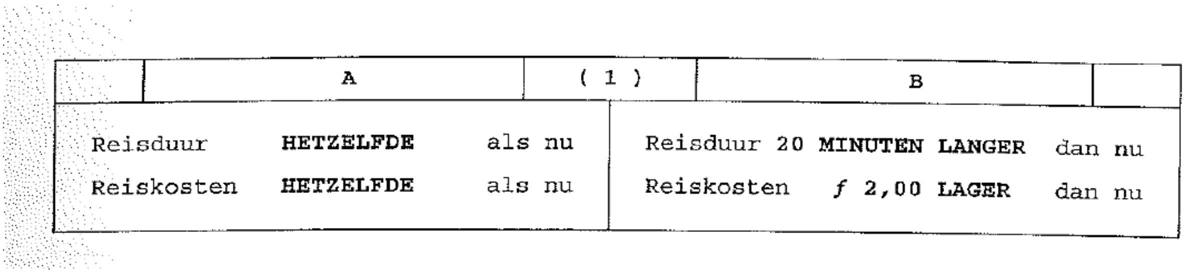
Note: Business values include employee and employer components.

The values of time for passenger transport that are used at the moment for CBA in The Netherlands (first column of Table 64) are based on SP research reported in Hague Consulting Group (1998). To get up-to-date values, the original outcomes from this SP survey were corrected for inflation using consumer price indices. Furthermore, the values were increased to account for real income growth, using an income elasticity of the value of time of 0.5. This elasticity is based on comparing outcomes from several previous Dutch value of time studies carried (Gunn, 2001) and is also consistent with the meta-analysis that Wardman (2001) carried out in the UK about ten years ago. However, later meta-analysis in the UK (with extended data sets) recently obtained an income elasticity of the value of time of 0.9 (Abrantes and Wardman, 2011). An even more recent meta-analysis on a combination of UK and international data (Wardman et al., 2012) found income elasticities varying between 0.68 and 0.85. These new findings are consistent with those of Börjesson et al. (2012b) for Sweden. So there is the possibility that the current values based on the 1997 survey should be higher, because the income elasticity used was not high enough.

8.12.2 Re-analysis of 1997 data using new methodology

In order to get a good picture of the differences between the VOTs based on the 1997 and on the new 2009/2011 data, we re-analysed the 1997 data (the details of this are in Appendix B) using the current methodology.

In 1997, a national stated preference study was conducted in the Netherlands to determine the VOT in passenger transport. The SP experiment was a simple binary time/cost experiment (Figure 14) very similar to experiment 1 of the 2009/2011 survey, however, with a different underlying statistical design.



	A	(1)	B	
Reisduur	HETZELFDE	als nu	Reisduur 20 **MINUTEN LANGER**	dan nu
Reiskosten	**HETZELFDE**	als nu	Reiskosten *ƒ* **2,00 LAGER**	dan nu

Figure 14: Example of an SP choice in the 1997 survey

An MNL utility model in preference space was estimated, equivalent to utility function [1] in Chapter 2, but with socio-economic interaction terms added to account for some observed heterogeneity in the cost and time coefficients and therefore also in the VOT. Expansion factors for each socio-economic segment were determined using the 1995 Dutch national travel survey. Finally, a weighted average VOT was determined (Hague Consulting Group, 1998), which was subsequently used in Cost-Benefit Analyses for major Dutch infrastructure project.

The influence of changes in methodology

The analysis of the 2009/2011 data differs from that of the analysis of the original 1997 data in seven ways:

1. The expansion procedure in 1997 differs from the procedure in the 2009/2011 analysis. In the current study we use a sample enumeration procedure in which each respondent gets a weight factor. These weight factors are determined in an iterative proportional fitting procedure such that the weighted sample population distribution matches all target distributions from OViN 2010. In 1997, a VOT was calculated for each combination of the socio-economic variable levels and the OVG was used to determine a weight factor for each of these cells.
2. In the 2009/2011 analysis, the simple MNL model was estimated in WTP space rather than in utility space.
3. The set of socio-economic factors used in the 2009/2011 study is not exactly the same as in the 1997 study (e.g. education was added in the new study).
4. In the 2009/2011 analysis, advanced MNL models were used that included non-linear terms (e.g. for the effect of base time and cost).
5. The 1997 results were expanded on the basis of OVG 1995, whereas the 2009/2011 results were expanded on the basis of OViN 2010.
6. The 1997 VOT did not depend on the travel time distribution and consequently in the expansion of that study, the travel time distribution of OVG was not used as a target that needed to be represented. However, the 2009/2011 VOT depends

on travel time and therefore the travel time distribution was added as a target in the expansion procedure using OViN 2010.

7. No panel-Latent Class models were estimated in 1997-1998.

We tested for the influence of each of these differences on the VOTs in both data sets.. The main outcomes are as follows (the details are in Appendix B):

1. The new expansion procedure clearly produces a different result by mode, however on average the VOT increases by only 4%.
2. Estimating models in WTP space rather than in utility space makes no difference. However, the 1997 survey had socio-economic interaction terms on the time and cost coefficient, whereas the 2009/2011 survey used socio-economic interaction terms directly on the VOT. This does cause some differences, but these are small. On average the VOT decreases by 4%.
3. The use of a different set of socio-economic factors has a very small effect. On average the VOT increases by 1%.
4. The use of a different base distribution for weighting (based on the 1995 or the 2010 travel survey) has some impact on the VOTs per mode, but on average the VOTs remain the same.
5. The use of the advanced MNLs has some impact, which can go either way relative to simple MNL. On average the VOTs remain the same.
6. Expanding the sample to the travel time distribution as well also has some impact by mode, but on average the VOT decreases by only 2%.
7. The use of the panel-Latent Class models increases the VOT substantially relative to advanced MNL models, for all purposes, both for the 1997 and the 2009/2011 data. On average the VOT increases by +31%.

In other words, had we in 1997-1998 applied a panel-Latent Class model (which was not available in its current form at that time) on the 1997 data, we would have obtained considerably higher VOTs. The shift from MNL models to LC models between the old and the new study leads to a substantial increase in the VOTs.

Comparison of 2009/2011 VOTs with 1997 VOTs

In order to make a fair comparison between the VOTs in 1997 and in 2010, we should calculate the VOTs for both years using the same method. As preferred method for this, we use the panel-Latent Class model, with socio-economic factors (but without education) and non-linear terms, expanded using the new methods and OViN 2010. The changes in the VOTs between 1997 and 2010 that we then observe are compared with our expectations about changes in the VOT over time.

Expected changes in the VOT

We expect the VOT between 1997 and 2010 to increase because of the change in price level over this period. Also, we expect a further increase as a result of real income changes (income increase over and above the price change). Between 1997 and 2010 consumer prices rose by 31.5%. Real income has increased by about 30% over the same period. So on the basis of price change up to 2010, one could expect the 1997 values to increase by 32%, and there would be a further increase (of about 15% using the currently employed income elasticity of 0.5) as a result of the real income increase over and above the price increase, giving about a 47% increase in the VOT.

On the other hand, VOTs could have been going down in the period 1997-2010 because it has become much easier to use travel time in all of the modes in a more productive and/or enjoyable way, since new technologies such as mobile phones (also hands free), laptops, iPads and smartphones with mobile internet have been introduced or become much more common in The Netherlands over this period.

In analysing the 1988 and the 1997 VOT data collected in The Netherlands, Gunn (2001) had already observed the same phenomenon. The VOTs did not change much in the period 1988-1997 in real terms. In this period real income substantially increased (e.g. wages rising more than prices), but this had not led to large increases in the VOT. Gunn explained this stability by hypothesising that the impact of real income growth was more or less balanced by the (technological) developments that allow travellers to make a better use of travel time (mobile phones, laptops).

Observed changes in the VOT

We now see that the VOT (over all modes) based on the 2009/2011 data (in euros of 2010) is 23% higher for commuting, 26% higher for business and 65% higher for other purposes relative to the VOT based on the same panel-Latent Class model estimated on the 1997 data (in euros of price level 1997). We therefore conclude that for commuting and business, the impact of the real income increase and even a part of the price increase on the VOT have been compensated by one or or more factors, such as technological innovations that affect the use of travel time. For other travel the increase in the VOT exceeds the sum of the price and half the real income increase (but not the full sum of price and real income change). It seems likely that the changes in travel-relevant in technology since 1997 have been less profound for other trips.

8.13 Comparison with the existing VOT literature

A comparison of the new recommended VOTs and the international literature can be found in the table below.

Table 64: Values of time (in euro of 2010 per hour) for car driver, train and BTM (bus/tram/metro) from various sources

	Value for NL from inter-national meta-study (Shires and de Jong, 2009)	Norway (Ramjerdi et al. 2010)	Sweden (Börjesson and Eliasson, 2011)	New value for The Netherlands
Commuting – car driver	11.05	12.13-26.95	9.2-12.1	9.25
Commuting – train	11.05		7.2	11.50
Commuting - BTM	9.14		5.3	7.75
Business – car driver	30.94	51.20		26.25
Business – train	30.94			19.75
Business – BTM	24.83			19.00
Other – car driver	8.85	10.37-19.67	5.9-7.8	7.50
Other – train	8.85		5.0	7.00
Other – BTM	6.21		2.8	6.00
Car – all purposes	-			9.25
Train – all purposes	-			9.50
BTM- all purposes	-			7.00

Business values include employee and employer components.

The first column of Table 64 gives the outcomes of an application of the estimation results of the international meta-analysis of Shires and de Jong (2009) to The Netherlands (e.g. using the Dutch GDP per capita, etc.). Here we adjusted for price changes since 2003 (the year that Shires and de Jong used to express their VOTs), but no correction for real income growth on top of that was applied. Income change-compensated values from the meta analysis would be slightly higher than the values in this column.

The most recent national VOT studies are those of Sweden (Börjesson en Eliasson, 2011) and Norway (Ramjerdi et al., 2010). The study teams in these countries estimated non-parametric models, which account for the sign and size of the travel time changes offered, observed heterogeneity and unobserved heterogeneity between respondents. In Table 64 for Sweden and Norway we sometimes give two values per cell: the lower values for Norway refer to short distances, the higher values refer to long distances (>100 km). For Sweden the lower values are valid outside Stockholm and the higher values are for Stockholm. All Swedish VOTs in this table refer to short distances; for longer distances, the Swedish VOTs are higher than presented (maximally 14.9 euro per hour), but for these there is no distinction between travel purposes. In Norway, train and BTM are not separate categories, whereas the Swedish study did not include business travel.

We see in Table 64 that the recommended new values for commuting and other purposes provide a good match with the international literature (represented by the meta-analysis and the most recent national VOT studies). The meta-analysis is mainly based on studies that use MNL models, and we have found in our study that these have a downward bias due to ignoring unobserved heterogeneity and panel effects. Recent studies that also account for unobserved heterogeneity (Sweden and Norway) produce values which are often higher than the older literature, and our new values for commuting and other are within the bandwidths provided by the Norwegian and Swedish studies.

For business travel our new values are somewhat lower than those of the meta-analysis. With the exception of car driver, the current CBA values for business travel are also smaller than those from the meta-analysis. The difference between the current CBA values and the new Dutch values on the one hand and those from the meta-analysis on the other hand may be caused by the latter being mainly based on countries that use the wage costs for the business VOT. The Dutch studies (1988-1990, 1997-1998 and this one) all accounted for the fact that not all saved time on business trips is used for the employer and that travel time is not necessarily unproductive; this reduces the business VOT.

The previous VOT study did not include air transport or recreational navigation. Below we compare the new value for air transport with the values in the Netscan model (SEO, 2011)[30], the current CBA values (which are based on those of SEO), and the meta-analysis. The new values for business travel and for other travel are about twice as high as the value in the Netscan model. However, the Netscan value refers to international travel by airplane, and our new value is for Dutch air travellers. The VOTs from MNL models (52 euro for business and 29 for other) were much closer to the NETSCAN values, but accounting for heterogeneity and panel effects in the panel-LC model increased the VOTs considerably. Apparently there is a lot of heterogeneity between air travellers in their VOT, both for business and other travel.

Table 65: Value of time (in euro of 2010 per hour) for air transport from various sources

	SEO (2011): values for international air travel	Current CBA value (in Euro of 2010)	Value for NL from international meta-study (Shires and de Jong, 2009)	New values for The Netherlands
Business – airplane	52[1]	47	43	85,75
Other – airplane	24[1]	20	-	47
All purposes - airplane		30		51,75

Notes:
- Business values include employee and employer components.
- [1] original values: 65 and 30 US dollars; exchange rate used: 1 US $ = 0.8 euro

---

[30] In 2010, SEO Economisch Onderzoek carried an analysis into the VoTs of air travellers. This was an analysis within the framework of the validation of two important parameters in the NetScan-model of SEO: the value of time (VoT) and the so-called “spread-coefficient”, that links generalised costs to utility. Both parameters have separate values for the business and the non-business segment. For the sake of the validation, a limited set of MIDT-data was used (trips from and via Schiphol), that among other things provide a specification for the total market volume, the chosen routes and the price levels used. The NetScan-model also generates this kind of information, with the parameters values mentioned as inputs. An investigation was carried out, using “trial and error”, to find out at which parameter values the NetScan-output best matched the MIDT-information. Those parameter values were then chosen as the validated NetScan-parameters. This gave a value of time for business and non-business of USD 65 and USD 30 (the MIDT-price information is also in US-dollars).

The latest and biggest meta-study, Wardman et al. (2012), comes to more than 50% higher than Shires and de Jong (2009) for the business VOT in air transport for a 500 km trip: about 74 euro per hour (and about 28 euro per hour for leisure). These are results of applications of meta-regressions estimated on the international VOT literature to The Netherlands (e.g. using the GDP per capita for The Netherlands).

Literature on values for recreational navigation is not available.

8.14 Comparison to the existing VOR literature

In Table 66 is a comparison of the values that we obtained for the reliability ratio (RR) against the empirical literature. There have been a few more studies on the value of reliability that provided numerical outcomes than listed here, but these provide metrics other than the RR, and are therefore not comparable. We also included in Table 66 the outcomes of the 2004 expert workshop which provided provisional values for the RR for use in CBA in The Netherlands, even though these are not empirical findings but expert judgments (see de Jong et al., 2004). In a number of CBAs in the Netherlands the reliability benefits have been calculated as 25% of the travel time benefits (based on Besseling et al. (2004)). From Table 66 we conclude that the new RRs that we obtained fit quite well within the range of values provided by the international literature. All values we now get (except the one for business for car) are lower than the provisional values from the expert workshop of 2004, but many recent empirical values are also lower than the workshop values. For air transport, we found just one other study that provided an RR (Norway), and that value is clearly lower than our value (which is more comparable to the RRs for other modes).

In order to apply the RRs in CBA, one should multiply the RR by the corresponding VOT in euros per hour. For commuting and other travel this can be done quite straightforwardly. For business travel we have the situation that the RRs come from the individual travellers, but that we have a VOT that consists of an employer and an employee component (through both were derived from interviewing the traveller). There is no information for a calculating a separate employer component in the business VOR. We think it is best to assume that the business RR applies to the sum of the employer and employee component that is to the total business VOT.

Table 66: Comparison against the empirical literature on the reliability ratio (for the value of the standard deviation of travel time versus travel time)

Study	Country	RR
Car		
MVA (1996)	UK	0.36 – 0.78
Copley et al. (2002)	UK	Pilot survey: 1.3
Hensher (2007)	Australia	0.3 – 0.4
Eliasson (2004)	Sweden	0.30 – 0.95
Mahmassani (2011)	USA	NCHRP 431: 0.80 – 1.10 SHRP 2 CO4: 0.40 – 0.90
Expert workshop of 2004	*The Netherlands*	*0.8*
This study	The Netherlands	Commuting: 0.4 Business: 1.1 Other: 0.6
Train		
ATOC (2002)	UK	is 0.6 – 1.5
Ramjerdi et al. (2010)	Norway	Short trips: 0.69 Long trips: 0.54
Expert workshop of 2004	*The Netherlands*	*1.4*
This study	The Netherlands	Commuting: 0.4 Business: 1.1 Other: 0.6
Bus/tram/metro		
MVA (2000)	France	0.24
Ramjerdi et al. (2010)	Norway	Short trips: 0.69 Long trips: 0.42
Expert workshop of 2004	*The Netherlands*	*1.4*
This study	The Netherlands	Commuting: 0.4 Business: 1.1 Other: 0.6
Air		
Ramjerdi et al. (2010)	Norway	0.20
This study	The Netherlands	Business: 0.7 Other: 0.7

CHAPTER 9 Further use of SP data in transport forecasting models

This section describes how the stated preference (SP) and revealed preference (RP) data gathered for freight and passenger transport in the project 'The value of travel time and travel time reliability' can further be used in other projects, more specifically in projects to include reliability in the national/regional freight transport models in The Netherlands.

9.1 The P-side and Q-side of reliability

The project 'The value of travel time and travel time reliability', carried out by Significance, VU University, John Bates, TNO, NEA, TNS NIPO and PanelClix for the Netherlands Ministry of Infrastructure and the Environment, was related to the P-side ('price') of the reliability of transport time: the research question was to provide monetary values of time and travel time reliability (variability) for passenger and freight transport by mode that can be used in cost-benefit analysis (CBA) of transport projects. But in order to carry out a CBA of a proposed transport infrastructure project or a transport policy, one not only needs monetary values, but also numerical estimates of what happens on the Q-side.

The Q-side ('quantity') of reliability of transport times relates to two different issues:

1. Predicting the impact of infrastructure projects and transport policies on reliability, so that the changes as a result of the project between the amount of unreliability in the reference situation and the policy situation can be quantified.
2. Predicting the reactions of the travellers, the shippers and the carriers to changes in reliability, so that the number of travellers and freight trips and the transport distances (by mode, segment, etc.) can be predicted for the reference and the policy situation.

For the first issue on the Q-side, the new SP (and RP) data on the choices of travellers, shippers and carriers is not useful. For the second issue, this information is potentially relevant. This will be further explained in this section.

9.2 The current transport models

For CBA of transport infrastructure and policy projects, the models regularly used are the LMS and NRM for passenger transport (for freight the LMS also gives the assignment of trucks to the road network).

The modules of the LMS contain: tour generation, mode/destination/time-of-day choice and network assignment (for road transport). This time-of-day choice model is not a full scheduling model: it does not contain deviations from the preferred arrival time (PAT) as explanatory variables for the departure time choice, but only travel time and cost for each time period of the day. The travel modes in the LMS are: car driver, car passenger, train, bus/tram/metro, cycling and walking.

For freight transport, a new national model, called BASGOED has recently been developed (Significance et al., 2010). This model includes the Economy Module from the earlier SMILE+ model to give transport production and attraction. New submodels for distribution and mode choice (road, inland waterways, rail) were estimated on aggregate data from the Basisbestand Goederenvervoer. The assignment takes place in existing unimodal models (LMS for roads, jointly assigning trucks and cars to the same network, BIVAS for inland waterways and ROUTGOED for rail). BASGOED is a relatively simple transport model. Its objective was to get an operational model in a relatively short time period. Development of more sophisticated freight transport models is planned for the coming years.

9.3 The use of SP experiments in freight and passenger transport forecasting models and in Cost-Benefit Analysis

For changes in travel time/transport time, the above models can give the reactions of passenger and freight transport. An important research question for the future is how the reactions to changes in reliability can be introduced in these models as well.

Data on the reactions of travellers, shippers and carriers to reliability is hard to get. A particular problem is getting enough variation in reliability to be able to estimate model coefficients. SP data then might be very helpful, because in SP the researcher can control the variation in reliability. For instance the departure time choice mode in the LMS was estimated on SP data. When using SP data in the estimation of a forecasting model, it is important to remember that it is bad practice to do forecasting on the basis of SP models alone. The variation in the unobserved component (that is related to the scale of the model) of the utility function differs between an SP context (where everything else is said to stay the same) and a real world context. For ratios of coefficient values (such as value of time and reliability), this is not an issue – since the influence of the unobserved variation cancels out –, but for model predictions (including elasticities) it is. To solve this problem, models are estimated simultaneously on SP and RP data (scaling the SP variance to the RP one) or SP models are further calibrated to the observed shares of the choice alternatives.

Passenger transport models

For passenger transport we see possibilities to do a joint estimation of the mode and destination choice models (possibly also the departure time choice models) on data sets from OVG/MON/OViN (national travel survey: RP) and the SP data collected in this project, to derive a model for these choices that will also have an impact of reliability.

Freight transport models

For freight transport, the possibilities for combining the new SP data with data on observed choices in a joint estimation are more limited. The current BASGOED model is not based on choice data for individual shippers and carriers (disaggregate data), but on aggregate data (zone-to-zone, by commodity type). Simultaneous estimation on aggregate and disaggregate data is statistically not infeasible, but hardly ever pursued, because the data sources are too different to expect joint coefficients. This would be different if disaggregate RP data was/were available for freight transport in The Netherlands, but no such surveys are planned as far as we know.

The SP models for freight can also not be used as a mode or distribution choice model on its own within a larger freight model system, because the choices in the SP are all within-mode choices. Possibly the variation between respondents would be sufficient for an SP mode choice model, but the fact that an RP model for mode choice could not be estimated successfully on the data in this project makes this not very likely. Using the theoretical results from Fosgerau and Karlström (2010) it might be possible to translate[31] a mean-variance model into a scheduling model (or to estimate this directly on the SP data) that could be used for departure time choice in freight transport.

The outcomes of the SP estimations for freight can be used to check the values of time that are implied by BASGOED (given the costs functions in BASGOED, especially the shippers VOTs for the cargo-related component will be especially relevant here). To get an approximation of the effect of a change in reliability on the shippers and carriers, the reliability ratios (value of reliability to value of time) from the SP research can be used. Using these, a change in unreliability (measured as the standard deviation) can be translated into a change in transport time. This change in transport time can then be inserted in the BASGOED model (independently or on top of the transport time change that results from the measure studied). BASGOED, possibly after a recalibration of the alternative-specific constants, will then give the freight transport reactions (tonnes, tonnes-km, by mode for reference and policy situation). Alternatively one may use the monetary values obtained for reliability changes in previous chapters and add these to the transport cost changes in BASGOED. The idea is that transport time and cost are already in the model and that our study has provided translation factors of reliability into these variables. In this way, reliability changes can be added to time and/or cost changes, and the model can then be run as it is.

Practical application of SP results in CBA for freight transport

The monetary transport time gain in a certain year in the future consists of a P (price: the VOT) times a Q (quantity). If the transport volumes increase over time, Q will increase each year (based on interpolation, since the transport models are only run for a very limited set of years).

---

[31] This translation is only possible if the ratio between ValEarly (value of arriving early) and ValLate (value of arriving late) and the travel time distribution is known.

The P-part itself will consist of the trade-off ratio TR (see Chapter 5) times the factor cost (per hour). If the transport costs increase with time, the factor cost will increase each year.

So, in a CBA, both Q and the factor cost need to be calculated for each year after the introduction of the new infrastructure that is evaluated. Using a different TR for each of those years then poses no extra complication.

It is sensible to assume that 10 years after the introduction of the project the maximal TR will be attained: all reactions of the freight sector have then been implemented (the remaining lifetime of transport equipment only seldom exceeds 10 years).

So, in practical CBA, we recommend using the minimal TR for year 1 and the maximal TR for year 10, and a linear interpolation in between.

The question then is which values should be used for the minimal and maximal TR. For the minimal TR we recommend the value from the SP. The maximum could in principle even exceed the factor costs per hour (because of the VOT component from the cargo itself), but this additional component usually is quite small, whereas it may also be difficult to reduce the transport costs all the way to zero to benefit from transport time gains, even in the long run. We recommend using TR=1 as the maximum for the VOT.

CHAPTER 10 Summary and conclusions

Aim and scope

This project aimed at providing values of time and of travel time reliability (variability) for use in cost-benefit analysis (CBA) of transport projects in The Netherlands. Values were sought for passenger transport by car, bus/tram/metro, train, airplane and ship (recreational navigation) and for freight transport by road, rail, air, inland waterways and sea transport.

The current values of time are based on surveys carried out in 1997 (passengers) and 2003/2004 (freight). There are also existing values of reliability for use in CBA, but these are based on expert opinions (passenger transport) or to a large degree based on many untestable assumptions (freight transport).

This was the first national study in The Netherlands that empirically investigated all of these topics in a joint framework. Previous national value of time studies for passenger transport did not include reliability and previous national freight value of time studies included a measure of reliability (the fraction of shipments that was delivered on time) that is not compatible with the definitions of reliability that are favoured for use in valuation studies for passengers and in developing prediction models for reliability.

As the operational measure for unreliability we use the standard deviation of travel time. The ratio of the value of the standard deviation to the value of travel time is called the reliability ratio. The main reason for choosing this definition was that all other possible measures of reliability would be much harder to incorporate in the national and regional transport models.

The SP surveys

In a previous project, questionnaires had been designed for interviewing travellers, shippers and carriers in The Netherlands. The focus in these interviews is on stated preference (SP) experiments, where the respondents are presented with hypothetical alternatives for a trip or transport that they had actually made. The hypothetical alternatives are described in terms of travel time, travel costs and reliability. Reliability is not presented to the respondents in the form of the standard deviation, because such concepts are not generally understood, but in the form of five possible travel times which are equally likely to happen.

For interviewing firms transporting goods by inland waterways or sea, we used an innovative choice context, namely that of waiting for a lock or bridge or to be loaded or unloaded at a quay, instead of the traditional transport time versus transport cost choice experiment.

For freight transport, 812 interviews were successfully carried out with shippers and carriers, using computer-assisted face-to-face interviewing. Shippers that contract the transport out are asked only to include the influences on the commodities themselves in their choice-making. Carriers are asked only to include the influences on their vehicles and their staff. This leads to values of time (and reliability) of shippers and carriers that can be added to obtain a total value.

For passenger transport, we first collected 5,760 interviews in 2009 using an internet panel. Initial models estimated on these data showed values of time that differed very much from the values that were derived on the basis of the 1988 and 1997 value of time surveys. These differences could not convincingly be explained by differences in the socio-economic composition of the sample, attributes of the trips or differences in the design of the SP experiments. It was therefore decided to do additional data collection (which took place in 2011) using the same method of recruiting respondents as had been used in 1988 and 1997, asking travellers at petrol stations/service areas, parking garages, railway stations, bus stops, airports and ports to participate in the survey. However, the survey itself was delivered using a weblink to an internet questionnaire in 2011, as opposed to the paper-based version sent by mail in 1988 and 1997. The 2011 survey includes 1430 interviews. Models estimated on the 2011 data or the combination of the 2009 and 2011 data yield values of time that are compatible with those of 1988 and 1997.

The use of internet panels

We conclude that the 2009 SP passenger survey using members of an internet panel leads to substantially lower VOTs than the 2011 SP passenger survey (with recruitment of respondents en-route). The most likely explanation is that the 2011 values are correct and that the 2009 values are biased downwards, mainly because persons with a lower value of time (in every socio-economic segment) have a higher probability of becoming a member of an internet panel (and among those panel members, the ones that are most likely to participate in this relatively long survey are the ones with even lower values of time). The shorter distances of the trips sampled in 2009 (and the corresponding smaller time savings offered in the SP) also played a role, but even correcting for this, there is clear evidence for a downward bias in the VOTs in the 2009 data. The reference values of time (and reliability) in our estimates are therefore based on 2011 only.

The analysis of the data and the results for freight transport

Discrete choice models were estimated on the SP data. The ratio of the time, or the reliability, coefficient to the cost coefficient of these models represents the trade-off between time and money, or between reliability and money. For the non-road models we use relative models, in which the attributes are measured relative to the observed levels. To obtain absolute money values of time and reliability from these models, data on the transport costs per hour (the so-called 'factor costs') are also required. These values were made available by DVS and used in this report in combination with the new estimates. This resulted in the following results for the freight transport value of time by mode.

Table Y1: Values of time for freight transport (Euro/hour per vehicle or vessel, price level 2010)

	Road	Rail	Air	Inland waterways	Sea
Container	*[2-40t truck]:* 59	*[full train]:* 880	*Not applicable*	*[ship waiting for a quay]:* 98 *[ship waiting for a lock/bridge]:* 340	*[ship waiting for a quay]:* 760
Non-container	*[2-15t truck]:* 23 *[15-40t truck]:* 44 *[all non-container]:* 37	*[bulk]:* 1200 *[wagonload train]:* 1100 *[all non-container]:* 1200	*[full freighter aircraft]:* 13000	*[ship waiting for a quay]:* 65 *[ship waiting for a lock/bridge]:* 300	*[ship waiting for a quay]:* 830
All	*[2-40t truck]:* 38	*[full train]:* 1100	*[full freighter aircraft]:* 13000	*[ship waiting for a quay]:* 69 *[ship waiting for a lock/bridge]:* 300	*[ship waiting for a quay]:* 820

Notes:

- All these values are combined values from shippers and carriers and were obtained after rounding off.
- The values for rail are for a train (not a wagon).
- The values for inland waterways and sea refer to a ship.
- These values do not include VAT.

The new values for road and rail are of the same order of magnitude as the (inflation-corrected) values from 2003/2004, and are also compatible with the international literature. For international waterway transport and sea transport we now obtain higher and more plausible values per hour than in 2003/2004. In Table Y2 we give the corresponding results for the freight transport VOR.

Table Y2: Values of reliability for freight transport (Euro/hour per vehicle or vessel, price level 2010)

	Road	Rail	Air	Inland waterways	Sea
Container	*[2-40t truck]:* 4	*[full train]:* 100	*Not applicable*	*[ship waiting for a quay]:* 18 *[ship waiting for a lock/bridge]:* 27	*[ship waiting for a quay]:* 45
Non-container	*[2-15t truck]:* 34 *[15-40t truck]:* 6 *[all non-container]:* 15	*[bulk]:* 260 *[wagonload train]:* 240 *[all non-container]:* 250	*[full freighter aircraft]:* 1600	*[ship waiting for a quay]:* 25 *[ship waiting for a lock/bridge]:* 25	*[ship waiting for a quay]:* 110
All	*[2-40t truck]:* 14	*[full train]:* 200	*[full freighter aircraft]:* 1600	*[ship waiting for a quay]:* 24 *[ship waiting for a lock/bridge]:* 26	*[ship waiting for a quay]:* 100

Notes:
- All these values are combined values from shippers and carriers and were obtained after rounding off.
- The values for rail are for a train (not a wagon).
- The values for inland waterways and sea refer to a ship.
- These values do not include VAT.

For all segments in Table Y2, except for road transport/non-container/2-15 tonnes (where the reliability ratio is 1.46) the carrier VOR was not significantly different from zero, and the resulting mean VOR only consists of the valuation of the shipper, which results in a mean VOR that is only 8-22% of the mean VOT for the corresponding segment (so the reliability ratio here is 0.08 – 0.22). For road transport other than non-container/2-15 tonnes and for air transport one might have expected higher VOTs (on the basis of the available literature, however limited and because of the association between road and air transport with high value commodities). For these segments we recommend, in studies using our estimated values, doing a sensitivity analysis also using a higher VOR.

The analysis of the data and the results for passenger transport

We estimated discrete choice models in which the values of time differ between trips with different time and costs levels, different time and costs changes offered in the SP, and different characteristics of the respondents (e.g. education, income, age, household composition). By using a panel latent class model, we also account for unobserved differences between respondents in the value of time and for repeated measurements/panel effects. The reference values of time and the reference reliability ratios were estimated on the 2011 sample only, but the effect of time and cost level, time and cost changes offered and socio-economic attributes was estimated on both the 2009 and 2011 samples.

These are absolute models in willingness-to-pay space. When including the dependencies of the VOT and the VOR on the observed levels of time and cost and on the magnitude of the changes in the attributes offered in the SP, models in willingness-to-pay space perform better than models in log willingness- to-pay space, and are therefore preferred.

The recommended values of time were calculated by weighting the sampled respondents to represent the distribution of time travelled in the trips recorded in the national travel survey OViN. The resulting VOTs are in Table Y3.

Table Y3: Values of time for passenger transport (Euro/hour, price level 2010)

	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Commute	9.25	11.50	7.75	9.75		
Business employee	*12.75*	*15.50*	*10.50*	*13.50*	*85.75*	
Business employer	*13.50*	*4.25*	*8.50*	*10.50*	*-*	
Business	26.25	19.75	19.00	24.00	85.75	
Other	7.50	7.00	6.00	7.00	47.00	8.25
All purposes	9.00	9.25	6.75	8.75	51.75	8.25

Notes:
- All values are rounded off to the nearest multiple of € 0.25
- These values include VAT.

Most of the new VOTs are not very different from the current official Dutch values, and are within the range of the recent international literature, especially that for comparable models.

For the value of reliability, we estimated the reliability ratio for each modelling segment (see Table Y4). The reliability ratio gives the monetary value of reliability (measured as standard deviation of transport time) divided by the value of time. To obtain a monetary value of reliability, the reliability ratio needs to be multiplied by the value of time. The outcomes for the VOR are given in Table Y5.

Table Y4: Reliability ratios for passenger transport

	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Commute	0.4	0.4	0.4	0.4		
Business	1.1	1.1	1.1	1.1	0.7	
Other	0.6	0.6	0.6	0.6	0.7	0

Notes:
- All values are rounded off to the nearest multiple of 0.1

Table Y5: Values of reliability for passenger transport (Euro/hour, price level 2010)

	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Commute	3.75	4.75	3.25	4.00		
Business employee	*14.50*	*18.00*	*12.00*	*15.50*	*56.00*	
Business employer	*15.50*	*4.75*	*9.75*	*12.25*	-	
Business	30.00	22.75	21.75	27.75	56.00	
Other	4.75	4.50	3.75	4.50	30.75	0
All purposes	5.75	5.50	3.75	5.25	33.75	0

Notes:
- All values are rounded off to the nearest multiple of € 0.25
- These values include VAT.

The VORs are generally somewhat lower than the expert judgement values that we had before, but correspond reasonably well to values obtained recently in the international literature.

Different model types for passenger and freight transport

For passenger transport we have estimated models with non-linear terms to account for the dependence of the VOT on the observed time and cost levels, and on the presented time and costs changes in the SP. Interaction terms represent the influence of socio-economic variables (observed heterogeneity). Furthermore, the panel latent class specification that we used accounts for unobserved heterogeneity and panel (repeated measurements) effects.

Even though we have collected one of the biggest ever SP samples in freight transport, we have far fewer observations in freight transport. Models with non-linear terms, such as those for passenger transport and panel latent class models or other mixed logit models, have been tried for freight, but did not lead to stable and significant estimates. On the other hand, a VOT that clearly increases with distance travelled has been found in many passenger transport studies, but not in freight.

It is generally accepted that freight transport is very heterogeneous, probably more so than passenger transport. Especially for the non-road transport modes, vehicles and vessels with a large variation in capacity are used. The way we accounted for heterogeneity in freight transport was by using a relative model: all SP attributes are measured relative to their observed levels, so that we estimate trade-offs between percentage changes, not absolute amounts of time and money. To calculate VOTs and VORs one then needs additional external inputs in the form of the factor cost: the transport cost per unit of time. To correct for repeated measurements in the freight SP, we used the Jack-knife method.

The relative model was used for all freight modes in the previous freight VOT study of 2003/2004. In the new study this was repeated, but for road transport there was a change of method: we now use an absolute model in logWTP space, that directly (without using an exogenous factor cost) yields a VOT. For road transport there probably is less heterogeneity in terms of the vehicle capacity used.

REFERENCES

Reference List

Abrantes, P.A.L., and M. Wardman (2011) Meta-analysis of UK values of time: an update, *Transportation Research A*, 45 (1), 1-17.

ALOGIT Software and Analysis Ltd. (2007) ALOGIT 4.2EC., http://www.alogit.com/

Association of Train Operating Companies (ATOC) (2002) Passenger Demand Forecasting Handbook, London: ATOC.

Avineri, E. & Bovy, P. (2008) Identification of parameters for prospect theory model for travel choice analysis.Transportation Research Record: Travel Behavior Analysis, 2082. pp. 141-147.

Bajari, P., Fox, J., Ryan, S., (2007) Linear regression estimation of discrete choice models with nonparametric distributions of random coefficients., *American Economic Review*, 97(2), 459-463.

Bastin, F., C. Cirillo, and Ph. L. Toint (2010) Estimating Nonparametric Random Utility Models with an Application to the Value of Time in Heterogeneous Populations. *Transportation Science* 44, 4 (November 2010), 537-549.

Bates, J., J. Polak, P. Jones and A. Cook (2001) The valuation of reliability for personal travel, *Transportation Research E (Logistics and Transportation Review)*, 37-2/3, 191-229.

Batley, R.P., S. Grant-Muller, J. Nellthorp, G.C. de Jong, D. Watling, J.J. Bates, S. Hess and J. Polak (2008) Multimodal Travel Time Variability, Final Report, report for the UK Department of Transport, ITS Leeds, John Bates and Imperial College.

Ben-Akiva, M.E. and Lerman, S.R. (1985) Discrete choice analysis: Theory and application to travel demand, MIT Press, Cambridge Massachusetts.

Besseling, P., de Groot, W. and Verrips, A. (2004) Economische toets op de Nota Mobiliteit, *CPB document 65*, CPB, The Hague.

Bhat, C., (1997) An endogenous segmentation mode choice model with an application to intercity travel, *Transportation Science* 31(1), 34-48.

Bierlaire, M., (2003) BIOGEME: A free package for the estimation of discrete choice models, *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland.

Bierlaire, M. (2008) An introduction to BIOGEME version 1.7, biogeme.epfl.ch.

Borger, B. de and M. Fosgerau (2008) The trade-off between money and travel time: A test of the theory of reference-dependent preferences, *Journal of Urban Economics*, 64, 101-115.

Börjesson, M. and J. Eliasson (2011) Experiences from the Swedish value of time study, CTS working paper, Centre for Transport Studies, Royal Institute of Technology, Stockholm.

Börjesson, M., J. Eliasson and J.P. Franklin (2011) Valuation of travel time variability in scheduling versus mean-variance models, Centre for Transport Studies, Royal Institute of Technology, Stockholm.

Börjesson, M., M. Fosgerau and S. Algers (2012a) On the income elasticity of the value of time, *Transportation Research A*, 46, 368-377.

Börjesson, M., M. Fosgerau and S. Algers (2012b) Catching the tail: Empirical identification of the distribution of the value of travel time, *Transportation Research A*, 46, 378-391.

Bruzelius, N. (2001) The valuation of logistics improvements in CBA of transport investments – a survey, SAMPLAN, SIKA, Sweden.

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

CBS (2012) CBS Statline – consumer price index, see http://statline.cbs.nl/statweb/

Chamberlain, G. (2010) Binary Response Models for Panel Data: Identification and Information. *Econometrica* 78 (1), 159-168.

Copley, G., Murphy, P. and Pearce, D. (2002) Understanding and valuing journey time variability; European Transport Conference – 2002, Cambridge.

Daly, A.J., F. Tsang and C. Rohr (2011) The value of small time savings for non-business travel, Paper presented at European Transport Conference 2011, Glasgow.

DeTar, C. (2002) http://www.physics.utah.edu/~detar/phycs6730/handouts/jackknife/jackknife/jackknife.html, written 30-03-2002, retrieved 31 July 2012

Eliasson, J. (2004) Car drivers' valuations of travel time variability, unexpected delays and queue driving, *Proceedings of the European Transport Conference*, 2004.

Feo-Valero, M., L. Garcia-Menendez and R. Garrido-Hidalgo (2011) Valuing freight transport time using transport demand modelling: a bibliographical review, *Transport Reviews*, 201, 1-27.

Fosgerau (2006) Investigating the distribution of the value of travel time savings, *Transportation Research Part B*, 40 (8), 688-707.

Fosgerau (2007) Using nonparametrics to specify a model to measure the value of time, *Transportation Research* A, 41 (2007), 842-856.

Fosgerau, M. and M. Bierlaire (2007) A practical test for the choice of mixing distribution in discrete choice models, *Transportation Research Part B: Methodological*, Volume 41, Issue 7, August 2007, 784-794

Fosgerau, M. and M. Bierlaire (2009) Discrete choice models with multiplicative error terms, *Transportation Research B*, 43 (2009), 494-505.

Fosgerau, M. and Karlström, A. (2010) The value of reliability, *Transportation Research B*, 44(1), 38-49.

Fosgerau, M. and Engelson, L. (2011) The value of travel time variance, *Transportation Research B*, 45(1), p.1-8.

Fowkes, A.S. (2006) The design and interpretation of freight stated preference experiments seeking to elicit behavioural valuations of journey attributes, ITS, University of Leeds.

Greene, W.H. and D.A. Hensher (2012) Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model, *Applied Economics*, 2012, 1-6

Gunn, H.F. (2001) Spatial and temporal transferability of relationships between travel demand, trip cost and travel time, *Transportation Research E*, Vol. 37, pp. 163-189.

Gunn, H.F. (2008) Valuation of travel time savings and losses, In D.A. Hensher and K.J. Button (Eds.): Handbook of Transport Modelling, Second Edition, Pergamon, Oxford.

Hague Consulting Group, Rotterdam Transport Centre and NIPO (1991) De reistijdwaardering in het goederenvervoer, eindrapport voorstudie, Rapport voor Rijkswaterstaat, Dienst Verkeerskunde, HCG, Den Haag.

Hague Consulting Group, Rotterdam Transport Centre and NIPO (1992) De reistijdwaardering in het goederenvervoer, rapport hoofdonderzoek, Rapport 142-1 voor Rijkswaterstaat, Dienst Verkeerskunde, HCG, Den Haag.

Hague Consulting Group (1998) The second Netherlands' value of time study: final report, Report 6098-1 for AVV, HCG, Den Haag.

Halse, A., H. Samstad, M. Killi, S. Flügel and F. Ramjerdi (2010) Valuation of freight transport time and reliability (in Norwegian), TØI report 1083/2010, Oslo.

Halse, A.H. and M. Killi (2012) Verdsetting av tid og pålitelighet for godstransport på jernbane (Values of transport time and reliability for railway freight), TØI report 1189/2012, TØI, Oslo.

Hamer, R., G.C. De Jong, and E.P Kroes. (2005) The value of reliability in Transport – Provisional values for the Netherlands based on expert opinion, RAND Technical Report Series, TR-240-AVV, Netherlands.

HEATCO (2006) Developing Harmonised European Approaches for Transport Costing and Project Assessment, Deliverable 5, Proposal for harmonized guidelines. IER, University of Stuttgart.

Hensher, D.A. (2007) Valuation of Travel Time Savings, prepared for the Handbook in Transport Economics, edited by André de Palma, Robin Lindsey, Emile Quinet, Roger Vickerman, Edward Elgar Publisher.

Hensher, D.A., W.H. Greene and Z. Li, (2011a) Embedding risk attitude and decision weights in non-linear logit to accommodate time variability in the value of expected travel time savings, *Transportation Research Part B: Methodological*, Volume 45, Issue 7, August 2011, 954-972

Hensher, D.A., Z. Li and J.M. Rose (2011b, Forthcoming) Accommodating Risk in the Valuation of Expected Travel Time Savings, Journal of Advanced Transportation, accepted for publication.

Hensher, D.A., J.M. Rose and Z. Li (2011c) Does the choice model method and/or the data matter? *Transportation*, DOI 10.1007/s11116-011-9329-x.

Hensher, D.A. and Z. Li (2012) Valuing travel time variability within a rank-Dependent Utility framework and an investigation of unobserved taste heterogeneity, *Journal of Transport Economics and Policy*, In press.

Hess, S., M. Ben-Akiva, D. Gopinath and J. Walker (2011) Advantages of latent class over continuous mixture of logit models, Working paper, ITS, University of Leeds.

Hjorth, K. and F. Ramjerdi (2011) A prospect theory approach to travel time reliability, Paper presented at Second International Choice Modelling Conference, Leeds

Jong, G.C. de, E.P. Kroes, R. Plasmeijer, P. Sanders and P. Warffemius (2004) The value of reliability, *paper presented at ETC 2004*, Strasbourg.

Jong, G.C. and M.E. Ben-Akiva (2007) A micro-simulation model of shipment size and transport chain choice, Special issue on freight transport of *Transportation Research B*, 41, 950-965.

Jong, G.C. de, A.J. Daly, M. Pieters. S. Miller, R. Plasmeijer and F. Hofman (2007) Uncertainty in traffic forecasts: literature review and new results for The Netherlands, *Transportation*, 34, 375-395.

Jong, G.C. de (2008) Value of freight travel-time savings, revised and extended chapter for Handbooks in Transport, Volume 1: *Handbook of Transport Modelling* (Eds: D.A. Hensher and K.J. Button), Elsevier.

Jong, G.C. de, M. Kouwenhoven, E.P. Kroes, P. Rietveld and P. Warffemius (2009) Preliminary monetary values for the reliability of travel times in freight transport, in *European Journal of Transport and Infrastructure Research,* Issue 9(2), pp. 83-99.

Kaa, E.J. van de (2008) Extended prospect theory, Findings on choice behaviour from economics and the behavioural sciences and their relevance for travel behaviour, PhD Thesis, Delft University of Technology.

Kahneman, D. and A. Tversky (1979) Prospect theory: an analysis of decision under risk, *Econometrica*, 47, 263-291.

Kahneman, D. and A. Tversky (1992) Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty*, 5 (4), 297-323.

Koster, P. and Verhoef, E.T. (2012) A rank dependent scheduling model, Forthcoming in *Journal of Transport Economics and Policy*.

Lancsar, E. and J. Louviere (2006) Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences, *Health Economics,* 15 797-811.

Mackie, P.J., M. Wardman, A.S. Fowkes, G. Whelan, J. Nellthorp and J. Bates (2003) Values of travel time savings in the UK, ITS Leeds and John Bates Services report for the UK Department for Transport, Leeds.

Mahmassani, H.S. (2011) Application to the New York metropolitan region of an integrated model of user responses to pricing and reliability with a state of the art simulation-based dynamic traffic assignment tool, Presentation to NYMTC, New York City.

McFadden, D. (1978) Modelling the choice of residential location. In Karlqvist, A., Lundqvist, L., Snickars, F. and Weibull, J. (eds) *Spatial Interaction Theory and Residential Location*. North-Holland, Amsterdam.

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

McKinsey (1986) Afrekenen met files, McKinsey & Company, Amsterdam.

Minken, H. and H. Samstad (2006) Appraising policies to reduce freight transport value of time, and its variability – a new method (in Norwegian), TØI report 825/2006, Oslo.

MVA (1996) Benefits of reduced travel time variability; report to DfT; MVA.

MVA (2000) Etude de l'impact des phénomènes d'irregularité des autobus – Analyse des resultats, MVA, Paris.

NEA (1990) Rekening Rijden en het goederenvervoer over de weg, Rapport 90183/12526, NEA, Rijswijk.

NEA, Transcare and TNO Inro (2003) Factorkosten van het goedererenvervoer: een analyse van de ontwikkeling in de tijd, rapport voor AVV, NEA, Rijswijk.

NEA (2011) Kostenbarometer goederenvervoer, rapport voor DVS, NEA, Zoetermeer.

NERA, MVA, STM and ITS Leeds (1997) The potential for rail freight, A report for the Office of the Rail Regulator, ORR, London.

OECD (2010) Improving reliability on surface transport networks, OECD, Paris.

Ploos van Amstel, W. and L. Tavasszy (2011) Praktijkvalidatie van uitkomsten VOT/VOR voor goederenvervoer, TNO, Delft

Ramjerdi, F., Flügel, S., Samstad, H. and Killi, M. (2010) Value of time, safety and environment in passenger transport – Time. TØI report 1053B/2010, Institute of Transport Economics, Oslo

RAND Europe (2004) De Waardering van kwaliteit en betrouwbaarheid in personen- en goederen vervoer (The valuation of quality and reliability in passenger and freight transport). AVV/RAND Europe, Rotterdam.

RAND Europe, SEO and Veldkamp/NIPO (2004) Hoofdonderzoek naar de reistijdwaardering in het goederenvervoer, rapport TR-154-AVV voor AVV, RAND Europe, Leiden.

Rijkswaterstaat (2011) Geindexeerde VoT's met 2010 als basisjaar: http://www.rws.nl/images/VoT%20GV%20Alle%20modaliteiten%20basisjaar%202010%20EC_tcm174-296215.pdf (retrieved 31 July 2012)

SEO (2011) Effecten van de voorgenomen wijziging van de tariefstructuur op Schiphol, report for DGLM, SEO, Amsterdam.

Shires, J.D. and G.C. de Jong (2009) An international meta-analysis of value of travel time savings, *Evaluation and Program Planning*, Vol. 32(4), pp. 315-325.

Significance, VU University Amsterdam and John Bates (2007) The value of travel time and travel time reliability, Survey design, Final Report prepared for the Netherlands Ministry of Transport, Public Works and Water Management, Significance. Leiden.

Significance, NEA and DEMIS (2010) Schatting BASGOED, rapportage DP1, Report for DVS Rijkswaterstaat, Significance, The Hague.

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

Stathopoulos, A.I. and S. Hess (2011) Revisiting reference point formation, gains-losses asymmetry and non-linear sensitivities: one size does not fit all! Paper presented at ETC 201, Glasgow.

Swahn, H. and J. Bates (2010) Time and quality in freight transport, A pre-study for VTI, VTI, Stockholm.

Swait, J., (1994) A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data, *Journal of Retailing and Consumer Services* 1(2), 77-89.

Tavasszy, L.A. (2008) Measuring value of time in freight transport: a systems perspective. in *Recent developments in transport modelling: lessons for the freight sector* (Eds: M.E. Ben-Akiva, H. Meersman and E. van de Voorde), Emerald, 2008.

Tavasszy, L.A. and F. Combes (2010) Endogenous value of time in freight transport models, in *Applied Transport Economics (*Eds: E. van de Voorde and T. Vanelslander), De Boeck, Antwerpen.

TNO-Inro en Muconsult (2002) Actualisering van de kengetallen voor tijdwaardering in het goederenvervoer: plan van aanpak, Eindrapportage; TNO-Inro rapport 2002-34, Delft.

Ton, J. and L. Tavasszy (2010) V&W onderzoek naar waardering reistijd en –betrouwbaarheid: quick scan representativiteit steekproef goederenvervoer, memorandum TNO, Delft.

Train, K.E., (2008) EM Algorithms for Nonparametric Estimation of Mixing Distributions., *Journal of Choice Modelling*, 1(1), 40-69.

Train, K.E., (2009) Discrete Choice Methods with Simulation, Second edition, Cambridge University Press, New York.

Tseng, Y.Y. and Verhoef, E.T. (2008) Value of time by time of day: A stated-preference study, *Transportation Research B*, 42(7-8), 607-618.

Tseng, Y.Y., E.T. Verhoef, G,.C. de Jong, M. Kouwenhoven and A.I.J.M. van der Hoorn (2009) A pilot study into the perception of unreliability of travel times using in-depth interviews, *Journal of Choice Modelling*, 2(1), 8-28.

Vickrey, W.S. (1973) Pricing, metering, and efficiently using urban transport facilities, *Highway Research Record* 476, 36-48.

Vickrey, W.S. (1969) Congestion theory and transport investment, *American Economic Review (Papers and Proceedings)* 59, 251-261.

Wardman, M. (2001) A review of British evidence on time and service quality valuations, in: *Transportation Research Part E* 37, p. 107-128.

Wardman, M., P. Chintakayala, G.C. de Jong and D. Ferrer (2012) European wide meta-analysis of values of travel time; Paper prepared for EIB, ITS Leeds,

Zamparini, L. and A. Reggiani (Freight transport and the value of travel time savings: A meta-analysis of empirical studies, *Transport Reviews*, 5, 621-635.

APPENDICES

Appendix A: Design tables

The following tables are the most important design tables from the design report (Significance et al. 2007) and are updated where necessary. For more explanation, see that report.

The underlying design tables for experiments 1, 2a and 2b are used for both the freight and passenger transport SP experiments. The attribute levels are different for each survey and also for most survey segments. They are presented separately.

Underlying design experiment 1

Set	CardNo	Time_L	Cost_L	Time_R	Cost_R
1	1	0	0	-1	2
1	2	2	-1	0	0
1	3	-1	1	0	0
1	4	2	0	0	1
1	5	-1	0	0	-1
1	6	0	-2	-1	0
2	1	1	0	0	2
2	2	0	1	1	0
2	3	0	-1	-2	0
2	4	0	0	-2	1
2	5	0	2	2	0
2	6	0	0	1	-1
3	1	0	-1	-2	0
3	2	1	-2	0	0
3	3	-2	0	0	-2
3	4	-2	2	0	0
3	5	0	0	2	-2
3	6	0	0	-1	2

Underlying design experiment 2a

Set	CardNo	Time_L	Cost_L	Reliab_L	PrefArr_L	Time_R	Cost_R	Reliab_R	PrefArr_R
1	1	-2	1	-2	2	0	-2	1	1
1	2	0	-1	2	-1	-1	2	2	-2
1	3	1	1	2	-2	2	1	0	-2
1	4	2	1	0	0	1	2	-2	2
1	5	2	-1	-2	-2	1	-2	-1	-2
1	6	1	-2	-1	0	1	-1	0	-1
2	1	-1	-2	-2	1	-2	-2	0	2
2	2	2	0	-1	-1	-2	-1	1	-2
2	3	-1	0	0	-2	-2	2	-1	1
2	4	0	-2	1	-2	-1	-2	-2	-1
2	5	-2	0	2	1	-1	0	0	1
2	6	-2	-1	1	0	-2	1	-2	0
3	1	-2	2	-1	-2	0	1	-1	-1
3	2	0	0	-2	0	0	-1	2	2
3	3	-1	2	2	0	2	0	-1	2
3	4	0	2	0	2	2	-1	-2	1
3	5	-1	-1	-1	2	1	1	2	1
3	6	-1	1	1	-1	-2	0	2	-1
4	1	-2	-2	0	-1	0	0	-2	-2
4	2	1	2	-2	-1	2	-2	2	0
4	3	2	-2	2	2	0	2	0	0
4	4	1	0	1	2	-1	1	1	2
4	5	0	1	-1	1	1	0	1	0
4	6	1	-1	0	1	-1	-1	-1	0
5	1	2	-1	-2	2	2	-2	1	0
5	2	1	1	2	2	0	-1	1	2
5	3	2	2	1	0	-1	-2	2	-1
5	4	-1	-2	-2	0	-2	0	1	-1
5	5	0	-2	1	2	0	1	-2	-1
5	6	2	0	-1	-2	-2	2	-2	1
6	1	-1	-1	-1	1	-2	-1	0	-2
6	2	0	1	-1	0	1	-2	-2	-2
6	3	-1	1	1	-2	0	2	-1	0
6	4	1	-1	0	0	0	0	2	-2
6	5	-2	2	-1	2	1	2	2	2
6	6	2	-2	2	1	-1	2	1	-2
7	1	-2	-2	0	-2	2	0	-2	2
7	2	-2	0	2	0	2	-1	2	1
7	3	1	-2	-1	-1	-1	-1	-2	0
7	4	-1	2	2	-1	1	1	1	1
7	5	1	0	1	1	2	1	-1	-2
7	6	-2	1	-2	1	2	2	0	-1
8	1	1	2	-2	-2	1	-1	-1	-1
8	2	-2	-1	1	-1	0	-2	0	1
8	3	0	2	0	1	1	0	0	0
8	4	0	-1	2	-2	-2	1	2	0
8	5	0	0	-2	-1	-2	-2	-1	2
8	6	2	1	0	-1	-1	0	-1	1
9	1	1	-2	1	2	1	1	-1	2
9	2	2	-2	-1	-1	0	-1	-1	-2
9	3	-1	1	-2	1	0	-2	-2	2
9	4	0	0	0	2	-1	-1	1	1
9	5	-1	-2	0	-2	0	2	2	1
9	6	2	1	2	2	2	2	-2	0
10	1	-1	2	-1	2	-2	1	0	1
10	2	-2	-1	-2	2	2	-2	-1	1
10	3	1	1	-1	0	-2	-1	-2	-1
10	4	1	0	-2	-1	-1	-2	0	0
10	5	0	1	1	-2	1	-2	1	-1
10	6	2	2	-2	-2	-2	-2	2	-2
11	1	0	-1	-1	1	-2	0	-1	0
11	2	1	2	0	1	1	-1	2	0
11	3	-2	0	-1	-2	2	1	2	-1
11	4	2	0	1	1	-1	2	-1	-1
11	5	-2	1	0	-1	2	0	1	-2
11	6	-1	-1	1	-1	0	0	0	-1
12	1	0	2	2	-1	1	2	0	-2
12	2	-2	2	1	0	0	1	1	0
12	3	2	-1	0	0	-2	2	1	2
12	4	1	-1	2	-2	-1	1	-2	-2
12	5	-1	0	2	0	-1	0	2	2
12	6	-2	-2	2	1	1	0	-2	1
13	1	-1	-1	1	2	-1	1	0	2
13	2	0	2	2	2	2	-2	1	0
13	3	-2	1	0	2	0	-2	0	1
13	4	-2	0	-1	1	-1	-1	-2	0
13	5	-2	2	1	-2	1	1	1	1
13	6	0	0	0	0	0	0	2	-2
14	1	2	2	-2	1	2	0	-2	2
14	2	-1	0	2	-2	2	-1	2	1
14	3	0	-2	-2	-2	-2	2	-2	1
14	4	2	1	2	0	-1	2	1	-2
14	5	1	-2	1	0	-2	1	2	0
14	6	1	1	-1	-2	0	-1	1	2
15	1	-1	2	-1	0	-1	-2	2	-1
15	2	2	-1	0	-2	1	0	0	0
15	3	2	-2	-1	2	-2	0	1	-1
15	4	1	2	0	-1	-2	-1	0	-2
15	5	0	-1	-1	-1	0	1	-2	-1
15	6	-2	-2	2	-1	1	-1	-1	-1
16	1	1	-1	2	1	2	2	0	-1
16	2	-1	1	-2	-1	-1	0	-1	1
16	3	-1	-2	0	1	-2	-2	-1	2
16	4	2	0	1	-1	2	1	-1	-2
16	5	-2	-1	-2	0	1	-2	-2	-2
16	6	1	0	-2	2	0	2	-1	0

Underlying design experiment 2b

Set	CardNo	Time_L	Cost_L	Reliab_L	Time_R	Cost_R	Reliab_R
1	1	0	0	0	-1	-1	2
1	2	1	2	0	0	0	1
1	3	0	0	-1	1	-1	0
1	4	2	0	2	0	2	0
1	5	0	-1	-2	-1	0	0
1	6	-1	-1	0	0	0	1
1	7	0	-1	2	-1	0	0
2	1	2	2	-2	0	0	0
2	2	0	0	0	-2	1	1
2	3	-2	2	0	0	0	1
2	4	0	1	0	1	0	-1
2	5	2	0	-2	0	-2	0
2	6	-1	-1	0	0	0	1
2	7	-2	0	0	0	2	-1
3	1	2	-2	2	0	0	0
3	2	-1	2	2	0	0	0
3	3	1	-2	0	0	0	-1
3	4	0	-1	0	-1	0	2
3	5	1	0	0	0	2	1
3	6	-1	-1	0	0	0	1
3	7	1	0	0	0	-2	1
4	1	-1	2	-2	0	0	0
4	2	0	0	0	2	-1	-2
4	3	0	0	1	-2	1	0
4	4	0	-1	0	-2	0	1
4	5	-1	0	2	0	2	0
4	6	-1	-1	0	0	0	1
4	7	-1	0	0	0	2	-2
5	1	0	0	0	1	1	-1
5	2	2	-2	-2	0	0	0
5	3	0	0	2	-1	1	0
5	4	-2	0	-1	0	-2	0
5	5	-1	0	0	0	-2	-2
5	6	-1	-1	0	0	0	1
5	7	0	-1	1	1	0	0
6	1	-2	-2	1	0	0	0
6	2	2	2	0	0	0	2
6	3	2	-2	0	0	0	-2
6	4	0	1	0	1	0	1
6	5	0	1	1	1	0	0
6	6	-1	-1	0	0	0	1
6	7	1	0	0	0	2	-1
7	1	-2	2	-1	0	0	0
7	2	-1	-2	0	0	0	-2
7	3	0	0	2	2	-1	0
7	4	1	0	1	0	2	0
7	5	0	-1	-1	-2	0	0
7	6	-1	-1	0	0	0	1
7	7	0	1	-1	-2	0	0
8	1	1	-2	1	0	0	0
8	2	1	-2	-1	0	0	0
8	3	1	-2	0	0	0	1
8	4	0	-1	0	-2	0	-1
8	5	0	1	0	-2	0	1
8	6	-1	-1	0	0	0	1
8	7	2	0	0	0	2	-2

Set	CardNo	Time_L	Cost_L	Reliab_L	Time_R	Cost_R	Reliab_R
9	1	0	0	0	-2	-1	1
9	2	0	0	-1	-2	-1	0
9	3	0	0	-2	2	-1	0
9	4	1	0	-1	0	2	0
9	5	0	-1	0	1	0	-1
9	6	-1	-1	0	0	0	1
9	7	2	0	0	0	-2	2
10	1	0	0	0	-2	1	-1
10	2	0	0	2	2	1	0
10	3	0	0	-2	-1	1	0
10	4	2	0	-2	0	2	0
10	5	1	0	-1	0	-2	0
10	6	-1	-1	0	0	0	1
10	7	-2	0	0	0	-2	1
11	1	0	0	0	2	1	-2
11	2	0	0	1	1	1	0
11	3	-1	2	0	0	0	2
11	4	-1	0	-2	0	-2	0
11	5	0	1	2	2	0	0
11	6	-1	-1	0	0	0	1
11	7	-1	0	0	0	-2	2
12	1	-1	-2	2	0	0	0
12	2	-2	2	1	0	0	0
12	3	2	-2	0	0	0	2
12	4	-2	0	1	0	-2	0
12	5	-2	0	1	0	2	0
12	6	-1	-1	0	0	0	1
12	7	0	1	-1	1	0	0
13	1	0	0	0	1	-1	1
13	2	0	0	0	1	-1	-1
13	3	-2	2	0	0	0	-1
13	4	0	1	0	2	0	-2
13	5	2	0	0	0	2	2
13	6	-1	-1	0	0	0	1
13	7	0	-1	1	-2	0	0
14	1	1	2	-1	0	0	0
14	2	0	0	-2	-1	-1	0
14	3	0	0	1	1	-1	0
14	4	0	-1	0	-1	0	-2
14	5	0	-1	0	2	0	-2
14	6	-1	-1	0	0	0	1
14	7	0	-1	2	2	0	0
15	1	0	0	0	-1	1	-2
15	2	0	0	0	-1	1	2
15	3	-1	2	0	0	0	-2
15	4	0	1	0	2	0	2
15	5	-2	0	0	0	-2	-1
15	6	-1	-1	0	0	0	1
15	7	0	1	-2	2	0	0
16	1	0	0	0	2	-1	2
16	2	-2	-2	0	0	0	-1
16	3	0	0	-1	-2	1	0
16	4	-1	0	2	0	-2	0
16	5	0	1	0	-1	0	2
16	6	-1	-1	0	0	0	1
16	7	0	1	-2	-1	0	0

Attribute levels – Passenger transport – Time

All segments, except recreational navigation

BaseTime	**Time level** *(relative to base)* **Passenger**				
(min.)	*level -2*	*level -1*	*level 0*	*level 1*	*level 2*
5 – 9	-2	-1	0	1	3
10 – 19	-3	-1	0	2	5
20 – 44	-5	-2	0	3	8
45 – 74	-10	-5	0	5	15
75 – 119	-15	-5	0	10	25
120 – 179	-15	-10	0	10	30
180 – 239	-20	-10	0	15	40
240 – 359	-40	-20	0	20	60
360 – 539	-60	-30	0	30	90
540 – 1439	-120	-60	0	60	180
1440 +	-240	-120	0	120	360

Attribute levels – Passenger transport – Cost

Car & Public Transport segment – Non-business

BaseTime	**Cost level** *(relative to base, in Euro)* **Non-business passenger**				
(min.)	*level -2*	*level -1*	*level 0*	*level 1*	*level 2*
5 – 9	-0.2	-0.1	0	0.1	0.3
10 – 19	-0.3	-0.1	0	0.2	0.8
20 – 44	-0.6	-0.2	0	0.4	1.5
45 – 74	-1	-0.4	0	0.8	3
75 – 119	-1.5	-0.8	0	1	4
120 – 179	-2	-1	0	1.5	6
180 – 239	-3	-1	0	2	8
240 – 359	-6	-1.5	0	3	12
360 – 539	-8	-2.5	0	4	18
540 – 1439	-18	-5	0	10	40
1440 +	-40	-10	0	20	75

Car & Public Transport segment – Business

BaseTime	**Cost level** *(relative to base, in Euro)*			**Business passenger**	
(min.)	*level -2*	*level -1*	*level 0*	*level 1*	*level 2*
5 – 9	-0.4	-0.1	0	0.3	0.5
10 – 19	-0.5	-0.2	0	0.3	1
20 – 44	-0.9	-0.3	0	0.6	2
45 – 74	-1.5	-0.5	0	1	5
75 – 119	-2	-1	0	1.5	6
120 – 179	-3	-1	0	2	10
180 – 239	-4	-1	0	4	15
240 – 359	-10	-2.5	0	5	20
360 – 539	-15	-4	0	6	30
540 – 1439	-30	-6	0	15	50
1440 +	-50	-15	0	30	100

Air transport segment – Non-business

BaseTime	**Cost level** *(relative to base, in Euro)*			**Non-business passenger**	
(min.)	*level -2*	*level -1*	*level 0*	*level 1*	*level 2*
5 – 9	-0.5	-0.2	0	0.3	0.75
10 – 19	-0.8	-0.2	0	0.5	2
20 – 44	-1.5	-0.5	0	1	3.75
45 – 74	-2.5	-1	0	2	7.5
75 – 119	-4	-2	0	2.5	10
120 – 179	-5	-2.5	0	4	15
180 – 239	-7.5	-2.5	0	5	20
240 – 359	-15	-3.75	0	7.5	30
360 – 539	-20	-6	0	10	45
540 – 1439	-45	-15	0	25	100
1440 +	-100	-25	0	50	200

Air transport segment – Business

BaseTime	**Cost level** *(relative to base, in Euro)*			**Business passenger**	
(min.)	*level -2*	*level -1*	*level 0*	*level 1*	*level 2*
5 – 9	-1	-0.4	0	0.6	1.5
10 – 19	-1.6	-0.4	0	1	4
20 – 44	-3	-1	0	2	7.5
45 – 74	-5	-2	0	4	15
75 – 119	-8	-4	0	5	20
120 – 179	-10	-5	0	8	30
180 – 239	-15	-5	0	10	40
240 – 359	-30	-7.5	0	15	60
360 – 539	-40	-12	0	20	90
540 – 1439	-90	-30	0	50	200
1440 +	-200	-50	0	100	400

Attribuut levels – Passenger transport – Reliability

Car & Air transport segment

Reliability *(relative to time level)*

Base time: 5 – 9 min. CAR

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-1	-1	-1	-1	-1
0	0	0	0	0
0	0	0	0	0
0	2	1	3	4
1	4	5	8	12

Base time: 10 – 19 min. CAR

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-2	-2	-2	-2	-2
0	0	0	0	0
0	0	0	0	0
0	2	4	7	7
2	5	8	15	20

Base time: 20 – 44 min. CAR

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-5	-5	-5	-5	-5
0	0	0	0	0
0	0	0	0	0
0	5	10	15	15
5	10	20	30	40

Base time: 45 – 74 min. CAR

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-10	-10	-10	-10	-10
0	0	0	0	0
0	0	0	0	0
0	10	20	30	30
10	20	40	60	80

Base time: 75 – 119 min. CAR

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-15	-15	-15	-15	-15
0	0	0	0	0
0	0	0	0	0
0	15	30	45	45
15	30	60	90	120

Base time: 120 – 179 min. CAR

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-20	-20	-20	-20	-20
0	0	0	0	0
0	0	0	0	0
0	20	40	60	60
20	40	80	120	160

Base time: 180 – 239 min. CAR

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-30	-30	-30	-30	-30
0	0	0	0	0
0	0	0	0	0
0	30	60	90	90
30	60	120	180	240

Base time: 240 – 359 min. CAR

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-40	-40	-40	-40	-40
0	0	0	0	0
0	0	0	0	0
0	40	80	120	120
40	80	160	240	320

Base time: 360 – 539 min. CAR

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-60	-60	-60	-60	-60
0	0	0	0	0
0	0	0	0	0
0	60	120	180	180
60	120	240	360	480

Base time: 540 – 1439 min. CAR

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-90	-90	-90	-90	-90
0	0	0	0	0
0	0	0	0	0
0	90	180	270	270
90	180	360	540	720

Base time: 1440+ min. CAR

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-240	-240	-240	-240	-240
0	0	0	0	0
0	0	0	0	0
0	240	480	720	720
240	480	960	1440	1920

Public transport segment

Reliability *(relative to time level)*

Base time: 5 – 9 min. PT

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-1	-1	-1	-1	-1
0	0	0	0	0
0	0	0	0	0
0	2	1	3	4
1	4	5	8	12

Base time: 10 – 19 min. PT

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-1	-1	-1	-1	-1
0	0	0	0	0
0	0	0	0	0
0	2	4	7	7
2	5	8	15	20

Base time: 20 – 24 min. PT

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-1	-1	-1	-1	-1
0	0	0	0	0
0	0	0	0	0
0	5	10	15	15
5	10	20	30	40

Base time: 45 – 74 min. PT

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-1	-1	-2	-2	-2
0	0	0	0	0
0	0	0	0	0
0	10	20	30	30
10	20	40	60	80

Base time: 75 – 119 min. PT

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-1	-1	-3	-5	-5
0	0	0	0	0
0	0	0	0	0
0	15	30	45	45
15	30	60	90	120

Base time: 120 – 179 min. PT

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-1	-2	-5	-5	-5
0	0	0	0	0
0	0	0	0	0
0	20	40	60	60
20	40	80	120	160

Base time: 180 – 239 min. PT

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-2	-2	-5	-5	-5
0	0	0	0	0
0	0	0	0	0
0	30	60	90	90
30	60	120	180	240

Base time: 240 – 359 min. PT

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-5	-5	-5	-10	-10
0	0	0	0	0
0	0	0	0	0
0	40	80	120	120
40	80	160	240	320

Base time: 360 – 539 min. PT

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-5	-5	-5	-10	-10
0	0	0	0	0
0	0	0	0	0
0	60	120	180	180
60	120	240	360	480

Base time: 540 – 1439 min. PT

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-5	-10	-10	-15	-15
0	0	0	0	0
0	0	0	0	0
0	90	180	270	270
90	180	360	540	720

Base time: 1440+ min. PT

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-10	-15	-15	-20	-20
0	0	0	0	0
0	0	0	0	0
0	240	480	720	720
240	480	960	1440	1920

Attribute levels – Passenger transport – Arrival Time

All segments, except recreational navigation

BaseTime	**Preferred Arrival Time level** *(relative to base)* **Passenger**				
(min.)	*level -2*	*level -1*	*level 0*	*level 1*	*level 2*
5 – 9	-2	-1	0	1	3
10 – 19	-3	-1	0	2	5
20 – 44	-5	-2	0	3	8
45 – 74	-10	-5	0	5	15
75 – 119	-15	-5	0	10	25
120 – 179	-15	-10	0	10	30
180 – 239	-20	-10	0	15	40
240 – 359	-40	-20	0	20	60
360 – 539	-60	-30	0	30	90
540 – 1439	-120	-60	0	60	180
1440 +	-240	-120	0	120	360

Attribute levels – Passenger transport – Recreational Navigation

Time

Time level *(absolute)* **Recreational Navigation**				
Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
5	10	15	30	60

Cost

Cost level *(absolute, in Euro)* **Recreational Navigation**				
Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
0	0.5	1.5	2.5	5

Reliability

Level 1
5

Reliability *(absolute in minutes)* **Recreational Navigation**				
Level 1	*Level 2*	*Level 3*	*Level 4*	*Level 5*
4	3	0	0	0
5	4	3	0	0
5	5	5	5	10
5	6	7	10	20
6	7	10	15	30

Level 2
10

Level 1	*Level 2*	*Level 3*	*Level 4*	*Level 5*
9	8	5	0	0
10	9	8	5	0
10	10	10	10	10
10	11	12	15	20
11	12	15	20	25

Level 3
15

Level 1	*Level 2*	*Level 3*	*Level 4*	*Level 5*
14	13	10	5	0
15	14	13	10	5
15	15	15	15	10
15	16	17	20	15
16	17	20	25	45

Level 4
30

Level 1	*Level 2*	*Level 3*	*Level 4*	*Level 5*
29	28	25	20	15
30	29	28	25	20
30	30	30	30	25
30	31	32	35	30
31	32	35	40	60

Level 5
60

Level 1	*Level 2*	*Level 3*	*Level 4*	*Level 5*
59	58	55	50	45
60	59	58	55	50
60	60	60	60	55
60	61	62	65	60
61	62	65	70	90

Attribute levels – Freight transport – Time

Road segment

BaseTime	**Time level** *(relative to base)* **Freight by road**				
(min.)	*level -2*	*level -1*	*level 0*	*level 1*	*level 2*
10 – 19	-3	-1	0	2	5
20 – 44	-5	-2	0	3	8
45 – 74	-10	-5	0	5	15
75 – 119	-15	-5	0	10	25
120 – 179	-15	-10	0	10	30
180 – 239	-20	-10	0	15	40
240 – 359	-40	-20	0	20	60
360 – 539	-60	-30	0	30	90
540 – 1439	-120	-60	0	60	180
1440 – 2879	-240	-120	0	120	360
2880 +	-480	-240	0	240	720

All segments except Road

BaseTime	**Time level** *(relative to base)* **Freight other**				
(min.)	*level -2*	*level -1*	*level 0*	*level 1*	*level 2*
10 – 59	-5	-2	0	5	10
60 – 179	-15	-10	0	10	30
180 – 599	-40	-20	0	30	60
600 – 1439	-120	-60	0	90	180
1440 – 2159	-240	-120	0	180	360
2160 – 2879	-360	-180	0	270	540
2880 – 4319	-480	-240	0	360	720
4320 – 5759	-720	-360	0	540	1080
5760 – 10079	-960	-480	0	720	1440
10080 +	-1920	-960	0	1440	2880

Attribute levels – Freight transport – Cost

All segments

Cost level *(relative)* **FREIGHT**				
Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-15%	-5%	0%	+10%	+25%

Attribute levels – Freight transport – Reliability

Road segment

Reliability *(relative to time level)*

Base time: 10 – 19 min. Road

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-2	-2	-2	-2	-2
0	0	0	0	0
0	0	0	0	0
0	2	4	7	7
2	5	8	15	20

Base time: 20 – 24 min. Road

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-5	-5	-5	-5	-5
0	0	0	0	0
0	0	0	0	0
0	5	10	15	15
5	10	20	30	40

Base time: 45 – 74 min. Road

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-10	-10	-10	-10	-10
0	0	0	0	0
0	0	0	0	0
0	10	20	30	30
10	20	40	60	80

Base time: 75 – 119 min. Road

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-15	-15	-15	-15	-15
0	0	0	0	0
0	0	0	0	0
0	15	30	45	45
15	30	60	90	120

Base time: 120 – 179 min. Road

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-20	-20	-20	-20	-20
0	0	0	0	0
0	0	0	0	0
0	20	40	60	60
20	40	80	120	160

Base time: 180 – 239 min. Road

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-30	-30	-30	-30	-30
0	0	0	0	0
0	0	0	0	0
0	30	60	90	90
30	60	120	180	240

Base time: 240 – 359 min. Road

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-40	-40	-40	-40	-40
0	0	0	0	0
0	0	0	0	0
0	40	80	120	120
40	80	160	240	320

Base time: 360 – 539 min. Road

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-60	-60	-60	-60	-60
0	0	0	0	0
0	0	0	0	0
0	60	120	180	180
60	120	240	360	480

Base time: 540 – 1439 min. Road

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-90	-90	-90	-90	-90
0	0	0	0	0
0	0	0	0	0
0	90	180	270	270
90	180	360	540	720

Base time: 1440+ min. Road

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-240	-240	-240	-240	-240
0	0	0	0	0
0	0	0	0	0
0	240	480	720	720
240	480	960	1440	1920

Base time: 1440+ min. Road

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-480	-480	-480	-480	-480
0	0	0	0	0
0	0	0	0	0
0	480	960	1440	1440
480	960	1920	2880	3820

Attribute levels – Freight transport – Reliability

All segments except Road

Reliability *(relative to time level)*

Base time: 10 – 59 min. Other

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-2	-2	-2	-2	-2
0	0	0	0	0
0	0	0	0	0
0	2	4	7	7
2	5	8	15	20

Base time: 60 – 179 min. Other

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-10	-10	-10	-10	-10
0	0	0	0	0
0	0	0	0	0
0	10	20	30	30
10	20	40	60	80

Base time: 180 – 599 min. Other

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-30	-30	-30	-30	-30
0	0	0	0	0
0	0	0	0	0
0	30	60	90	90
30	60	120	180	240

Base time: 600 – 1439 min. Other

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-90	-90	-90	-90	-90
0	0	0	0	0
0	0	0	0	0
0	90	180	270	270
90	180	360	540	720

Base time: 1440 – 2159 min. Other

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-240	-240	-240	-240	-240
0	0	0	0	0
0	0	0	0	0
0	240	480	720	720
240	480	960	1440	1920

Base time: 2160– 2879 min. Other

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-360	-360	-360	-360	-360
0	0	0	0	0
0	0	0	0	0
0	360	720	1080	1080
360	720	1440	1920	2880

Base time: 2880 – 4319 min. Other

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-480	-480	-480	-480	-480
0	0	0	0	0
0	0	0	0	0
0	480	960	1440	1440
480	960	1920	2880	3840

Base time: 4320 – 5759 min. Other

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-720	-720	-720	-720	-720
0	0	0	0	0
0	0	0	0	0
0	720	1440	1920	1920
720	1440	2880	3840	5760

Base time: 5760 – 10079 min. Other

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-960	-960	-960	-960	-960
0	0	0	0	0
0	0	0	0	0
0	960	1920	2880	2880
960	1920	3840	5760	7640

Base time: 10080 min. Other

Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
-1920	-1920	-1920	-1920	-1920
0	0	0	0	0
0	0	0	0	0
0	1920	3840	5760	5760
1920	3840	7680	11520	15280

Attribute levels – Freight transport – Arrival Time

Road segment

BaseTime	**Preferred Arrival Time level** *(relative to base)* **Freight by road**				
(min.)	*level -2*	*level -1*	*level 0*	*level 1*	*level 2*
10 – 19	-3	-1	0	2	5
20 – 44	-5	-2	0	3	8
45 – 74	-10	-5	0	5	15
75 – 119	-15	-5	0	10	25
120 – 179	-15	-10	0	10	30
180 – 239	-20	-10	0	15	40
240 – 359	-40	-20	0	20	60
360 – 539	-60	-30	0	30	90
540 – 1439	-120	-60	0	60	180
1440 – 2879	-240	-120	0	120	360
2880 +	-480	-240	0	240	720

All segments except Road

BaseTime	**Preferred Arrival Time level** *(relative to base)* **Freight other**				
(min.)	*level -2*	*level -1*	*level 0*	*level 1*	*level 2*
10 – 59	0	0	0	0	0
60 – 179	-3	-1	0	2	5
180 – 599	-10	-5	0	5	15
600 – 1439	-20	-10	0	15	40
1440 – 2879	-120	-60	0	60	180
2880 – 5759	-240	-120	0	120	360
5760 – 10079	-480	-240	0	240	720
10080 +	-960	-480	0	480	1440

Attribute levels – Freight transport – Inland Waterways and Sea Transport

Notes:

- All numbers below are **multiplicative** factors on the BaseWaitTime, , presented wait time, total transport costs and cost for use of the quay, loading and unloading.
- The following minimum BaseWaitTimes apply
  - If experiment for bridge & BaseWaitTime < 10 min.then BaseWaitTime = 10 min.
  - If experiment for lock & BaseWaitTime < 15 min.then BaseWaitTime = 15 min.
  - If experiment for quay & BaseWaitTime < 60 min.then BaseWaitTime = 60 min.

A. Experiment for locks and bridges

Wait time level *(factor on BaseWaitTime)* **IWW and sea transport**				
Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
0.6	0.85	1.0	1.2	1.4

Total transport Cost *(BaseCost)* **IWW and sea transport**				
Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
0.94	0.98	1.0	1.02	1.05

Reliability *(factor on Wait time as presented)* **IWW and sea transport**				
Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
0.95	0.9	0.75	0.65	0.6
1.0	0.95	0.9	0.75	0.65
1.0	1.0	1.0	1.0	0.75
1.0	1.05	1.1	1.25	1.0
1.05	1.1	1.25	1.35	2.0

B. Experiment for quays/port terminals

Wait time level *(factor on observed)* **IWW and sea transport**				
Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
0.6	0.85	1.0	1.2	1.4

Cost for quay, (un)loading *(factor on observed)* **IWW and sea transport**				
Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
0.70	0.90	1.0	1.15	1.25

Reliability *(factor on presented wait time)* **IWW and sea transport**				
Level -2	*Level -1*	*Level 0*	*Level 1*	*Level 2*
0.95	0.9	0.75	0.65	0.6
1.0	0.95	0.9	0.75	0.65
1.0	1.0	1.0	1.0	0.75
1.0	1.05	1.1	1.25	1.0
1.05	1.1	1.25	1.35	2.0

Appendix B: Re-analysis of the 1997 and 2009/2011 data for passenger transport

In this appendix we report on the analysis of the data collected in 1997 using the same methodology as we have used for the analysis of the 2009/2011 data. In order to make a fair comparison, it was necessary to re-analyse the 2009/2011 as well.

To be able to determine the effect of the changes in methodology for analysing the data, we report the VOTs for each step in Figures B1 – B3 (for commute, business and other purpose trips respectively). Each step is numbered (1A, 1B etc.). The VOTs for each step are displayed to the right of each block in green (C, T, B, A for car, train, BTM and all surface modes). Note that all 1997 VOTs are in Euro (price level 1997) and all 2009/2011 VOTs are in Euro (price level 2010). The original money units in 1997 were guilders, but we converted these to 1997 euros using a factor *ƒ*2.20371 = € 1.

The percentages in red indicate the relative changes of the VOT compared to the previous step. The percentages in blue indicate the differences between the 1997 and the 2009/2011 VOTs at the final step of the analysis.

The 1997 analysis can be summarised in three steps:

- Step 1A: a default MNL model estimated in preference space
- Step 2A: addition of socio-economic factors, as reported in Hague Consulting Group (1998) using the so-called "new specification"
- Step 3A: expansion based on OVG 1995.

The 2009/2011 analysis is discussed in the main part of this report and consists of 5 steps

- Step 1H: a default MNL model estimated in WTP space
- Step 12H: an advanced MNL model
- Step 13H: addition of socio-economic factors (note that these can also be added to the simple MNL model, see step 2I)
- Step 5H/6H: expansion to OViN 2010 excluding (step 5H) and including (step 6H) expansion to the travel time distribution
- Step 7H: estimation of a latent class model

In Section 8.12.2 we already identified seven methodological changes between the two studies. The effects of each of these changes are discussed below.

Methodological change 1: sample enumeration

We were able to re-estimate exactly the same models (step 1A and 1B), with exactly the same estimation results as in the analysis that was carried out in 1998 (Hague Consulting Group, 1998). The expansion in 1998 was done in a different way than the sample enumeration approach that is currently used. We have not tried to re-do the expansion in the same way as we was done in 1998, instead we have used the new sample enumeration technique for the 1997 models as well (step 4A). It is striking that for commuting the differences between the results from step 3A (published result in 1998) and step 4A (same models but with sample enumeration technique) are quite different. The average over all modes does not change much, but in 1997 the VOTs for car, train and BTM were quite similar, whereas according to the new sample enumeration technique they are quite different.

The VOTs per mode change between -21% and +10%. The average VOT over all modes changes by +4%.

Methodological change 2: utility space versus WTP space

We now move to estimate and apply models that were not reported in 1998. First, we re-estimate the 1997 data in WTP space. From Figures B1-B3, we see that -as expected- a simple MNL model in utility space (step 1A) and a simple MNL model in WTP space (step 1B) lead to the same VOT.

The model used in 1998 to derive the recommended VOTs included socio-economic factors. If these are added, we do obtain differences in the average VOT between preference space (step 2A) and WTP space (step 2B), but these differences are small. The same is true if we explicitly add mode specific interaction factors. (steps 1C and 2C). If we expand these results using the weights based on the OVG 1995 (step 3C), we obtain VOTs that can be directly compared with the published results and with our own results from step 4A.

The VOTs per mode change between -6% and +3%. The average VOT over all modes changes by -4%.

Methodological change 3: different set of socio-economic factors

In the 2009/2011 analysis we tested whether the inclusion of several socio-economic interaction variables improved the model. Something similar was done in 1997. However, the set of significant interaction variables differed between the two studies. In order to understand the effect of this, we introduced the 2009/2011 set of interaction variables into the 1997 models (step 2D before expansion and 3D after expansion). However, the education variable, which was one of the socio-economic variables in the 2009/2011 models, is not available in the 1997 survey and was left out.

If we compare the results from step 3C with 3D, we see that the VOTs per mode change between +1% and -2%. The average VOT over all modes changes by +1%.

In order to make a fair comparison with the 2009/2011 survey, we re-analysed this survey as well including mode-specific interaction variables (step 1G and 2G) and excluding the education factors (step 2F and 3F).

Methodological change 4: OVG-1995 versus OViN-2010

In the expansion procedure of the 2009/2011 analysis we use a set of variable distributions (income, age, education, household composition, mode, peak/off-peak) as targets for the iterative proportional fitting technique that determines the weight factors. This set differs from the 1997 set. Furthermore, the target distributions themselves will be different between the OVG-1995 and the OViN-2010. These changes are made in the analysis in step 13D and 4D. Note that for commute and for other the effect of step 13D is roughly opposite to the effect of step 4D.

If we compare the results from step 4D with 3D (i.e. the combined effect of these two changes), we see that the VOTs per mode change between -8% and +9%. The average VOT over all modes changes between -7% and +9%. Averaged over all modes and purposes the VOT remains the same.

Methodological change 5: simple versus advanced MNL models.

In step 5D we have replaced the simple MNL model with an advanced MNL model that includes gamma and lambda exponents that indicate possible decreasing sensitivities with increasing (delta) times and costs. The effect of this on the 1997 commute data is quite small (-2% to +1%). On the business data, the effect is somewhat larger (-5% to -8%) and on the other data the effect is again stronger (-11% to -12%). Note that there is very little differences of the effect between the modes.

We can also look at the effect of this methodological change in the 2009/2011 analysis. For this, we need to compare step 4F with 5F[32]. For commute and business, the general effect is small (-1% to -4%), but for BTM the effect is quite large (-19% to +19%). Also for other purposes, the effect is larger (20% on the average VOT, up to +47% for an individual mode).

Averaged over all modes and purposes and over both data sets the VOT remains the same.

Methodological change 6: expansion towards the travel time distribution

This additional target for the iterative proportional fitting procedure is only necessary if the VOT depends on the travel time / cost, i.e. for advanced MNL models. From the comparison between steps 4D and 10D (1997 data) and 4F and 10F (2009/2011 data) it can be seen that the addition of this target has some effect on the outcomes of simple MNL models (probably because the other socio-economic variables have some correlation with travel time/cost), but this effect is rather small.

In step 6D (1997 data) and 6F (2009/2011 data) this target is added to the procedure. The effect per mode is in most cases very small (about 0%, especially for the 1997 data), but in individual cases it can be up to -22% / +19% (for the 2009/2011 data). Averaged over all modes and purposes and over both data sets the VOT decreases by -2%.

---

[32] We can also compare steps 3I with 5H. (this pair differs from the 4F/5F pair because of the inclusion of mode specific factors, the exclusion of the education interaction variables and the education variable distribution in the expansion procedure)

Methodological change 7: panel-LC models instead of MNL models

Panel-LC models have a strong upwards effect on the VOT. This can be explained by the fact that these types of models take more heterogeneity into account, whereas the bias due to panel effects (i.e. one respondent has given multiple (and possibly correlated) answers) has been deleted. Generally, the VOTs increase by 20% to 40% (both by mode and on average). Averaged over all modes and purposes and over both data sets the VOT increases by +31%.

Fair comparison between 1997 and 2009/2011 results.

The best comparison between 1997 and 2009/2011 can be made by comparing similar LC models based on both data sets, i.e. comparing models in step 7D and 7F. The results are in Table B1. Note that the resulting VOTs from 7F are different from the ones published in the main part of the report, since we made some changes to the analysis in order to make a fair comparison (see also footnote 32).

Table B1: Growth of VOT over the period 1997 – 2010 based on a fair comparison between the two surveys (no correction for inflation)

	Car	Train	Bus, tram, metro	All surface modes
Commute	+13%	+49%	+9%	+23%
Business	+13%	+60%	+98%	+26%
Other	+71%	+59%	+52%	+65%

InTable B1 we see that in the period 1997-2010 for all purposes there is a substantial increase in the train VOT, whereas for car the increase on average is much more limited (and bus/tram/metro have an intermediate position). Possible explanations are that in 1997 the mobile phone was already rather common (as was the laptop), so this could be used in the train, but not so much in the car, since the mobile phone car kits became popular later than 1997. Increased crowding in the trains could also have had an impact.

We also see a substantial increase in the other VOT for all modes, which does not occur with the same consistency for commuting and business travel. We think that the importance of new communication technology is greater for commuting and business trips than for other trips, since there is less pressure to use communication technology for work-related purposes and often the other trips are made together with other people and these trips are rather short.

With regards to the very high increase in the VOT for business travel using bus/tram/metro, one should keep in mind that this is a very small group, and also the changes will be based on small numbers and be less reliable.

Figure B1: Step-by-step analysis of 1997 and 2009/2011 commute data

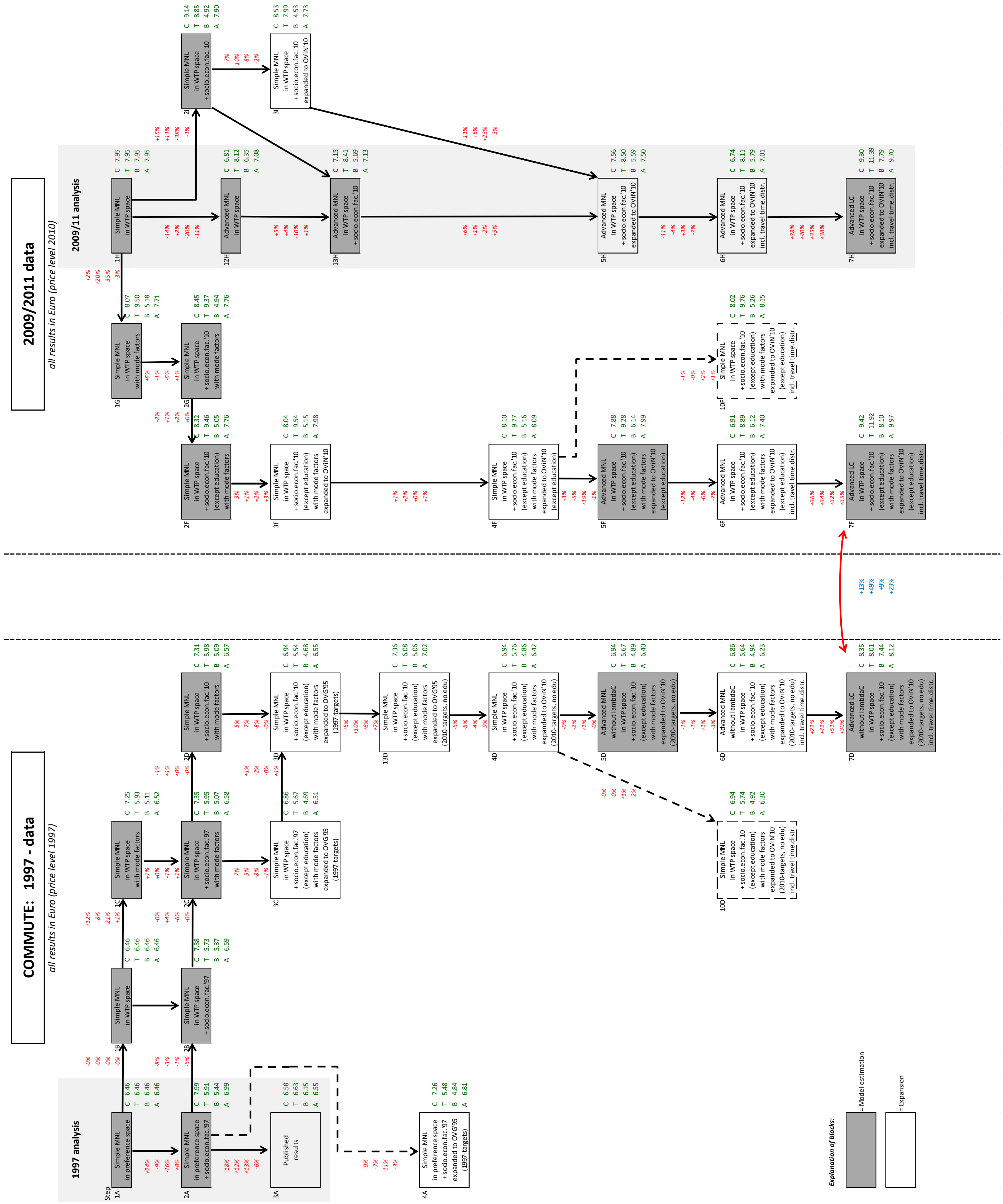


Figure B2: Step-by-step analysis of 1997 and 2009/2011 business data

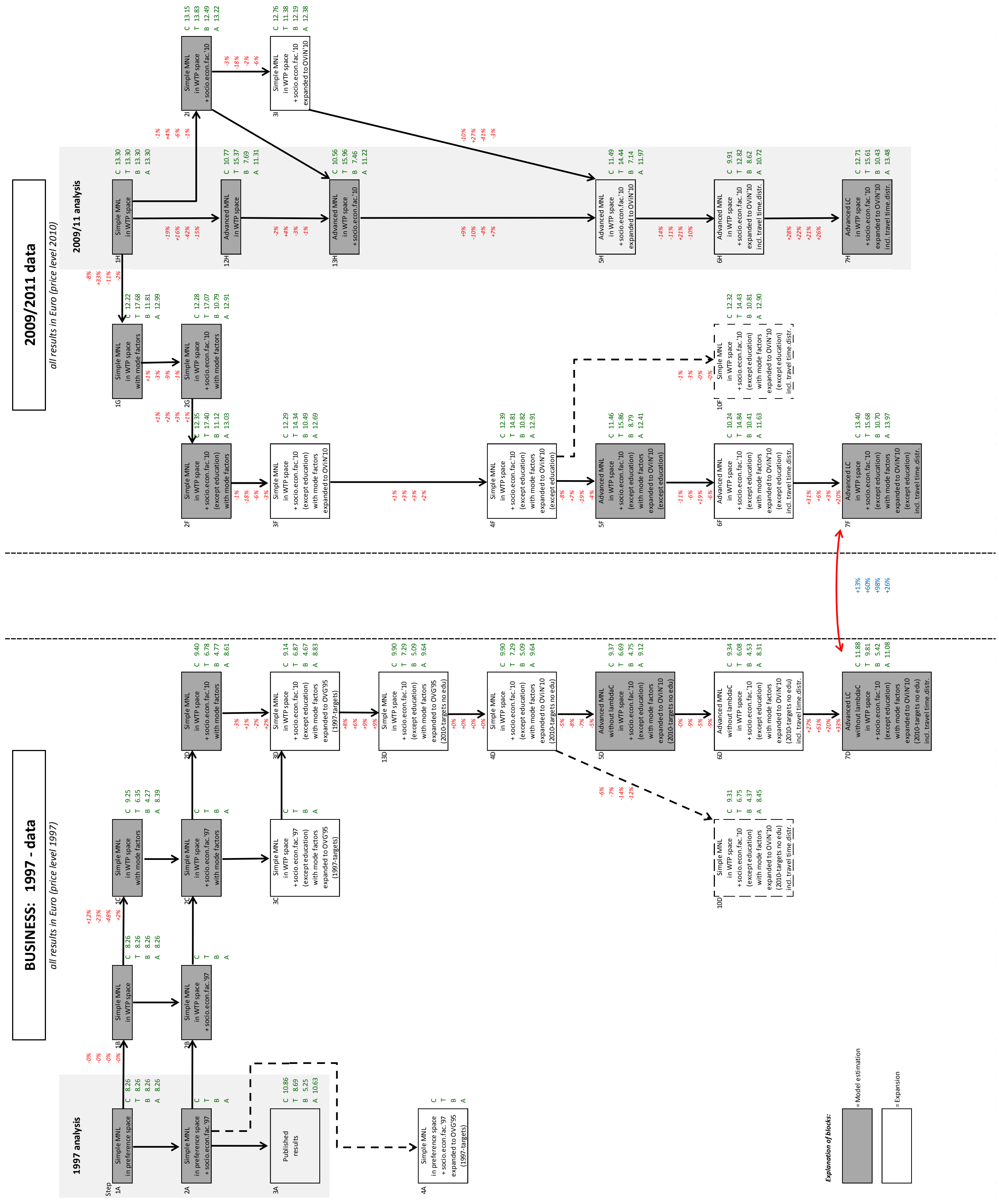


Figure B3: Step-by-step analysis of 1997 and 2009/2011 other data