

## **Updating the Value of Travel Time: Scoping Study**

24 March 2025

Institute for Transport Studies, Andrew Gordon Ltd., Intersection Analytics Ltd., Peak Economics Ltd.  
Report to the Department for Transport

# **Updating the Value of Travel Time: Scoping Study**

Report to the Department for Transport

Institute for Transport Studies (ITS), University of Leeds

Andrew Gordon Ltd.,

Intersection Analytics Ltd.,

Peak Economics Ltd.

## **Authors:**

Thijs Dekker (ITS)

Richard Batley (ITS)

Manuel Ojeda-Cabral (ITS)

Gerard de Jong (ITS)

Stephane Hess (ITS)

Andrew Gordon (Andrew Gordon Consulting Ltd.)

Jacek Pawlak (Intersection Analytics Ltd.)

Aruna Sivakumar (Intersection Analytics Ltd.)

James Laird (Peak Economics Ltd.)

## **With inputs from:**

Rob Sheldon (Accent)

## Contents

|       |                                                                                              |    |
|-------|----------------------------------------------------------------------------------------------|----|
| 1.    | Introduction .....                                                                           | 8  |
| 1.1   | Part 1: Review of methods and evidence.....                                                  | 8  |
| 1.2   | Part 2: Scoping of research options and recommended research programme .....                 | 10 |
| 2.    | History of UK VTT studies, recent international VTT studies, and societal developments ..... | 11 |
| 2.1   | History of UK VTT studies.....                                                               | 11 |
| 2.2   | Overview of the UK 2014/15 study.....                                                        | 11 |
| 2.3   | Follow-up work following ARUP et al. (2015a) .....                                           | 15 |
| 2.4   | Recent national VTT studies outside of the UK.....                                           | 16 |
| 2.4.1 | Overview of the German 2012/13 study .....                                                   | 16 |
| 2.4.2 | Overview of the Singapore 2015 study .....                                                   | 19 |
| 2.4.3 | Overview of the Norwegian 2018/19 study.....                                                 | 22 |
| 2.4.4 | Overview of the Dutch 2022 study .....                                                       | 24 |
| 2.4.5 | Summary of recent national VTT studies .....                                                 | 27 |
| 2.5   | Societal development since the last UK national VTT study .....                              | 30 |
| 3.    | Developments in surveying and estimation methods .....                                       | 35 |
| 3.1   | The potential use of Revealed Preference data .....                                          | 35 |
| 3.2   | State-of-the-art design of VTT SP studies .....                                              | 40 |
| 3.2.1 | Simple time vs. cost trade-offs.....                                                         | 40 |
| 3.2.2 | Experimental designs .....                                                                   | 41 |
| 3.2.3 | Short vs long term choices .....                                                             | 44 |
| 3.3   | Survey administration .....                                                                  | 44 |
| 3.3.1 | Representativeness .....                                                                     | 44 |
| 3.3.2 | Selection bias and methods of recruitment.....                                               | 45 |
| 3.4   | State-of-the-art estimation of VTT .....                                                     | 48 |
| 3.4.1 | Random Utility vs. Random Valuation .....                                                    | 48 |
| 3.4.2 | Additive vs Multiplicative error terms.....                                                  | 49 |
| 3.4.3 | Reference dependence or size and sign effects .....                                          | 51 |
| 3.4.4 | Unobserved preference heterogeneity .....                                                    | 53 |
| 3.4.5 | Observed preference heterogeneity.....                                                       | 53 |
| 3.5   | Summary of development in surveying and estimation methods .....                             | 54 |
| 4.    | Segmentation of the Value of Travel Time .....                                               | 56 |
| 4.1   | User group and trip purpose segmentation in appraisal .....                                  | 56 |
| 4.1.1 | UK-appraisal guidance with respect to segmentation by trip characteristics.....              | 56 |
| 4.1.2 | User group and trip purpose segmentation in appraisal – a look forward. ....                 | 58 |
| 4.2   | Segmentation in UK modelling practice.....                                                   | 60 |

|       |                                                                                   |     |
|-------|-----------------------------------------------------------------------------------|-----|
| 4.3   | Agent-based modelling .....                                                       | 63  |
| 4.4   | Mode segmentation: mode vs. user effects.....                                     | 63  |
| 4.4.1 | Context and problem statement .....                                               | 63  |
| 4.4.2 | Review of evidence .....                                                          | 64  |
| 4.4.3 | Taking stock, challenges, and opportunities .....                                 | 70  |
| 4.5   | Treatment of income.....                                                          | 71  |
| 4.5.1 | Estimation of income effects and elasticities in behavioural VTT models .....     | 71  |
| 4.5.2 | The role of income in appraisal practices .....                                   | 74  |
| 4.6   | Conclusions on segmentation .....                                                 | 77  |
| 5.    | Valuing savings in business value of travel time and worthwhile use of time ..... | 81  |
| 5.1   | The history of valuing savings in business travel time .....                      | 81  |
| 5.2   | The alternative methodologies to valuing savings in business travel time.....     | 81  |
| 5.2.1 | The Cost-Saving Approach.....                                                     | 81  |
| 5.2.2 | The Hensher Equation.....                                                         | 82  |
| 5.2.3 | The Willingness-to-pay approach.....                                              | 83  |
| 5.3   | Application of the WTP approach in the 2014/15 study.....                         | 84  |
| 5.4   | Application of the Hensher Equation to the 2014-15 data .....                     | 86  |
| 5.5   | Capturing of VBTT in other national VTT studies .....                             | 88  |
| 5.6   | Conclusions on the VBTT.....                                                      | 89  |
| 5.7   | Worthwhile use of travel time – general state of research.....                    | 89  |
| 5.8   | Worthwhile use of travel time – theoretical and conceptual developments .....     | 95  |
| 5.9   | Estimating the effects of travel time use on the VTT .....                        | 98  |
| 5.10  | VTT impacts of travel activities in autonomous vehicles .....                     | 101 |
| 5.11  | Novel aspects concerning measurement of travel time activities .....              | 105 |
| 5.12  | Key takeaways on worthwhile use of travel time.....                               | 106 |
| 6.    | Congestion, reliability, and crowding .....                                       | 110 |
| 6.1   | Congested values of travel time (CVTT) and congestion multipliers.....            | 111 |
| 6.1.1 | UK position – 2014/15 study.....                                                  | 111 |
| 6.1.2 | UK position – CVTT workstream.....                                                | 112 |
| 6.1.3 | Empirical evidence on congestion multipliers .....                                | 116 |
| 6.2   | The Value of travel time reliability (VTTR) and the reliability ratio.....        | 120 |
| 6.2.1 | Empirical evidence on the VTTR and the reliability ratio .....                    | 120 |
| 6.2.2 | Modelling of travel time reliability in the UK and beyond .....                   | 122 |
| 6.3   | Interaction between congestion and reliability .....                              | 124 |
| 6.4   | Crowding effects.....                                                             | 124 |
| 6.5   | Conclusions in relation to congestion, reliability, and crowding.....             | 129 |

|       |                                                                    |     |
|-------|--------------------------------------------------------------------|-----|
| 7.    | Active and emerging modes of travel .....                          | 131 |
| 7.1   | Survey design in ARUP et al. (2015b) .....                         | 131 |
| 7.2   | Active travel in other national VTT studies .....                  | 134 |
| 7.3   | Moving forward with obtaining VTT estimates for active travel..... | 140 |
| 8.    | Aviation .....                                                     | 143 |
| 8.1   | Official national guidance.....                                    | 143 |
| 8.1.1 | UK position .....                                                  | 143 |
| 8.1.2 | US position .....                                                  | 144 |
| 8.1.3 | Dutch position .....                                               | 145 |
| 8.1.4 | Norwegian position .....                                           | 146 |
| 8.1.5 | German position .....                                              | 146 |
| 8.2   | Other studies.....                                                 | 147 |
| 8.3   | Findings and conclusions .....                                     | 148 |
| 9.    | Is this the right time for a new national VTT study? .....         | 149 |
| 10.   | Evaluating the design of the ARUP et al. (2015a) study.....        | 150 |
| 10.1  | Segmentation by mode and purpose.....                              | 150 |
| 10.2  | SP-Experiments .....                                               | 151 |
| 10.3  | Covariates.....                                                    | 152 |
| 10.4  | RP component.....                                                  | 153 |
| 10.5  | Recruitment methods .....                                          | 153 |
| 10.6  | Sampling targets.....                                              | 154 |
| 10.7  | Developing and testing approach .....                              | 154 |
| 10.8  | Data analysis and behavioural VTTs .....                           | 154 |
| 10.9  | Translation into appraisal VTTs.....                               | 156 |
| 10.10 | Conclusions regarding the design of ARUP et al. (2015a).....       | 157 |
| 11.   | Reflection on the level of segmentation .....                      | 158 |
| 12.   | Essential areas to consider beyond ARUP et al. (2015a).....        | 160 |
| 12.1  | Separation of mode- and user-effects.....                          | 160 |
| 12.2  | Revealed preferences.....                                          | 161 |
| 12.3  | Active travel .....                                                | 162 |
| 12.4  | Aviation .....                                                     | 164 |
| 13.   | Exploratory areas of research .....                                | 165 |
| 13.1. | Relationship between congestion and reliability .....              | 165 |
| 13.2. | Long term SP choices .....                                         | 166 |
| 14.   | Summary of recommendations .....                                   | 166 |
| 15.   | References.....                                                    | 171 |



## Glossary

|             |                                  |
|-------------|----------------------------------|
| <b>BVTT</b> | Boundary Value of Travel Time    |
| <b>CAA</b>  | Civil Aviation Authority         |
| <b>CVTT</b> | Congested Value of Travel Time   |
| <b>DfT</b>  | Department for Transport         |
| <b>EG</b>   | Equivalent Gains                 |
| <b>EL</b>   | Equivalent Losses                |
| <b>RP</b>   | Revealed Preference              |
| <b>RU</b>   | Random Utility                   |
| <b>RV</b>   | Random Valuation                 |
| <b>SP</b>   | Stated Preference                |
| <b>TAG</b>  | Transport Analysis Guidance      |
| <b>VBTT</b> | Value of Business Travel Time    |
| <b>VTT</b>  | Value of Travel Time             |
| <b>VTTR</b> | Value of Travel Time Reliability |
| <b>WTA</b>  | Willingness to Accept            |
| <b>WTP</b>  | Willingness to Pay               |

## 1. Introduction

The Value of Travel Time (VTT; also commonly referred to as the Value of Travel Time Savings or simply the Value of Time) is one of the fundamental sets of values used by the DfT to appraise transport schemes (Daly and Hess, 2020). They represent the monetised value of changes in forecast travel time resulting from an intervention, as well as capturing the ability of travellers to both spend time more productively, and more comfortably.

Technical details of the current approach to VTT are set out in Section 4 of TAG Unit A1.3 of the DfT's Transport Analysis Guidance (TAG - Department for Transport, 2022a), alongside details of the related Values of Travel Time Reliability (VTTR) (Section 6). The current national average monetary values used by the DfT are displayed in Table A1.3.1 of the TAG Data Book (Department for Transport, 2024f).

The current VTT and VTTR recommended in TAG were estimated in the most recent national UK study, carried out by a consortium of ARUP, ITS Leeds and Accent (ARUP et al., 2015a). Whilst these values are still considered robust and appropriate for the analysis of transport schemes, travel patterns and behaviours continuously change, as do the demands and techniques of transport modelling and appraisal.

This report presents the outcomes of a scoping study commissioned by the Department for Transport (DfT) into updating the VTT and the VTTR. This scoping study is considered as a first step to help establish the key theoretical and practical considerations, areas of advancement since the last study, and to provide recommendations with which to progress to a full primary study. We structure this report in two parts. Part 1 considers a review of methods and evidence, whereas Part 2 scopes out research option and recommends a research programme leading to a potential new national VTT study.

### 1.1 Part 1: Review of methods and evidence

The invitation to tender for this study set out the requirements for a review of methods and evidence, considering '*the state of the art*' with regards to surveying and estimation methods used in the valuation of VTT, drawing out developments in methods since the last national study in 2014/15. It was noted that this scoping study should focus on the estimation of passenger VTT and VTTR since a recent study on estimating the VTT and VTTR for road freight has recently been completed by ARUP et al. (2023).

Besides reviewing the development of methods and evidence of VTT and VTTR research to date, with a particular focus on national studies in the UK and internationally, the DfT set out eight key areas of interest for consideration. These areas of interest are set out in **Table 1**, including a reference to the chapters in which these topics are covered.

*Table 1: Key areas of interest set out in the invitation to tender by the DfT*

| <b>Key area of interest</b> | <b>Topic</b>                                          | <b>Covered in chapter</b> |
|-----------------------------|-------------------------------------------------------|---------------------------|
| KAI1                        | <i>Worthwhile (productive) travel time</i>            | 5                         |
| KAI2                        | <i>Willingness to Pay (WTP) estimation techniques</i> | 3                         |
| KAI3                        | <i>Modal effects and disaggregation</i>               | 4, 7, 8                   |
| KAI4                        | <i>Sampling and representation</i>                    | 3                         |
| KAI5                        | <i>User group and trip purpose segmentation</i>       | 4                         |
| KAI6                        | <i>Pandemic effects</i>                               | 2                         |
| KAI7                        | <i>Congestion and Reliability</i>                     | 6                         |
| KAI8                        | <i>Income, wealth, and distributional effects</i>     | 4                         |



In response to this specification, we brought together a consortium with a broad range of expertise in the context of VTT and wider transport modelling related research and set out a work programme. This report largely follows the structure of the work programme addressing closely linked key areas of interest in the same chapters.

**Chapter 2** summarises the history of national UK VTT studies, including the 2014/15 study, and international national VTT studies in Germany, Singapore, Norway, and the Netherlands that have taken place since 2014/15. Throughout the report the international developments are positioned against the 2014/15 study, the 2017/18 feasibility study on updating the VTT (ARUP and ITS, 2017, ARUP and ITS, 2018), and any unresolved topics in UK and international practice. The second half of **Chapter 2** reflects on societal developments which have taken place since the 2014/15 study, including the impact of the pandemic (KAI6). From our point of view, the relevant question related to the pandemic is whether travel patterns have sufficiently stabilised to commission a new study. Although the pandemic has had significant impacts on travel patterns as such, and potentially the VTT and the VTTR, we do not consider it to have posed challenges to the way in which VTT and VTTR related data are collected and analysed.

**Chapter 3** reviews developments in surveying and estimation methods and thereby covers KAIs 2 and 4. More specifically, the chapter reviews the potential for relying (more) on Revealed Preference (RP) data relative to the current dominant application of Stated Preference (SP) data; recognises that some form of SP will be included in a new study and therefore reviews the state-of-the-art of VTT SP surveys in terms of their presentation and experimental designs; reviews international best practice with established and novel recruitment and sampling methods; and evaluates how the approach to data analysis has changed since the 2014/15 study which introduced significant methodological advances (see Hess et al., 2017a).

**Chapter 4** focuses on the segmentation of the VTT and thereby covers KAIs 3,5 and 8. One of the main features discussed is the extent to which a new study can deal with the separation of mode- and user-effects. In the 2014/15 study there was insufficient confidence that for non-business travel differences in the VTT across modes could entirely be attributed to differences in comfort across modes of transport due to potential issues of self-selection. The extent to which the VTT should be further disaggregated by trip and user characteristics are discussed in terms of the empirical feasibility, but also reflected on in terms of the implications for transport modelling efforts which may need to deal with these levels of segmentation in assignment. One specific user characteristic which receives specific attention is the treatment of income, which plays an important role in explaining the variation in (behavioural) VTT in the population, in appraisal and distributional implications, and uprating the VTT over time.

**Chapter 5** covers KAI1 and reviews the literature on the worthwhile (productive) use of travel time. The first half of **Chapter 5** covers the literature on the business value of travel time (VBTT) which has always been a specific area of interest because the benefits of time savings, and the productive use of travel time, can (partly) be attributed to the employer and not just the employee. The second half of the chapter adopts a broader perspective and analyses how the worthwhile use of time can be measured and is translated in potentially lower VTTs across different modes of transport and journey purposes other than business. This includes a review on the literature estimating the VTT for autonomous vehicles which are typically claimed to have the potential to generate such benefits in the future.

**Chapter 6** covers KAI7 and brings together the recent methodological developments and evidence in relation to congestion, reliability, and crowding. This chapter particularly draws on follow-up research

commissioned by the DfT following the 2014/15 into the area of congested values of travel time (CVTT) and reflects on the implications for modelling from the chosen metrics to monetise the referred quality factors.

Finally, **Chapters 7 and 8** reflect on the application of VTT to active modes of travel and aviation respectively. As described in ARUP et al. (2015b), the estimation of a VTT for walking and cycling was challenging in the previous study due to these modes not being directly associated with a monetary cost component which lends itself to study (within-mode) time vs cost trade-offs central to the notion of VTT. Attention is paid to the extent to which emerging modes of travel, such as e-bikes and e-scooters, offer new opportunities to overcome these challenges. Aviation was not part of the 2014/15 study, but was explicitly covered in the ITT and different elements of travel time, such as access, wait times, and in-vehicle time, are commonly covered in other national VTT studies and these are reviewed in **Chapter 8**.

## 1.2 Part 2: Scoping of research options and recommended research programme

Following from the review of methods and evidence described in the first part of this scoping study, the next stage of the project involved consolidating observations from the review and scoping up possible approaches to a potential study for updating recommended values of travel time in the UK. A workshop was organised at the Department for Transport on the 14th of November 2024 involving the study team, the Department for Transport, and other stakeholders. The objective of the workshop was to discuss the results from the review of methods and evidence and identify requirements for a future study. The topics covered ranged from relevant policy applications to essential areas of improvement, and more exploratory options to be included in an upcoming study. The outcomes of the workshop are reflected in **Chapters 9 - 14**.

**Chapter 9** argues that with an evidence base of nearly ten years old, it is about the right time to commission a new national VTT study, and that this can build on the progress made by the ARUP et al. (2015a) study. **Chapter 10** continues by evaluating the design of the ARUP et al. (2015a) and more importantly highlight areas of attention where there is specific scope for improvement when conducting a new study. These suggestions range from the design of the SP experiments to recruitment methods, data analysis, and the translation of behavioural VTTs into appraisal VTTs. In relation to the appraisal VTT, **Chapter 11** concludes that the present level of segmentation by mode and journey purpose is still fit for purpose, and that any further level of segmentation by, for example distance or income, has implications for the study design and is associated with practical challenges for modelling and appraisal. **Chapter 12** highlights four specific areas which are labelled as 'essential' and in need of specific consideration in a new national VTT study. These include the separation of mode- and user-effects, the use of revealed preferences, and deriving a VTT for active travel and aviation, respectively. Finally, **Chapter 13** presents two exploratory areas of research in relation to the relationship between congestion and reliability, and the use of long-term values of travel time. **Chapter 14** summarises the set of recommendation and sets out a high-level timeline for a new national VTT study.

## 2. History of UK VTT studies, recent international VTT studies, and societal developments

### 2.1 History of UK VTT studies

The history of national VTT studies in the UK is well-documented and a detailed overview of the work conducted prior to the latest national VTT study in 2014/15 is documented in its associated scoping study (ITS, 2010). The first two UK national VTT studies are therefore briefly summarised below alongside a more detailed description of the latest study and related subsequent work.

The early work on the value of travel time savings in the UK originates from a series of research studies during the 1960s, the results of which were adopted and synthesised by the Department. Noteworthy, this work was largely based on revealed preference (RP) studies of mode choice for the journey to work. The first UK national value of time study by MVA et al. (1987) in the 1980s was the first VTT study to rely substantially on stated preference (SP) data, presenting respondents with mode and route choice situations enabling a wider empirical base to also include other leisure purposes in the analysis. The second UK national value of time study (Accent and Hague Consulting Group, 1999) dates from 1994 and was further analysed by ITS Leeds in 2003 (Mackie et al., 2003). The study solely focussed on car travel for non-work purposes and did not consider other modes and trip purposes. Again, SP data formed the basis for most of the analysis with the primary results coming from trade-offs between time and cost on route choices. The outcomes of the study provided updated guidance for the Department's Transport Analysis Guidance (TAG) in 2003.

These values were subsequently revised and updated by the Department to reflect changes in incomes and travel patterns over time. With an evidence base dating more than 20-years back, the Department commissioned the third national VTT study, the scope of which was larger and more complex than many previous national VTT studies. In particular, the ARUP et al. (2015a) study covered non-work and business travel purposes, several transport modes (car, bus, train, and other public transport), and covered not only travel time savings, but also travel time reliability, and several aspects of quality and comfort, including crowding. The design of the study was largely informed by a scoping study on updating the values for non-work travel (ITS, 2010), the ITS (2013) review of methods for estimating the VTT for business travellers, and related meta-analysis work by Wardman et al. (2015). Batley et al. (2019) provides an overview of the ARUP et al. (2015a) study design and its related outcomes.

### 2.2 Overview of the UK 2014/15 study

The third national VTT study took place over a period of 11 months and comprised two phases. Phase 1 of the work was undertaken between June and September 2014 and involved the designing and testing of the study in preparation for the actual market research. The data collection, including a large-scale field study, modelling and reporting took place between October 2014 and May 2015. The results of the study are published in full in ARUP et al. (2015a).

The scope of the work included business and non-work travel, but excluded freight, and aimed at identifying segmentations of the VTT by relevant factors including mode of transport, journey purpose, traveller characteristics such as income, and trip characteristics such as time and cost. To this end, four modes of transport were analysed, respectively car, bus, rail and 'other PT'<sup>1</sup>. Despite efforts to include active modes of travel (e.g. walking and cycling), challenges were encountered during the study which are discussed in more detail in **Chapter 7** of this report and accordingly only tentative recommendations were provided. Across the study the main form of data collection was that of Stated

---

<sup>1</sup> 'other PT' refers to other public transport and represents trams, light rail and the London Underground.

Preference (SP) studies, but for business travel there was also an attempt to validate with Revealed Preference (RP) data.

Business travel followed a specific study design relative to non-work travel and elicited values of business travel time savings using willingness-to-pay (WTP) methods. The decision was based on the earlier scoping studies by ITS (2013) and replaced the Cost-Savings Approach (CSA) traditionally used by the Department. **Chapter 5** provides a more in-depth review of the alternative approaches to valuing business time, including the use of the related ‘Hensher Equation’ which enables capturing the productive use of travel time.

The fieldwork took place between October and December 2014 during which just over 8,500 SP interviews were completed. These were segmented by eleven mode-purpose combinations (car/bus/rail/other PT vs commute/other non-work / employee’s business) noting that employee’s business trips by bus were considered out of scope. 89% of the interviews were collected using intercept-recruitment and 11% were Computer Assisted Telephone Interviews (CATI) and most surveys were completed online. **Table 2** summarises the number of completed interviews by mode-purpose combination.

*Table 2: Total complete SP interviews (online and CATI) by mode and purpose – adapted from Batley et al. (2019)*

| <b>Mode</b>  | <b>Commute</b> | <b>Other Non-work</b> | <b>Employee’s business</b> | <b>Total</b> |
|--------------|----------------|-----------------------|----------------------------|--------------|
| Car          | 1,032          | 1,037                 | 956                        | 3,025        |
| Bus          | 371            | 672                   | N/A                        | 1,043        |
| Rail         | 998            | 1,128                 | 1,010                      | 3,136        |
| Other PT     | 614            | 540                   | 265                        | 1,419        |
| <b>Total</b> | <b>3,015</b>   | <b>3,377</b>          | <b>2,231</b>               | <b>8,623</b> |

The core section of each SP survey comprised three ‘SP games’. Each game comprised five choice tasks such that each respondent made fifteen choices overall, each choice reflecting a trade-off between travel time and travel costs, and other relevant journey characteristics (attributes) as described by **Table 3**. The presented choices were all presented in the context of a reference trip, and presented levels of travel time and costs were pivoted around these reference levels. Screenshots of the different SP games can be found in Batley et al. (2019). Specific design considerations are discussed in more detail in **Section 3.2** of this report.

Table 3: Description of SP format by game and mode used in ARUP et al. (2015a)

| Game and mode  | Description of SP format                                                                                                                                                                                                                                                                                                                                                                                                     |
|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| SP1            | SP1 used a generic format across all modes, presenting respondents with an ‘abstract’ choice between two options described only on the basis of travel time and travel cost, where one option was cheaper, but the other option was faster.                                                                                                                                                                                  |
| SP2            | SP2 also presented respondents with an abstract binary choice, still focussing on travel cost and travel time but where, for travel time, five different typical trip outcomes were presented for each alternative as a representation of travel time variability.                                                                                                                                                           |
| SP3            | SP3 used somewhat different presentations across modes, whilst nevertheless retaining an abstract binary choice context, as described for each mode below.                                                                                                                                                                                                                                                                   |
| SP3 car        | For car, the two options were described in terms of travel cost for each trip and the amount of time that each trip spends in three types of driving conditions (free-flow, light traffic, heavy traffic).                                                                                                                                                                                                                   |
| SP3 rail       | For rail, two different experiments were used: <ul style="list-style-type: none"> <li>a) For the first group, we presented a choice similar to SP1, with the difference that for each alternative we additionally defined the level of crowding applying to the trip.</li> <li>b) For the second group, we presented a choice between up to three operators, described in terms of travel time, fare and headway.</li> </ul> |
| SP3 bus        | For bus, two different experiments were also used. <ul style="list-style-type: none"> <li>a) For the first group, we presented a crowding game analogous to the rail game, albeit with different crowding definitions.</li> <li>b) For the second group, we presented a choice between two bus routes described in terms of free-flow time, slowed down time, dwell time, headway and fare.</li> </ul>                       |
| SP3 ‘other PT’ | For ‘other PT’, two different experiments were again used. For the first group, we presented a crowding game analogous to the bus game. For the second group, we presented a mode choice game (‘other PT’ against either bus or rail) using time, headway and cost as attributes.                                                                                                                                            |

The survey responses formed the basis for the modelling efforts and ultimately the translation into recommended values of travel time and reliability to the Department. In terms of estimating so-called ‘behavioural models’, which explain the variation in the VTT for different modes of transport and journey purposes, several state-of-the-art features were tested and included in the final model specifications (Hess et al., 2017a). These features include:

- *Multiplicative error terms* – accounting for the fact that error variance increases when utility increases, for example with longer distance journeys (Fosgerau and Bierlaire, 2009).
- *Reference dependence* – all behavioural models considered changes in travel time and cost relative to the trip on which respondents were intercepted. Interest was paid to the ‘size’ and ‘sign’ of these changes, with the assumption that respondents prefer positive changes (time and cost savings) over losses (increases in time and cost), i.e. so-called loss aversion; and non-linearities in the impacts of gains and losses. The latter enables the presence of behavioural effects like cost damping. Implementation was based on the principles of De Borger and Fosgerau (2008).
- *Valuation space* – As opposed to modelling utility functions from which values of travel time can be derived post estimation, modelling in valuation space enables direct estimation of the VTT and its variations (Ojeda-Cabral et al., 2016).
- *Random taste heterogeneity* – any unexplained variation in the VTT, after controlling for trip and respondent characteristics, is accommodated by a random parameter specification on

the VTT parameter using the log-uniform distribution which limits issues with the tail of the distribution identified in other studies (e.g. Börjesson et al., 2012b).

- *Joint modelling across SP games*, to improve identification of variation in the VTT, the analysis combined the data from SP1-3, but the VTT was allowed to vary across SP games.

The eleven mode-purpose specific models describing the functional form of the '*behavioural VTT*' were taken forward to generate recommended '*appraisal values*'. To this end, a sample enumeration approach was adopted using the National Transport Survey (NTS) over the years 2010-2102. In simple terms, the behavioural models allow calculating the expected VTT for each trip included in the NTS using the relevant covariates identified in the modelling process (e.g. travel time, cost, income, gender etc.). Having a measure of the '*behavioural VTT*' for each trip in the NTS then allowed deriving a weighted national average VTT measure, a process called sample enumeration. To determine the weight associated with each trip two factors were used. The first weighting component concerns the weights (or expansion factors) included in the NTS enabling achieving a representative picture of national travel. The second weighting component is distance, assigning a higher weight to longer distance journeys based on these journeys having a larger probability experiencing the benefits of transport investments reducing travel times (Mackie et al., 2003).

In translating the behavioural VTTs into appraisal VTTs some important considerations were made. Firstly, the values taken forward for policy purposes were based on SP1 (i.e. simple time-cost trade-offs) and assumed a change in travel time of  $\Delta t = 10$  minutes<sup>2</sup>. This arbitrarily chosen value for  $\Delta t$  is consistent with values implemented in preceding international VTT studies (Börjesson and Eliasson, 2014). Secondly, all non-work trips appraisal values were not segmented by mode because some of the identified differences in the VTT across modes may have been the result of self-selectivity instead of pure modal preferences. This was not considered to be an issue for business travel. Thirdly, business values were based on employees reporting their employer's willingness-to-pay for time savings. Fourthly, it was recommended to further segment the VTT for employee's business by distance band highlighting an increase in the VTT in this dimension. **Table 4** presents the recommended appraisal values by ARUP et al. (2015a), which have ultimately been translated into TAG Table A.1.3.1 using DfT's own distance band calculations (Department for Transport, 2024f).

---

<sup>2</sup> This assumption is needed due to the presence of reference dependence in the behavioural VTT model.

Table 4: Reproduction of Table 7.28 of ARUP et al. (2015a)

| Source             | Distance        | Commute   | Other non-work | Employees' business |       |           |            |       |
|--------------------|-----------------|-----------|----------------|---------------------|-------|-----------|------------|-------|
|                    |                 | All modes | All modes      | All modes           | Car   | Bus       | 'Other PT' | Rail  |
| WebTAG             | All             | 7.62      | 6.77           | 25.47               | 24.43 | 15.6<br>4 | 24.72      | 30.07 |
| Re-surveyed values | All             | 11.21     | 5.12           | 18.23               | 16.74 | N/A       | 8.33       | 27.61 |
|                    | <20 miles       | 8.27      | 3.62           | 8.31                | 8.21  | N/A       |            | 10.11 |
|                    | 20 to 100 miles | 12.15     | 6.49           | 16.05               | 15.85 | N/A       |            | 28.99 |
|                    | >= 100 miles    |           | 9.27           | 28.62               | 25.74 | N/A       |            |       |

Notes: Distance weighted, 'all distance' values based on income option 1, for distance-banded values non-work based on income option 2 (household income = £49,684) and business on income option 1, VTT imputed for PT trips with zero cost, SP1 VTTs,  $\Delta t=10$ , employers paying for EB trips, Tool version 1.1.

### 2.3 Follow-up work following ARUP et al. (2015a)

The recommendations of the 2014/15 study were peer reviewed and audited by a consortium involving SYSTRA, Imperial College London, and the Danish Technical University (SYSTRA et al., 2015) before being implemented into TAG Unit A1-3 in 2017 (Department for Transport, 2022a).

In 2018 ARUP and ITS Leeds conducted a feasibility study setting out a 'programme for maintaining a robust valuation of travel time savings'. The Phase 1 report identified a potential structure for the maintenance programme, which has the overall aim of keeping the VTT fit-for-purpose over longer periods of time (ARUP and ITS, 2017). Four streams were identified building on i) regular updates of the VTT; ii) monitoring of changes in the VTT; iii) improving the robustness of base and future VTT values and iv) commissioning work on outstanding issues. The Phase 2 report primarily concerns the first two streams as together these formed the core of a proposed maintenance programme and place confidence in the derived 2014/15 behavioural models (ARUP and ITS, 2018). Two prominent proposals from the Phase 2 report were to i) validate the results from the current VTT uprating procedure by GDP growth using meta-analysis; and ii) collecting small SP samples on an annual basis using exact repeats of the 2014/15 study to pick up changes in the VTT over time quicker than by collecting a large sample every 10 to 20 years through a new national UK VTT study. The current scoping study highlights the Department's decision to continue the tradition of updating the VTT using a large-scale study.

ARUP and ITS (2018) also concluded that despite promising developments in the context of big and emerging data sources, the use of RP data for estimating the VTT were still associated with practical limitations at the time. This is a topic which is revisited in **Section 3.1** of this report.

A separate body of work has been taking place exploring the issue of congested values of time (WSP et al., 2018, WSP and Mott MacDonald, 2019, ITS, 2020, WSP and Mott MacDonald, 2022, ITS, 2022). The latter two reports form the culmination of this stream of work, where WSP and Mott MacDonald (2022) investigates the use of congested values of travel time (CVTT) in model forecasting and scheme appraisal. It develops a 'proof of concept' using the policy-responsive integrated strategy model (PRISM) of the West Midlands Combined Authority; and ITS (2022) re-analyses the 2014 to 2015 UK values of travel time (VTT) study to address concerns around the robustness of the UK congestion multipliers estimated at the time. In doing so, the report analyses how CVTT values differ by journey

purpose and length; reviews the approaches to deriving UK 'headline VTT'; and considers the need for, and possible approaches to, a new CVTT study where respondents are re-surveyed and values are updated. This body of literature is discussed in more detail in **Chapter 6** of this report.

## 2.4 Recent national VTT studies outside of the UK

In this subsection an overview of four recent official national VTT studies that have taken place outside of the UK is provided. These are respectively the German, Singapore, Norwegian and Dutch national VTT study. Switzerland has recently completed a new national VTT study (ETH, 2024). The results of the study were, however, not available in time to be included in this report. The design of the latest Swiss study is, however, very close to that of the German study and its approach to analysis closely resembles that of the latest UK VTT study.

### 2.4.1 Overview of the German 2012/13 study

Germany conducted its first national VTT study in between 2012-2014 (Axhausen et al., 2014). The purpose of the study was to provide inputs for cost-benefit analysis as part of Germany's Federal Ministry of Transport and Digital Infrastructure (BMVI) new Federal Transport Infrastructure Plan (BVWP) 2030. A detailed description of the survey design and data collection process can be found in Dubernet and Axhausen (2020), and a detailed analysis of the reliability part of the study in Ehreke et al. (2015).

The design of the German VTT (and value of reliability) study builds on the experience of earlier studies conducted in Switzerland where, compared to international practice, more complex SP experiments including multiple modes and multiple elements of the generalized costs of travel are employed in a series of overlapping choice contexts (Axhausen et al., 2004, Axhausen et al., 2008, Weis et al., 2011, Fröhlich et al., 2013). Most significantly a distinction between short- and long-term travel choices is made as discussed in more detail below.

In terms of its segmentation, the German VTT study covered a wider set of journey purposes, respectively education, commuting, shopping, business and leisure; and several modes of transport, including car, public transport, airplane, cycling and walking (with the presented modes depending on the distance class).

The data collection was carried out between 2012 and 2013. Around 4,000 RP interviews were collected by phone (CATI): phase 1. In phase 2, more than 3,000 SP interviews were carried out. Respondents of the business sample were recruited by an online access panel. The non-business sample was drawn from a dual frame of landline and mobile numbers (60% and 40%). The RP data were collected over the phone, and they were asked to participate in a follow-up SP survey.

The RP component requested respondents to report three recent trips: commuting to work, and the trips to the most important shopping and leisure (< 50 km) destinations. Also, information on the last long-distance trip over 50 km distance was collected, where, if the latter was ground-based, data on the most recent air trip was also collected. From these reported trips, one recent trip was randomly selected as the reference trip and information regarding non-chosen alternatives was externally inferred. In assigning the reference trip a slight bias was implemented to have a sufficient representation of long-distance trips during data collection. The most recent trip became the reference in the business sample.

Respondents to the RP phase 1 of the study received the SP survey within two weeks of having participated. The SP survey itself was carried either by pencil and paper or online. In common with the UK VTT study, the SP part of the data collection was constructed around the reference trip.



Table 5 provides an overview of the types of choices and attributes presented in the short-term part of the SP study. Three different types of choices were presented, respectively being mode, route, and departure time choice. Most notably, a significantly larger number of attributes is included in the choice settings relative to the other national VTT studies reviewed here. Each SP setting was associated with eight choice tasks and in the non-business sample respondents were presented with all three SP settings, where respondents to the business trips were only presented with two out of three SP settings and these respondents only had to answer sixteen questions in total.

In terms of the long-term SP questions, either eight workplace or residential location choices were presented to non-business respondents only as per **Table 6**. In the workplace games respondents were asked to choose between their current workplace and an alternative workplace that varied in commute times, commute costs, salary, and other workplace attributes. In the residential location games, the current location was traded-off against a new location varying in travel cost and time for commute trips, the time and cost for car and public transport to the nearest shopping location, and differences in residential attributes such as appearance and location of the dwelling.

The results from the long-term SP questions were not taken forward in deriving recommended values for policy appraisal. Only a brief overview of these results is presented in Sections 10.1.2 and 10.1.3 of Axhausen et al. (2014), where the residential location choice results in lower VTT estimates relative to the short term mode choice situations, and the estimates for job location suffer from non-significant parameter estimates for travel cost and travel time limiting the identification of the VTT in this context. A more detailed analysis of the same data is presented in Dubernet et al. (2020) also indicating towards challenges in separating the role of travel costs and other monetary components such as income in this setting.

The emerging VTT estimates are based on models estimated jointly on the three short-term SP experiments and the RP data. These are multinomial logit utility functions, with non-linear influences for transport time and costs and interactions with income. Sign/size effects of the time and cost differences were not included in the models. For business trips, a cost-savings approach, the Hensher equation and the willingness to pay (WTP) method were considered, and the results from the WTP method were selected for implementation.

Translating the obtained behavioural models of the VTT followed a different process relative to the most recent UK national VTT study. In this case, the estimated behavioural models are applied to the sample of respondents, *not the German equivalent of the NTS*, to calculate their personal VTT. Dubernet and Axhausen (2020, p. 1499) discuss three specific weights to achieve representativeness which account for respectively population characteristics, trip characteristics (e.g. trip frequency and trip length, mode and trip purpose), and business trips. The application of these weights is, unfortunately, not clearly described in Axhausen et al. (2014), but it can only be assumed that a combination of the first two weight elements, whilst ignoring mode and trip purpose, are applied to obtain mode-purpose specific VTT estimates. In short, a combination of population and trip weights is included in the sample enumeration here. To our knowledge distance weighting, as applied in the sample enumeration process in the UK, is not implemented in the German national VTT study.

Table 5: Overview of the short-term experimental study designs used in Dubernet and Axhausen (2020)

| Attribute                            | Attribute levels                                                              | Alternative |      |     |                  |       |       |
|--------------------------------------|-------------------------------------------------------------------------------|-------------|------|-----|------------------|-------|-------|
|                                      |                                                                               | Walk        | Bike | Car | Public transport | Coach | Plane |
| Mode choice (SP1)                    |                                                                               |             |      |     |                  |       |       |
| Travel time                          | -30%, -10%, +20% of current state                                             | x           | x    | x   | x                | x     | x     |
| Access/egress time                   | 5%, 10%, 20% of travel time                                                   | -           | -    | x   | x                | x     | x     |
| Congestion/waiting time              | 5%, 10%, 20% of travel time                                                   | -           | -    | x   | x                | x     | x     |
| Travel cost                          | -20%, +10%, +30% of current state                                             | -           | -    | x   | x                | x     | x     |
| Monthly travel cost                  | Travel cost (to and from destination) * trip frequency from RP data (rounded) | -           | -    | x   | x                | x     | x     |
| Transfers                            | -1, ±0, +1 time                                                               | -           | -    | -   | x                | x     | x     |
| Headway                              | -1, ±0, +1 step                                                               | -           | -    | -   | x                | x     | x     |
| Share delayed trips                  | 5%, 10%, 20%                                                                  | -           | -    | x   | x                | x     | x     |
| Route choice (SP2)                   |                                                                               |             |      |     |                  |       |       |
| Travel time                          | -30%, -10%, +20% of current state                                             | -           | -    | x   | x                | -     | -     |
| Access/egress time                   | 5%, 10%, 20% of travel time                                                   | -           | -    | x   | x                | -     | -     |
| Congestion/waiting time              | 5%, 10%, 20% of travel time                                                   | -           | -    | x   | x                | -     | -     |
| Travel cost                          | -20%, +10%, +30% of current state                                             | -           | -    | x   | x                | -     | -     |
| Transfers                            | -1, ±0, +1 time                                                               | -           | -    | -   | x                | -     | -     |
| Crowding                             | Low, medium, high                                                             | -           | -    | -   | x                | -     | -     |
| Delay every x. trip                  | 5%, 10%, 20%                                                                  | -           | -    | x   | x                | -     | -     |
| Departure time and reliability (SP3) |                                                                               |             |      |     |                  |       |       |
| Travel time                          | -30%, -10%, +20% of current state                                             | -           | -    | x   | x                | -     | -     |
| Access/egress time                   | 5%, 10%, 20% of travel time                                                   | -           | -    | x   | x                | -     | -     |
| Congestion/waiting time              | 5%, 10%, 20% of travel time                                                   | -           | -    | x   | x                | -     | -     |
| Travel cost                          | -20%, +10%, +30% of current state                                             | -           | -    | x   | x                | -     | -     |
| Transfers                            | -1, ±0, +1 time                                                               | -           | -    | -   | x                | -     | -     |
| Share arriving early                 | 5%, 10%, 20%                                                                  | -           | -    | x   | x                | -     | -     |
| Share arriving on time               | 100%-share early-share delayed                                                | -           | -    | x   | x                | -     | -     |
| Share arriving delayed               | 10%, 20%, 40%                                                                 | -           | -    | x   | x                | -     | -     |
| Time arriving early                  | 5%, 15%, 25% of travel time                                                   | -           | -    | x   | x                | -     | -     |
| Time arriving late                   | 10%, 20%, 30% of travel time                                                  | -           | -    | x   | x                | -     | -     |

Table 6: Overview of the long-term experimental study designs used in Dubernet and Axhausen (2020)

| Attribute (current alternative (RP)) | Unit  | Attribute levels (new alternative (SP)) | Alternative |     |
|--------------------------------------|-------|-----------------------------------------|-------------|-----|
|                                      |       |                                         | Current     | New |
| Workplace choice (SP4)               |       |                                         |             |     |
| Car commute time                     | (min) | -30%, -10%, +20% of current state       | x           | x   |

|                                          |                        |                                    |    |   |
|------------------------------------------|------------------------|------------------------------------|----|---|
| Car commute cost                         | (€/month)              | -20%, +10%, +30% of current state  | x  | x |
| Public transport commute time            | (min)                  | -30%, -10%, +20% of current state  | x  | x |
| Public transport commute cost            | (€/month)              | -20%, +10%, +30% of current state  | x  | x |
| Salary before tax                        | (€/month)              | -10%, ±0%, +10% of current salary  | x  | x |
| Staff managed                            | (Number)               | -50%, +20%, +100% of current state | x  | x |
| Budget managed                           | (Million €/year)       | -50%, +20%, +100% of current state | x  | x |
| Change of industry needed                | (Yes/no)               | No, yes                            | No | x |
| Change of company needed                 | (Yes/no)               | No, yes                            | No | x |
| <i>Residential location choice (SP5)</i> |                        |                                    |    |   |
| Type                                     | (House/apartment)      | House, apartment                   | x  | x |
| Size                                     | (m2)                   | -20%, +10%, +30% of current state  | x  | x |
| Standard                                 | (New/renovated/old)    | New, renovated, old                | x  | x |
| Exterior                                 | (None/garden/balcony)  | None, garden, balcony              | x  | x |
| Rent/mortgage                            | (€/month)              | -20%, +10%, +30% of current state  | x  | x |
| Area                                     | (Urban/suburban/rural) | Urban, suburban, rural             | x  | x |
| <i>Car travel time</i>                   |                        |                                    |    |   |
| Commute                                  | (min)                  | -30%, -10%, +20% of current state  | x  | x |
| Shopping                                 | (min)                  | -30%, -10%, +20% of current state  | x  | x |
| <i>Car travel costs</i>                  |                        |                                    |    |   |
| Commute                                  | (€/month)              | -20%, +10%, +30% of current state  | x  | x |
| Shopping                                 | (€/month)              | -20%, +10%, +30% of current state  | x  | x |
| <i>Public transport travel time</i>      |                        |                                    |    |   |
| Commute                                  | (min)                  | -30%, -10%, +20% of current state  | x  | x |
| Shopping                                 | (min)                  | -30%, -10%, +20% of current state  | x  | x |
| <i>Public transport travel costs</i>     |                        |                                    |    |   |
| Commute                                  | (€/month)              | -20%, +10%, +30% of current state  | x  | x |
| Shopping                                 | (€/month)              | -20%, +10%, +30% of current state  | x  | x |

Altogether this results in the following set of recommended values for the main motorised modes of transport (see **Table 7**). A more detailed segmentation by distance is available for journeys shorter than 50km.

*Table 7: Recommended values by Axhausen et al. (2014) in 2012 €/hr*

| <b>Journey purpose</b> | <b>Car</b> | <b>Public transport</b> | <b>Air</b> | <b>All</b> |
|------------------------|------------|-------------------------|------------|------------|
| Education              | 3.80       | 4.28                    |            | 4.15       |
| Shopping               | 4.31       | 5.15                    |            | 4.65       |
| Business               | 8.40       | 7.03                    | 36.82      | 8.53       |
| Leisure                | 4.75       | 4.35                    |            | 4.68       |
| Commute                | 4.00       | 4.33                    | 24.88      | 4.32       |

#### 2.4.2 Overview of the Singapore 2015 study

Hess et al. (2017b) report on Singapore's most recent VTT study conducted in 2015 by the Land Transport Authority (LTA). The LTA is the government agency tasked with the development and regulation of Singapore's land transport system. The study updates the preceding VTT values derived in 2008.

The scope of Singapore's VTT study included eight different modes of transport, respectively being car, motorcycles, mass rapid transit (MRT), bus, taxi, walking, and cycling. Unlike the other national VTT studies reviewed here, the sampling strategy did not specifically target specific journey purposes; journey purposes were part of the analysis but are not used in obtaining mode specific VTT estimates. Ultimately, VTTs are derived for cars, motorcycles, MRT, bus and taxis and walking but not for cycling.

Most interviews were computer-assisted personal interviews (CAPI) in respondents' homes and partially through intercept surveys and SP was the only source of data collected. The design of Singapore's SP study is distinct in that it does not include the traditional SP1 within-mode route choice format trading off travel time and cost. Instead, a separation is introduced between time- and accident-based SP games. For journeys by car, motorcycle, walking and cycling, the accident-based SP games introduced public safety programs reducing the number of accidents (of different types). These games did not involve trade-offs with travel time and hence are not covered in the remainder of this review. For the time-based SP games, different time components were specified (e.g. free-flow, light- and heavy-congestion), and in public transport crowding was also covered. Travel time reliability was not covered by the study design. Standard short term within mode route choices were presented for each mode. A detailed overview of the SP-games and the included attributes is presented in **Table 8** below, including the number of choice tasks per game. In line with the standards on experimental design, the levels for travel time and travel cost were pivoted around a recent reference trip to increase reliability of the design. Respondents were presented with up to 15 choice tasks each.

*Table 8: Overview of SP games in Singapore's VTT study (replicating Table 1 from Hess et al., 2017b)*

| Game | Description                      | Attributes                                                                                                                                                                                  |
|------|----------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CT1  | Car: congestion and costs        | Free-flow travel time, light congestion, heavy congestion, parking cost, petrol cost, ERP cost                                                                                              |
| CT2  | Car: parking choice              | Walking time, queuing time, search time, parking cost                                                                                                                                       |
| CA1  | Car: accidents                   | Fatalities, serious and minor injuries per year, change in annual tax burden                                                                                                                |
| MCT1 | Motorcycle: congestion and costs | Free-flow travel time, light congestion, heavy congestion, parking cost, petrol cost, ERP cost                                                                                              |
| MCA1 | Motorcycle: accidents            | Fatalities, serious and minor injuries per year, change in annual tax burden                                                                                                                |
| MT1  | MRT: time and crowding           | Walking time, waiting time, in-vehicle time in five crowding levels (3 seated, 2 standing), interchanges, fare                                                                              |
| MT2  | MRT: walking                     | Crossing type (at grade, uncovered bridge, covered bridge without lift, covered bridge with lift, air-conditioned underpass), covered and uncovered walking time to and from crossing, fare |
| BT1  | Bus: time and crowding           | Walking time, waiting time, in-vehicle time in five crowding levels (3 seated, 2 standing), interchanges, fare                                                                              |
| BT2  | Bus: excess waiting time         | Bus arrival times, fare                                                                                                                                                                     |
| TT1  | Taxi: access, time, and costs    | Walking time, waiting time, in-vehicle time, prebooked or on street, fare, booking fee                                                                                                      |
| PT1  | Pedestrian: walking environment  | Crossing type (at grade, uncovered bridge, covered bridge without lift, covered bridge with lift, air-conditioned underpass), covered and uncovered walking time to and from crossing, fare |
| PA1  | Pedestrian: accidents            | Fatalities, serious and minor injuries per year, change in annual tax burden                                                                                                                |
| CYA1 | Cycling: accidents               | Fatalities, serious and minor injuries per year, change in annual tax burden                                                                                                                |

Note: ERP=electronic road pricing

A sample of 5,000 households were selected for interviews, but unfortunately the study report is insufficiently clear on final sample sizes per segment and any associated data cleaning activities. The modelling approach is also distinct from other national VTT studies. Additive RUM models are estimated in WTP-space directly estimating the VTT. Bayesian estimation methods are applied introducing correlated random heterogeneity in the VTT components, which is not standard in other

national VTT studies. Instead of using interaction terms, posterior densities examine the variation in the VTT (and its components) in the sample. That is, individual-level posterior means are correlated with sample characteristics.<sup>3</sup> The posteriors thereby account for variation in the journey purpose (including commute, shopping, other non-work and business trips). For car and motorcycles no clear income effects could be identified in this process.

Multiple cost components were accounted for in the car and motorcycle models, but it was decided to rely on the VTT based on the sensitivity to electric road pricing (ERP) charges. Recommended values are presented by mode, but not by journey purpose, and are based on the overall sample means of the estimated correlated random coefficients (i.e. mixed logit models). As such, no weighting is applied to arrive at a national average VTT for a give mode-purpose combination. Distinctions in the VTT for car and motorcycles were identified based on traffic conditions (alike SP3 in the UK's 2014/15 study). A weighted value across road conditions is based on the share of travel conditions estimated from the 2012 (household travel survey) HTS model. When additionally weighted by daily mode share of cars and motorcycles provides an equity weighted VTT of \$16.39 per hr (2015 Singapore dollars). Note that distance weighting is not applied in this weighting process. **Table 9** and **Table 10** present the recommended values for different time components by car, motorcycle, bus and MRT.

*Table 9: Recommended values of the 2015 Singapore study for cars and motorcycles (replicating Table 10 from Hess et al., 2017b)*

| Variable (cents/min)                        | Cost  |            |
|---------------------------------------------|-------|------------|
|                                             | Car   | Motorcycle |
| Value of free-flow time versus ERP          | 34.32 | 21.20      |
| Value of light congestion time versus ERP   | 36.58 | 24.91      |
| Value of heavy congestion time versus ERP   | 46.38 | 24.92      |
| Value of weighted travel time versus ERP    | 36.80 | 23.02      |
| Value of walking time versus parking cost   | 36.39 | na         |
| Value of queuing time versus parking cost   | 32.62 | na         |
| Value of searching time versus parking cost | 40.02 | na         |

Note: na=not applicable

*Table 10 Recommended values of the 2015 Singapore study for bus and mass rapid transit (MRT) (replicating Table 12 from Hess et al., 2017b)*

| Variable (cents/min)                                  | Cost  |       |
|-------------------------------------------------------|-------|-------|
|                                                       | Bus   | MRT   |
| Value of walking time                                 | 15.96 | 22.83 |
| Value of waiting time                                 | 15.06 | 17.00 |
| Value of interchanges (cents/interchange)             | 40.82 | 68.16 |
| Value of in-vehicle time, seated, with empty seats    | 10.67 | 17.39 |
| Value of in-vehicle time, seated, quite packed        | 10.88 | 17.78 |
| Value of in-vehicle time, seated, completely packed   | 13.01 | 18.01 |
| Value of in-vehicle time, standing, quite packed      | 16.55 | 22.08 |
| Value of in-vehicle time, standing, completely packed | 17.01 | 24.50 |

<sup>3</sup> Note that this is not a recommended approach because individual-level posterior means are associated with a large degree of uncertainty and are drawn to the population mean. It is better to re-estimate the model using the identified explanatory variables of interest as interaction terms to obtain more accurate results.

|                                                          |       |       |
|----------------------------------------------------------|-------|-------|
| Value of in-vehicle time weighted by crowding conditions | 11.73 | 20.71 |
| Value of EWT                                             | 71.74 | na    |

### 2.4.3 Overview of the Norwegian 2018/19 study

In Norway, national VTT studies for passenger transport have been conducted in 1996-1997 (Ramjerdi et al., 1997), 2007-2009 (Samstad et al., 2010), and 2018-2020 (Flügel et al., 2020). The purpose of the 2018-2020 study was to estimate new unit value for economic appraisal of transport projects in Norway. It contains values of travel time and time-related factors.

The scope of the 2018-2020 Norwegian VTT study covered the following ten modes of transport: car driver, car passenger, ferry, bus, train, metro/tram/light rail, passenger boat, airplane, cycling and walking, and segments the VTT by three journey purposes, respectively being commute, business, and leisure. Moreover, the value of time is segmented into different components, including value of travel time and multipliers for headway, transfers, travel time variability, public transport, airport, and ferry access, waiting and delay time, comfort, and congestion. The method for business travel considers both the value for the employer and the employee and the opportunities for working while travelling.

The data collection was carried out in 2018 and in 2019 and only SP data were collected. Notably, the data collection was split into four separate efforts. The main survey covered the VTT and its core components such as reliability. The second survey considered crowding in public transport. The third survey considered station quality factors in public transport (e.g. seating, covering, tidiness). The fourth survey covered access and egress to the airport. Since the main values of time are based on the first survey effort, these will form the core of the review here. The other surveys are referred to where relevant in the remainder of this report.

Respondents were recruited partly from an internet panel (Norstat), partly from an alternative email register owned by the Postal service (Bring), and partly on-site (onboard public transport, at stops/stations or on the street). The largest survey in the autumn of 2018, covering the value of travel time and several other attributes, involved all three recruitment methods. Those recruited on site (intercept) could choose between answering straight away and answering later. The results show that recruited method matters for the value of travel time and suggest that internet panel members are not representative for the traveling population in this dimension. Due to not being representative for the travelling population, internet panel members are given a lower weight in the analysis and as a result higher values of time are obtained (Halse et al., 2023, Flügel et al., 2020).

**Figure 1** provides an overview of the different SP formats presented by mode of transport in the Norwegian VTT study. Alike most other national VTT studies, SP1 takes the within-mode route choice design trading off time and costs for all motorised forms of transport. This is then followed by SP2, adding additional attributes like congestion and reliability to the within mode route choice context. Note that from some respondents (SP2-CE2d) a mode choice component is included. The choice experiments were dynamically constructed by a pivotal experimental design, so that each respondent received realistic choice tasks relatable to their current travel context and based on attribute levels of a reference trip.

The modelling takes account of the heterogeneity in preferences and characteristics, by employing a large set of interaction factors based on personal, trip and geographic characteristics and by estimating a standard deviation for unobserved (random) heterogeneity. In common with the national UK VTT study, the VTT is estimated directly (i.e. random valuation) with multiplicative error terms (i.e. log

random valuation), and sign- and size-effects of the time and cost differences were included in the models.

The implementation of the estimated models follows the same approach as the German study, by applying the model to the interviewed sample and using weights from the 2018 national travel survey (RVU) to obtain a national average VTT measure for different VTT segmentations. **Table 11** reports the recommended values for the main motorised modes of transport. Most notably a segmentation is made by three distance segments (under 70km; 70-200km; and over 200km). Apart from business trips by rail we see that the VTT is increasing with distance. Additional recommended values are available for active travel and ferries.

Whilst the official values of travel time are only segmented by three distance segments, subsequent work by Flügel and Madslien (2020) develops a continuous functions describing the relationship between the value of travel time and travel distance for different combinations of transport models and travel purposes. Not unexpectedly, the implementation of such a continuous relationship between distance and the VTT highlights an increase in user benefits for transport projects targeting, on average, longer travel distances compared to the typical trip included in the RVU.

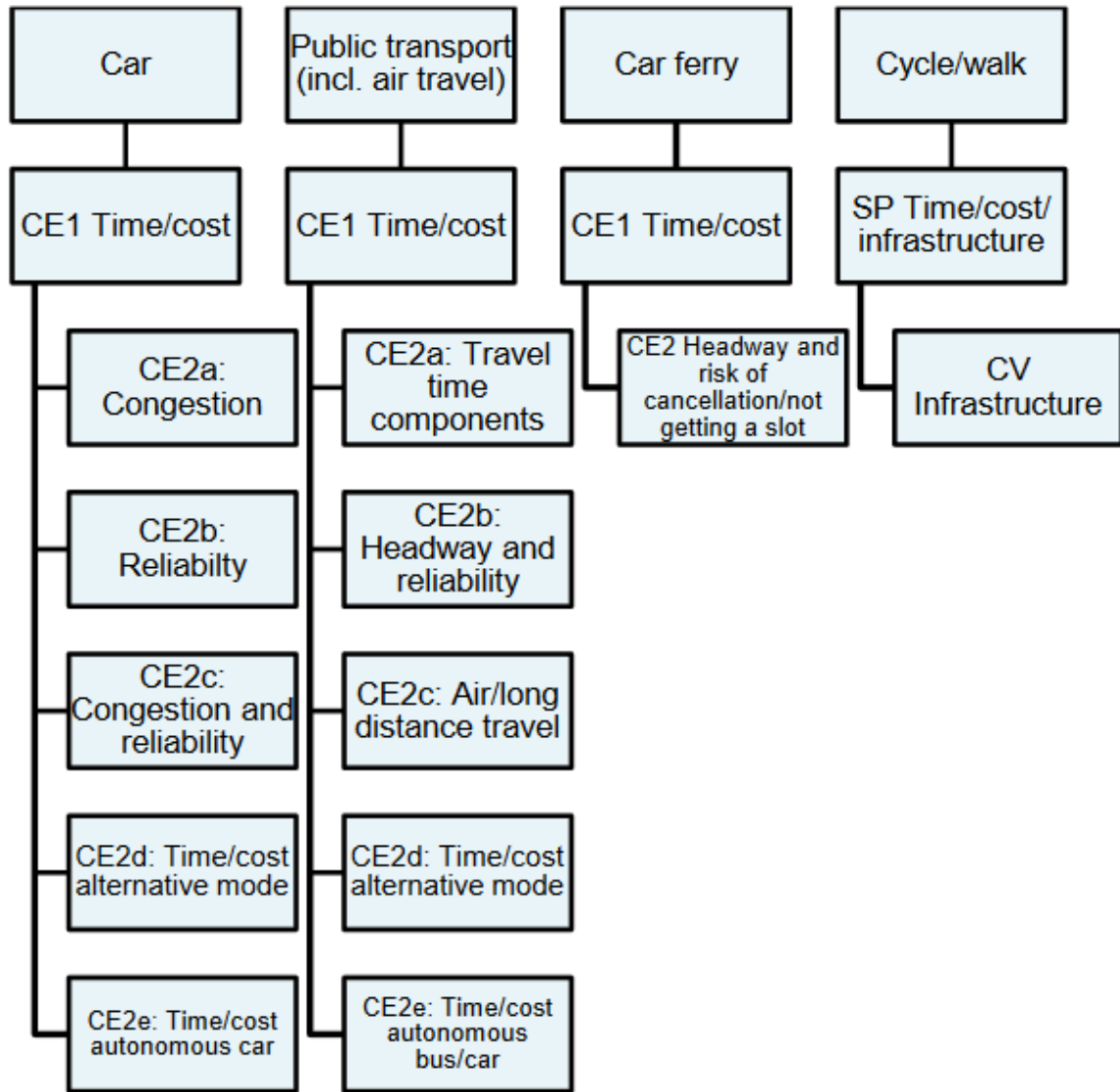


Figure 1: Overview of SP experiments in the Norwegian VTT study (Flügel et al., 2020)

Table 11: Overview of recommended values for the Norwegian study in 2018 NOK/hr (Flügel et al., 2020)

|                        |            | Business | Commuting | Leisure | All purposes |
|------------------------|------------|----------|-----------|---------|--------------|
| Car driver             | under 70km | 512      | 93        | 77      | 167          |
|                        | 70-200km   | 524      | 232       | 130     | 182          |
|                        | over 200km | 631      | 316       | 187     | 223          |
| Bus                    | under 70km | 450      | 79        | 56      | 75           |
|                        | 70-200km   | 447      | 170       | 94      | 118          |
|                        | over 200km | 447      | 170       | 94      | 132          |
| Rail                   | under 70km | 451      | 108       | 94      | 109          |
|                        | 70-200km   | 391      | 183       | 120     | 162          |
|                        | over 200km | 419      | 233       | 150     | 193          |
| Metro/tram/ light rail | under 70km | 478      | 79        | 71      | 86           |
| Air                    | 70-200km   | 792      | 450       | 267     | 495          |
|                        | over 200km | 792      | 450       | 267     | 495          |



#### 2.4.4 Overview of the Dutch 2022 study

In the Netherlands, national VTT studies for passenger transport have been conducted in 1988 (Hague Consulting Group, 1990), 1997 (Hague Consulting Group, 1998) and 2009/2011 (Significance et al., 2013, Kouwenhoven et al., 2014). In the latter study, VTT values for freight transport were also determined. For other years, it was recommended to use the VTT from the most recent study and correct this for inflation and (in case of passenger transport) for 50% of the real income growth. In contrast, international practice including the UK instead relies on an income elasticity of 1 (e.g. Wheat and Batley, 2015). However, this correction method does not take changes in the VTT due to changes in comfort or due to intrinsic changes in the preferences of travellers and transporters into account. Furthermore, the 50% correction for the real income growth is also uncertain. Therefore, a new national study is conducted approximately every 10 years. This also allows to incorporate the latest insights in how the VTT can best be determined into the national travel time valuation.

In 2019, the KiM Netherlands Institute for Transport Policy Analysis published a call for a new project to determine new VTT values for the Netherlands in 2020. A consortium led by the research agency Significance was awarded this project. Due to the Covid-19 pandemic the main data collection was postponed until 2022. The analysis of the data was finalised in 2023 and published in April 2024 (Significance, 2023).

The scope of the Dutch study was slightly broader than the most recent UK study and covered eight modes of transport, respectively being car, train, local public transport, recreational navigation, airplane, airport access/egress, walking and cycling. In terms of trip purpose, however, a similar segmentation was adopted using commute, business, and leisure (similar in definition to 'other non-work'). There was specific attention for business travel, but one of the lessons that was drawn from the 2014/2015 UK study was not to interview employers regarding business travel but test different ways of eliciting the business VTT from interviewing employees (i.e. Hensher method versus willingness-to-pay method; the outcomes of the latter method were selected in the end).

The data collection took place between June and September 2022. In contrast to the UK study, roughly 80% of respondents were recruited from a high-quality internet panel and about 20% were recruited by an interviewer intercepting them during their travel. The overall sample size (see **Table 12**) is slightly larger than the most recent UK national VTT study, partly due to the inclusion of more modes of transport, but for the individual mode-purpose combinations coverage is somewhat lower. After applying data cleaning 7,644 responses were used for data analysis.

Table 12: Completed survey responses in the 2022 Dutch national value of time study before cleaning (Significance, 2023)

| Mode             | Purpose |          |       | Total |
|------------------|---------|----------|-------|-------|
|                  | Commute | Business | Other |       |
| Car              | 935     | 541      | 854   | 2,330 |
| Train            | 689     | 282      | 657   | 1,628 |
| Bus, tram, metro | 591     | 208      | 695   | 1,494 |
| Air              | 0*      | 467      | 1,305 | 1,772 |
| Cycling          | 451     | 180      | 521   | 1,152 |
| Walking          | 256     | 154      | 403   | 813   |
| Recr. navigation | 0       | 0        | 347   | 347   |
| Total            | 2,922   | 1,832    | 4,782 | 9,536 |

\* One respondent travelling by plane selected commute as travel purpose and was labelled as “business”.

The data collection is entirely based on an SP survey collecting information regarding trip characteristics and personal characteristics alongside a range of choice experiments asking respondents to make trade-offs between travel cost, travel time and where relevant other features of the trip (reliability, congestion etc.). A total of ten different SP survey formats are applied in the study, as presented by **Table 13**. Alike the most recent UK national value of time study, each respondent started with within-mode route choices based on travel time and travel cost trade-offs in SP1. In the Dutch study respondents were presented with eight choices per SP game and two SP games in total. This is in contrast with the UK study where respondents were presented with three SP games answering five choice tasks in each. For car drivers this implied that only an SP2 design was included covering the value of travel time reliability, but that the value of travel time is not decomposed into different traffic conditions (as per SP3 in the UK study). For public transport SP3A considers the following attributes: in-vehicle travel time; access/egress travel time; total wait and transfer time; number of transfers; and travel cost. The comfort experiment (SP4A) was rather like the crowding experiment in the UK study and included in-vehicle travel time; level of crowding; indicator whether you were able to sit, or whether you had to stand; and travel costs as core attributes, but in addition also added the frequency as an additional attribute. Details of the non-conventional modes (air, cycling and walking) are discussed in **Chapter 7** as these are not relevant for comparison with the UK study.

Table 13: Overview of the ten SP formats used in the 2022 Dutch national value of time study (Significance, 2023)

| Travel mode                                  | First experiment                                     | Second experiment                                                    |                                                               |                                                            |
|----------------------------------------------|------------------------------------------------------|----------------------------------------------------------------------|---------------------------------------------------------------|------------------------------------------------------------|
| Car driver                                   | <b>SP1A</b><br>Value of travel time                  | <b>SP2A</b><br>Value of travel time reliability                      |                                                               |                                                            |
| Train                                        |                                                      | <b>SP2A</b><br>Value of travel time reliability                      | <b>SP3A</b><br>Value of travel time comfort (trip components) | <b>SP4A</b><br>Value of travel time comfort (crowding)     |
| Local public transport<br>(bus, tram, metro) |                                                      | <b>SP2A</b><br>Value of travel time reliability                      | <b>SP5A</b><br>Value of access travel time to an airport      | <b>SP6A</b><br>Value of egress travel time from an airport |
| Air                                          |                                                      |                                                                      |                                                               |                                                            |
| Cycling                                      | <b>SP1B</b><br>Value of travel time                  | <b>SP2B</b><br>Value of travel time comfort                          |                                                               |                                                            |
| Walking                                      |                                                      |                                                                      |                                                               |                                                            |
| Recreational navigation                      | <b>SP1C</b><br>Value of waiting time for bridge/lock | <b>SP2C</b><br>Value of reliability of waiting times for bridge/lock |                                                               |                                                            |

The analysis of the Dutch data followed a similar process to that of the UK. The final models consider the heterogeneity in preferences and characteristics, by employing random parameter (or mixed) logit choice models, extended with a large set of interaction factors based on personal and trip characteristics. These models were developed in accordance with the linear Random Utility Maximisation (RU-Lin) framework. The alternative methods eventually adopted in the UK (e.g. random valuation with multiplicative error terms, sign/size effects of the time and cost differences) were tested but not selected. For example, rigorous evaluation revealed that the resulting VTT of a utility function with, versus without, the inclusion of sign- and size-effects rendered nearly identical results.

Translating the obtained behavioural models of the VTT followed a different process relative to the most recent UK national VTT study. In this case, the estimated behavioural models are applied to the sample of respondents, *not the Dutch equivalent of the NTS*, to calculate their personal VTT. The weighted mean VTT over the relevant subsets of respondents was calculated to obtain the national average VTT. The sample enumeration process included two types of weights and thereby mirrors the sample enumeration process used in the UK. The first set of weights correct for biases in the sampling and generates a set of weights to obtain a representative population of travellers and trips. This is akin to the weighting implemented in Germany and Norway. The sample weights were obtained using an iterative proportional fitting method (IPF), where targets for the IPF were obtained from the national travel survey ODin 2022, accounting amongst others for geographical coverage, time-of-day, personal and household characteristics. The second set of weights accounts for trip duration – akin to the use of distance weighting in the UK – and provides a higher weight to longer journeys. Alike the UK national VTT study the results from SP1 forms the base level for the national average VTT. The final national average VTT measures for the main modes of transport are presented below in **Table 14**.

Table 14: Value of travel time for (motorised) land modes, in € / hr (price level 2022) (Significance, 2023)

| Mode/purpose               | Commute | Business | Other | All purposes |
|----------------------------|---------|----------|-------|--------------|
| Car                        | 10.78   | 21.20    | 9.60  | 10.42        |
| Train                      | 12.05   | 17.96    | 8.64  | 10.08        |
| Local public transport     | 7.62    | 14.39    | 6.66  | 7.12         |
| All (motorised) land modes | 10.76   | 20.63    | 9.34  | 10.19        |

#### 2.4.5 Summary of recent national VTT studies

**Table 15** summarises the five national VTT studies on key points, which are briefly discussed below before comparing the main VTT values for motorised transport. Considering the UK 2014/15 study against the four subsequent national VTT studies presented, it can be concluded that the data collection strategies and data analysis methods applied remain relevant.

Stated Preference remains the predominant way of collecting data, and the majority of national VTT studies present respondents with simple within-mode route choice tasks trading-off travel time and travel costs. Differences across studies emerge in the extent to which SP scenarios are presented for other journey components, such as travel time reliability, congestion, and crowding. These differences are largely presentational and result in slightly different units being valued, but these do not capture a step change in the way the data are collected. The German study, and more recently the Swiss national VTT study, take an alternative approach and present multi-attribute mode choice and route choice SP studies to elicit the VTT. Moreover, they also contrast short- (mode and route) and long-term (residential and job location) choices showing signs that the field is evolving. However, all recommended values are based on short-term mode and route choice modelling. Germany is also the only study which has a more active revealed preference mode choice component as part of its study supplemented with SP data. The Singapore study stands out, not only by its geographical location, but also by its somewhat unconventional modelling approach, and hence we do not cover it in detail in the remainder of this report.

The core modes and journey purpose segmentations covered across the VTT studies are largely the same in that car, bus, rail, and other public transport modes are included in all studies, but that the pairing of segmentation somewhat varies across studies. For example, the Netherlands considers bus as part of local public transport, whereas these are separated out by most other national VTT studies. All studies account for commute, business, and leisure trips, but the German study additionally separated out education and shopping trips.

Recruitment strategies are discussed in more detail later in this report, but all studies rely to a large extent on intercept sampling supplemented with online, and (or) computer-assisted telephone or personal interviews. Particularly the Norwegian and Dutch studies have investigated the potential differences in the elicited VTT across data collection methods and flag that online panels typically come out with lower VTT values partially due to self-selection effects.

What does stand out is that the approach taken in the UK to derive recommended national average VTT values across segments is distinct from the approach taken in Germany, Norway, and the Netherlands. First, all the other national studies derive national averages by applying the estimated behavioural models to the estimation sample, as opposed to the National Travel Survey (NTS) in the UK. The largest benefit of using the SP-based estimation sample is that all modelled variables are included in the dataset. In the latest UK study specific assumptions had to be made regarding several

variables which were collected in the SP survey and used as explanatory variables of variation in the VTT, but these were not included in the NTS. A downside of using the SP sample may, however, be that elicited travel patterns are less representative and thereby puts more emphasis on the representativeness of the implemented sampling protocol both in terms of its socio-economic characteristics and in terms of their trip characteristics.

The national studies in Germany, Norway, and the Netherlands all recognise that the SP sample is not representative in terms of the travelling population and the type of trips they make. All referred national studies, including the latest UK study, implement trip weights to correct for this bias present in the SP sample and the NTS, respectively. It is only in the Netherlands and in the UK that an additional weight is implemented. The Dutch study implements a time weight, whereas the UK study implements a distance weight. In both studies this additional weight recognises that travellers with longer journeys are more likely to experience the benefits from travel time savings and should be given more weight accordingly. This process inflates the VTT in the Netherlands and the UK relative to not implementing these “distance weights”.

The UK is the only country reducing the segmentation of the VTT for commute and leisure-based trips. Specifically, an equity weighted VTT for these journey purposes is derived which is identical across all modes of transport. The argument used in the UK context is the limited ability of the study to confidently separate mode- from user- effects, and that accordingly lower estimates for the VTT by bus may undesirably deviate transport investments from lower income groups. The other national VTT studies reviewed in this report do not report on these considerations and provide estimates of the VTT by journey purpose and by mode.

To make the recommended national VTTs comparable across countries and across time, we obtain country specific Consumer Price Indices (CPIs) between 2010 and 2022 and national currency exchange rates for 2010 from the OECD Data Archive. This allows expressing the recommended national VTTs for the main motorised modes of transport car, rail and bus, and journey purposes commute, business, and leisure in 2010 GBP in **Table 16** and **Figure 2**. Across the board, the German VTT values come out as the lowest, whereas the Dutch values for car and rail, and commute and leisure are in the same order of magnitude. For bus journeys, the UK values for commute and leisure are higher, due to the aforementioned issue of not separating out the VTT across modes for these purpose segments in the UK. Relative to the Dutch study, the business values of time in the UK are 50-150% higher, particularly for rail. The Norwegian study, however, comes out with much higher business values time, whereas the values for the other mode-purpose combinations are a lot more comparable to those in the UK. Apart from the lower values found in the German study, which was conducted two years before the UK study and made use of a distinct survey design, there is no clear pattern of change in the recommended VTTs. This does, however, not preclude those societal developments, as reviewed in the next section, having had an impact on the VTT in the UK since completion of the last study in 2014/2015.

Table 15: Summary of national VTT studies between 2012-2023

|                           | <b>UK</b>                                                                                               | <b>Germany</b>                                                                                                                     | <b>Singapore</b>                                                            | <b>Norway</b>                                                                                                      | <b>Netherlands</b>                                                                                                |
|---------------------------|---------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|
| <b>Year of study</b>      | 2014-2015                                                                                               | 2012-2013                                                                                                                          | 2015                                                                        | 2018-2019                                                                                                          | 2019-2023                                                                                                         |
| <b>History</b>            | 4 <sup>th</sup> national study                                                                          | 1 <sup>st</sup> national study                                                                                                     | 2 <sup>nd</sup> national study                                              | 3 <sup>rd</sup> national study                                                                                     | 4 <sup>th</sup> national study                                                                                    |
| <b>Data collection</b>    | Primarily SP<br>RP for rail only                                                                        | RP included in mode choice supplemented with broader SP design.                                                                    | SP only                                                                     | SP only                                                                                                            | SP only                                                                                                           |
| <b>Included modes</b>     | Car, Rail, Bus, Other PT, <i>walking and cycling</i>                                                    | Car, PT, Aviation, Cycling, Walking                                                                                                | Car, motorcycles, mass rapid transit (MRT), bus, taxi, walking, and cycling | Car driver, car passenger, ferry, bus, train, metro/tram/light rail, passenger boat, airplane, cycling and walking | Car, train, local public transport, recreational navigation, airplane, airport access/egress, walking and cycling |
| <b>Journey purposes</b>   | Business, commute, other non-work                                                                       | Education, commuting, shopping, business, and leisure                                                                              | Not directly sampled for                                                    | Commute, business, and leisure                                                                                     | Commute, business, and leisure                                                                                    |
| <b>Recruitment</b>        | Intercept + CATI                                                                                        | Business (online)<br>Other (CATI)                                                                                                  | CAPI + intercept                                                            | Online panel, email, and intercept                                                                                 | Online panel and intercept                                                                                        |
| <b>SP choice</b>          | Route choice                                                                                            | <b>Short term:</b><br>Mode choice, route choice, departure time choice<br><b>Long term:</b><br>location choice (residence and job) | Route choice                                                                | Route choice                                                                                                       | Route choice                                                                                                      |
| <b>SP setup</b>           | Time vs Cost, added experiments with additional attributes (reliability, comfort, congestion, crowding) | All experiments include more than two attributes, crowding is not included                                                         | All experiments include more than two attributes                            | Time vs Cost, added experiments with additional attributes (reliability, comfort, congestion, crowding)            | Time vs Cost, added experiments with additional attributes (reliability, comfort, congestion, crowding)           |
| <b>Sample enumeration</b> | National Travel Survey                                                                                  | SP sample                                                                                                                          | Mean of estimated random parameters                                         | SP sample                                                                                                          | SP sample                                                                                                         |
| <b>Trip weights</b>       | Yes                                                                                                     | Yes                                                                                                                                | No                                                                          | Yes                                                                                                                | Yes                                                                                                               |
| <b>Distance weights</b>   | Yes                                                                                                     | No                                                                                                                                 | No                                                                          | No                                                                                                                 | Yes (time weights)                                                                                                |

Table 16: Recommended national VTTs for the main motorised modes (expressed in 2010 £)

|                             |          | UK      | Germany | Norway* | Netherlands |
|-----------------------------|----------|---------|---------|---------|-------------|
| Car                         | Commute  | £ 7.14  | £ 3.27  | £8.46   | £6.92       |
|                             | Business | £ 22.91 | £ 6.87  | £46.57  | £13.60      |
|                             | Leisure  | £ 6.35  | £ 3.89  | £7.00   | £6.16       |
| Rail                        | Commute  | £ 7.14  | £ 3.54  | £9.82   | £7.73       |
|                             | Business | £ 28.19 | £ 5.75  | £41.02  | £11.52      |
|                             | Leisure  | £ 6.35  | £ 3.56  | £8.55   | £5.54       |
| Bus                         | Commute  | £ 7.14  | £ 3.54  | £7.19   | £4.89       |
|                             | Business | £ 14.66 | £ 5.75  | £40.93  | £9.23       |
|                             | Leisure  | £ 6.35  | £ 3.56  | £5.09   | £4.27       |
| Avg wage per hr (2010£ PPP) |          | £ 23.99 | £ 28.37 | £ 29.41 | £34.26      |

\* values based on trips under 70km and assuming car drivers

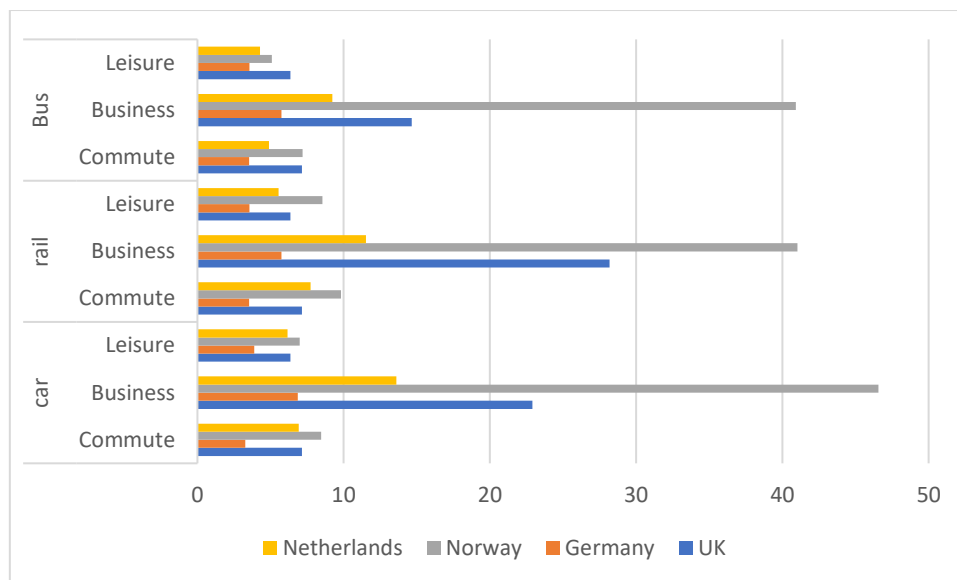


Figure 2: Bar chart with recommended national VTT values by mode/purpose combinations (2010 £)

## 2.5 Societal development since the last UK national VTT study

Since the 2014/15 UK national VTT study society has been exposed to a series of rather challenging circumstances, including Covid-19 and a 'cost of living' crisis. On top of that, particularly urban travel has seen the emergence and maturing of e-mobility in the form of e-scooters and e-bikes both in terms of private ownership and sharing schemes in larger cities. Electric vehicles have expanded their representation in the vehicle fleet, including buses, and public transport has seen increases in availability of WiFi and charging facilities potentially facilitating productive use of travel time and increasing the comfort of the journey. This is complemented by improved mobile signal facilitating connectivity across the country, and transport providers' ambition to provide mobility as a service (MaaS) intended to increase the information provision to travellers and facilitate multi-modal travel through integrated ticketing.

**Figure 3** highlights that between 2002 and 2022 bus patronage has seen a steadily decline across England, although this pattern has been less pronounced for buses in London. Surface rail saw strong growth up to 2019, and the London underground observed slight increases and stable patronage

numbers respectively until the Covid-19 pandemic. Across all public transport modes, a steep decline was observed for 2020 and continued into 2021 due to subsequent full lockdown and travel restrictions put in place. For the year 2022 initial signs of recovery in patronage levels are observed for most modes of public transport, except the London underground, although patronage levels remain below pre-pandemic levels.

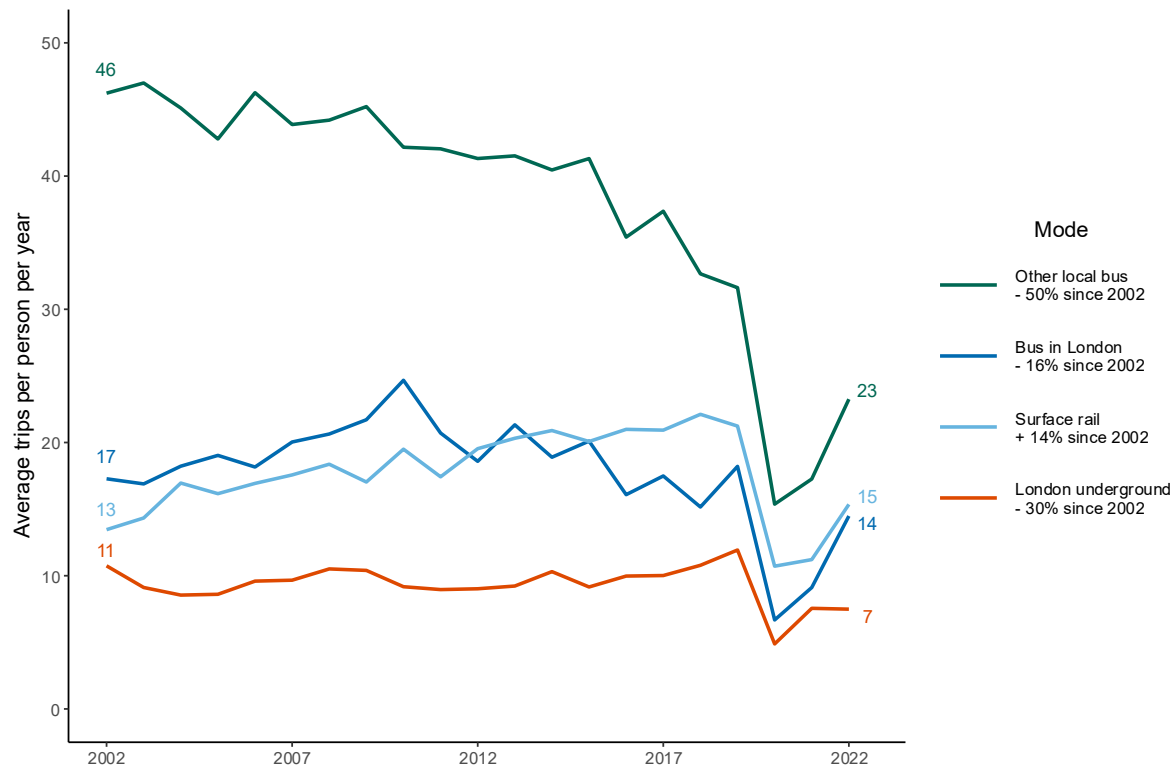


Figure 3 Average trips per person per year in England by public transport (Department for Transport, 2023c)

More recent figures from the Department for Transport reveal that private motorised transport has recovered to pre-pandemic levels in Great Britain, but rail and bus patronage levels are still seeing lower levels of demand in **Figure 4**. There is considerable variability within the headline figures, including notable shifts in terms of day of week, time of day and journey purpose. Outside London, bus travel appears to continue its downward trend again, despite fixing bus fares at £2 for single journeys for most of the country until 31 December 2024 (Department for Transport, 2024a).

In a response to high levels of inflation the UK government capped bus fares to £2 between January 2023 and 2024, and evidence shows that patronage was recovering from the Covid-19 pandemic to which the change in fares may have contributed. However, evidence also highlights that mostly existing bus users make a small number of additional trips and that the change in fares does not necessarily attract passengers from other modes (Department for Transport, 2023a). Fountas et al. (2023) identify that for Scotland the cost-of-living crisis influenced mode choices to save money and that the number of trips made also declined. Notably, the cost-of-living crises also exposed significant inequalities in transport accessibility (Johnson, 2023).



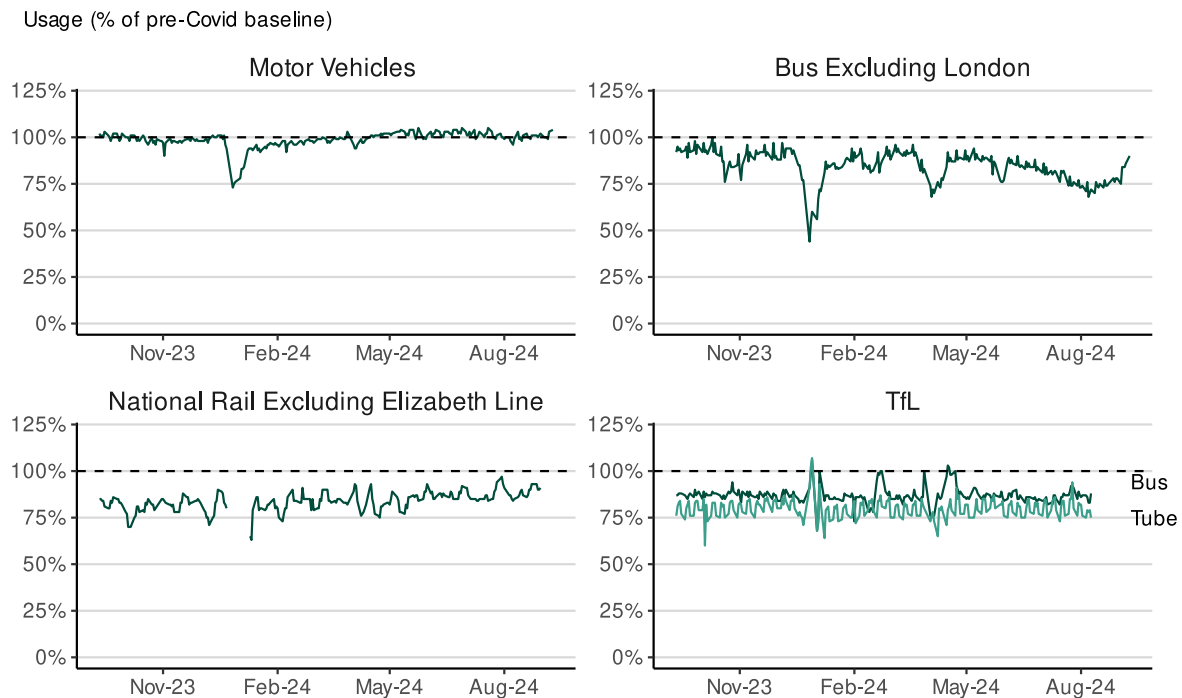


Figure 4: Transport usage as a proportion of pre-COVID-19 levels, Great Britain, September 2023 to September 2024 (Department for Transport, 2024c)

The above evidence highlights that mode choices and travel patterns have changed because of the Covid-19 pandemic and cost-of-living crisis. Although travel patterns seem to have stabilised over the last year for most modes of public transport, demand levels are still below the level it would be expected based on pre-Covid trends. It is not unlikely that the VTT has changed accordingly with these societal challenges.<sup>4</sup>

Apart from the short-term responses and recovery to the Covid-19 pandemic, working patterns have significantly changed. Data from the ONS (2024) 'Opinions and Lifestyle Survey' (OPN) related to the working arrangements of people in Great Britain reveal that the pandemic increased the proportion of adults working from home significantly and this has consistently remained high despite lifting the restrictions on travel. Shah et al. (2024) argue, based on firm-level data, that managers forecast levels of remote work within their own firms in 2028 to be almost identical to the levels in 2023. Higher levels of working from home are simultaneously associated with higher productivity but lower wage growth. Finally, Ilham et al. (2024) highlight that the pandemic has affected residential location choices with households moving towards desirable dwellings at affordable prices, which typically are located further away from the city centre. This affects commuting patterns, but simultaneously the increase in online shopping has also led to permanent changes in retail destinations.

Besides the societal challenges highlighted, significant changes have taken place in the electrification of modes of transport. CoMoUK (2023a) highlight that in September 2023 the number of e-bikes in UK bike sharing schemes outnumbered traditional pedal bikes for the first time. Specifically, the number of e-bikes in bike sharing schemes increased from around 150 in 2020 to nearly 26,000 in less than four years, and presently nearly 60% of bike share hires are made with e-bikes. Key reasons for choosing electric bicycles include saving time, preventing fatigue and sweat, cycling up hills, and

<sup>4</sup> Task BC.5 explores further the impact of the Covid-19 pandemic on crowding levels.

travelling longer distances. Across the UK trials for the use of e-scooters have taken, and are still taking, place since July 2020 (Department for Transport, 2024d) and CoMoUK (2023b) reports on various user statistics highlighting that e-scooter trips largely replace walking trips, but are increasingly replacing trips by private motor vehicle or taxi. With e-bikes and e-scooters becoming increasingly available through private ownership and sharing schemes there is a significant opportunity to change individual mode choices. Public transport, and particular bus journeys, are also seeing increased levels of electrification through the Zero Emission Bus Regional Areas (ZEBRA) programme (Department for Transport, 2024n). The programme is associated with an expectation that electrification of the bus fleet will not only reduce emissions, but also increases user satisfaction by improving the comfort of travelling on buses. Finally, **Figure 5** highlights that the percentage of battery electric vehicle (BEV) ownership has significantly increased in recent years but that there is still ample opportunity for increasing its market share and overtake conventional petrol and diesel cars as being the dominant drive train in the UK's vehicle fleet.

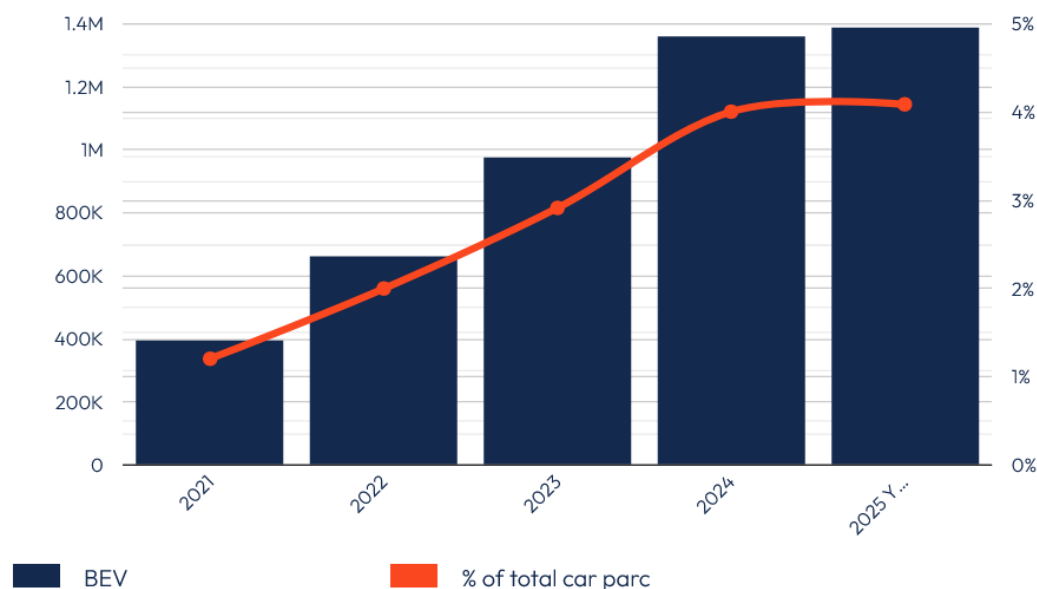


Figure 5: Percentage of BEV ownership. Source: SMMT and DfT January 2025, <https://www.zap-map.com/ev-stats/ev-market>

Mobility as a Service or MaaS forms one of the key components as part of the '*Future of Mobility: urban strategy*' (Department for Transport, 2019). MaaS is typically defined as the integration of various modes of transport along with information and payment functions into a single mobility service. In theory, such levels of integration have the potential to significantly improve the satisfaction of travel. The Department for Transport (2023b) recently set out a code of practice for MaaS highlighting that extensive user testing in terms of accessibility and data security amongst other things is required before implementation. The code of practice echoes recommendations made by Enoch (2018) based on a review of MaaS in the UK.

Overall, the societal challenges and changes in the transport environment articulated above may very well have changed the VTT. On the one hand, increasing pressures on private (travel) budgets may have affected the marginal utility of income potentially reducing the population's willingness-to-pay for travel time reductions. On the other hand, improvements in the quality and comfort of travel may

affect the marginal (dis)utility of travel time and accordingly raise the VTT. As indicated by **Figure 4**, travel patterns in the UK have largely stabilised since the Covid-19 pandemic. Accordingly, there is no indication that the transport market is in such significant turmoil that a new national VTT study should not be conducted to pick up changes in the VTT since the last national VTT study. Moreover, any changes which have taken place in travel behaviour are unlikely to have significant implications for the design of a new VTT study. In fact, the availability of a wider range of micromobility options (e-bikes and e-scooters) may offer the opportunity to resolve some of the challenges eliciting the VTT for walk and cycling in the previous national study. Nevertheless, the observed changes in demand profiles (e.g. time of day, frequency, and purpose) may have impacts on implemented sampling strategies.

### 3. Developments in surveying and estimation methods

#### 3.1 The potential use of Revealed Preference data

Whilst the present national VTT values adopted in TAG are entirely based on evidence from Stated Preference (SP), a recurring question is the extent to which Revealed Preference (RP) data are suitable for obtaining a representative VTT measure.

The strong interest in RP methods comes from the interest in studying real word choices, where travellers must pay for reductions in travel time, as opposed to deriving a WTP estimate from choices made in hypothetical SP scenarios. It is believed that real world choices are a closer reflection of the population's 'true' preferences for travel time savings.

In ARUP et al. (2015a) RP was only used for validating the outcomes from SP in the case of business travellers. The RP exercise was targeting within-mode route choices for rail journeys between Birmingham and London, Peterborough and London, and Stoke-on-Trent, Stafford, and Rugby to London. In these locations travellers had a choice between different train operators on the same route.<sup>5</sup> In the empirical analysis it was, however, concluded that it was rather difficult to disentangle the VTT from operator preferences, such that the RP was only moderately useful as an exercise in validation. **Section 2.4** of this report provided further evidence that SP is the dominant means of data collection. Only the German VTT study (Axhausen et al., 2014) and more recently the Swiss VTT study have made some use of RP, but the RP mode choices studied were supplemented with SP data.

From a historical perspective, much of the early core developments in VTT took place with the use of RP data, especially in the 1970s and 1980s. Early attempts at estimating the VTT in the UK were based on the use of RP methods (e.g. Beesley, 1965). Amongst the core reasons for abandoning RP in this context were the limited volumes of available data, and the lower informational content of RP data relative to SP due to high levels of correlation between the key variables of interest, i.e. travel time and travel costs. Particularly the high levels of correlation limit the estimation of accurate VTT values and explains the use of more complex models in SP research (ARUP and ITS, 2017, ARUP and ITS, 2018).

SP data gradually became increasingly popular, to the point of becoming the dominant source in both academia and practice. Notable exceptions remain, such as in large scale modelling work. The easy availability of SP data and ability of analysts to customise the contents of their surveys facilitated many of the methodological developments from the 1990s onwards. However, doubts about the reliability of SP data have never gone away, and there remains a lack of work on validation of SP results, including for VTT. In addition, there are concerns that many of the behavioural phenomena uncovered in work using SP data, notably in terms of heterogeneity, reference dependence and other behavioural phenomena, may in fact be amplified by the hypothetical setting (e.g. Hess et al., 2020).

The ARUP and ITS (2018) scoping study concluded that despite emerging big data sources increasing the availability of large volumes of data, which may overcome some of the correlation concerns but not all, RP was at the time not (yet) considered a suitable approach for estimating a nationally representative VTT value. The primary reason for reaching this conclusion was that passive big data sources do not include the same level of detail as SP data. In other words, passive data sources are unlikely to include information on mode of travel, journey purpose, travel cost and travel time, related travel conditions, and socio-economic characteristics of the traveller. These limitations apply to the

---

<sup>5</sup> Notably, the different train operators make use of different origin and destination stations potentially limiting the extent to which the revealed operator choices are a pure representation of the traveller's VTT.

chosen alternative and are generally even more severe for unchosen alternatives. All relevant variables are directly elicited in SP surveys and used in the analysis and segmentation of the VTT.

In a more recent review study, Flügel et al. (2022) cover a wide range of RP data sources and evaluate their potential for use in valuation. All data sources are scored on a range of criteria including 'data access and general quality', 'opportunity for analysis', and 'flexibility, synergies and future perspective'. The conclusions of the review are to a certain extent comparable with ARUP and ITS (2018) in that passive data sources still offer limited potential due to the lack of background data limiting the opportunity for analysis, and limited degree of access to, for example, mobile phone data. In the eventual ranking of options, however, data from apps tracking individuals by means of GPS are ranked highest due to their high levels of accuracy, frequent collection of trips, and the opportunity to combine them with surveys collecting the necessary background data for model development.

Although GPS-based tracking with background surveys are considered the new gold standard for valuation by Flügel et al. (2022), they are still recommending the use of small scale supplementary SP experiments for estimation of unit values that may be difficult to estimate from RP data.

Continuing with the currently dominant approach is only recommended by Flügel et al. (2022) when combined with several well-crafted RP case studies to validate/adjust the overall level of VTT. Reflecting on the difficulty to find like-for-like RP route choice situations mimicking the dominant within-mode route choice format in most SP-based national VTT studies (e.g. ARUP et al., 2015a), this may require reformulating the types of SP choices presented. The combined use of RP-SP data in the German VTT study (Axhausen et al., 2014) using a mode choice situation is one way forward.

Flügel et al. (2022) furthermore consider the use of passive RP data more as an opportunity for quality assurance of VTT values estimated from surveys or app-based studies, due to potential selection biases in these methods arising from recruiting and maintaining respondents.

The conclusions from Flügel et al. (2022) resonate with a growing body of academic work showing the value of RP data, in terms of the richness of results. In a range of studies using controlled experimental settings, RP data have been shown to allow estimating the WTP of taxi users to reduce waiting time (Goldschmidt et al., 2020, Buchhold et al., 2024), and to estimate the VTT based on congestion pricing policies (Kreindler, 2024). Although these studies are impressive, they are very reliant on the researchers' ability to control the experimental settings and thereby not suited for implementation in the context of national VTT studies with a broader range of time components in need of valuing.

Tsoleridis et al. (2022) is closer in design to Flügel et al. (2022)'s recommended approach. The study uses GPS-data from a two-week trip diary collected in Yorkshire and the Humber, not supplemented with additional SP tasks but with generic background data questions, and estimates mode choice models. The VTT estimates obtained after sample enumeration resemble those of ARUP et al. (2015a). Tsoleridis et al. (2022) interpret this as a strong case for moving towards RP generated data sources.

Tabasi et al. (2023) make use of the suggested combination of GPS data supplemented with SP (mode choice) data in the context of Sydney, Australia. The study estimates a wide range of VTT components whilst clearly highlighting which VTT estimate is related to RP/SP based evidence. For in-vehicle time for car journeys the RP-based VTT estimates is somewhat higher than its SP-based equivalent, but the confidence intervals are largely overlapping.

Tabasi et al. (2023) recommend that travel behaviour modellers not only collect RP data for their research but also allocate some time and effort to collect complementary SP data. This is because SP data collection is relatively cheap and contains in general much more information than RP. On its own,

this conclusion is not new to the literature (e.g. Hensher and Bradley, 1993), but it is another example illustrating the increasing potential of GPS-based data for VTT estimation purposes.

Although GPS-based data is accurate, some essential information at the trip-level – such as journey purpose, and travel cost – is still missing. This can either be inferred from external sources (e.g. Tsoleridis et al., 2022) or directly verified by the surveyed travellers (e.g. Winkler et al., 2024). However, not all surveyed travellers will verify the requested information (Ali et al., 2024).

Another set of RP-based surveys, not relying on GPS data but on traditional trip-diaries, collects travel behaviour alongside activity and expenditure patterns in Austria and Switzerland (Schmid et al., 2021, Hössinger et al., 2020, Aschauer et al., 2019, Schmid et al., 2019, Jokubauskaitė et al., 2019). This Mobility Activity and Expenditure Diary (MAED) study design is discussed in more detail in **Chapter 4** and allows estimating the VTT and the value of leisure (VoL) and thereby the researchers can infer the value of time assigned to travel (VTAT) due to comfort by mode. Due to its close relationship with economic theory, MAED may, in principle, be considered a preferred method of data collection. The MAED is, however, a rather intensive and perhaps its data collection method asking respondents to reveal a lot of personal information can be considered somewhat intrusive by its respondents. This is therefore considered a less feasible data collection method for deriving national VTT estimates.

The challenges of using passive data sources, like mobile phone data, for the estimation of VTT are illustrated by Andersson (2024). The study explores the potential of using mobile phone data for estimating long-distance mode choice models in Sweden. The emerging VTT estimates are associated with large variations across model specification, and wide confidence intervals. Moreover, the values are largely inconsistent with existing evidence in Sweden (Börjesson and Eliasson, 2014).

With the presented evidence highlighting that there is potential for increased usage of RP methods in the estimation of the VTT, it is important to reflect on what this means in terms of continuity. A core interest in VTT work is studying trends overtime and where possible explain these variations (e.g. Börjesson et al., 2023). Potential reasons for changes in the VTT may include economic growth, but also changes in comfort or productivity whilst travelling. For this reason, it can be seen as problematic if successive national studies rely on different methods, be they different sources of data (RP vs SP), different experimental designs, or even different models. Accordingly, there has been a reluctance to make significant changes to the data that is used in these studies. This has meant not just a reliance on SP data per se, but also the use of the same types of experiments. A core example is the use of ‘simple’ time vs cost trade-offs in the within-mode route choice settings used in the UK, the Netherlands and Norway, but also Denmark and Sweden.

Conversely, the type of model has in fact received little attention in this context, with a ready acceptance that any new study should make use of state-of-the-art modelling techniques. It is well known that results differ across model specifications, with more advanced techniques typically avoiding the biases of simpler methods and thus leading to different VTT measures. For reasons of comparability, model specifications implemented in earlier studies are often replicated to separate out the effect of the newly adopted and improved model specification. This is, however, fully comparable only where the experimental design and questions included in the previous survey are also retained..

Changes in the experimental design of the same SP format may introduce another source of difference. Namely, the collected data will have different levels of informational content and may thereby influence the precision by which certain parameters of interest can be estimated (i.e. wider standard errors). If changes in the experimental design unintentionally influence response behaviour even bias may arise in the VTT estimates. Not only would that be a concern with the reliability of SP, but it would

also hamper the intertemporal comparison of VTT estimates. Without proper experimental settings in the research design it is nearly impossible to identify the impact of the underlying experimental design on the VTT estimates. Significance (2023) present a percentage of their respondents with the 'old' designs.

Appendix A of Significance (2023) shows that even within the simple time vs cost trade-offs the experimental designs have changed significantly across the presented range of VTT studies. **Figure 6**, where individual respondents within a given national VTT study are stacked on top of each other based on the average boundary VTT presented to respondent (blue line), presents in black the range of boundary VTTs presented across the different choice tasks to a given individual. The message from **Figure 6** is that variations in SP designs over time can be one explanatory factor for observed changes in the VTT estimates over time, and across countries.

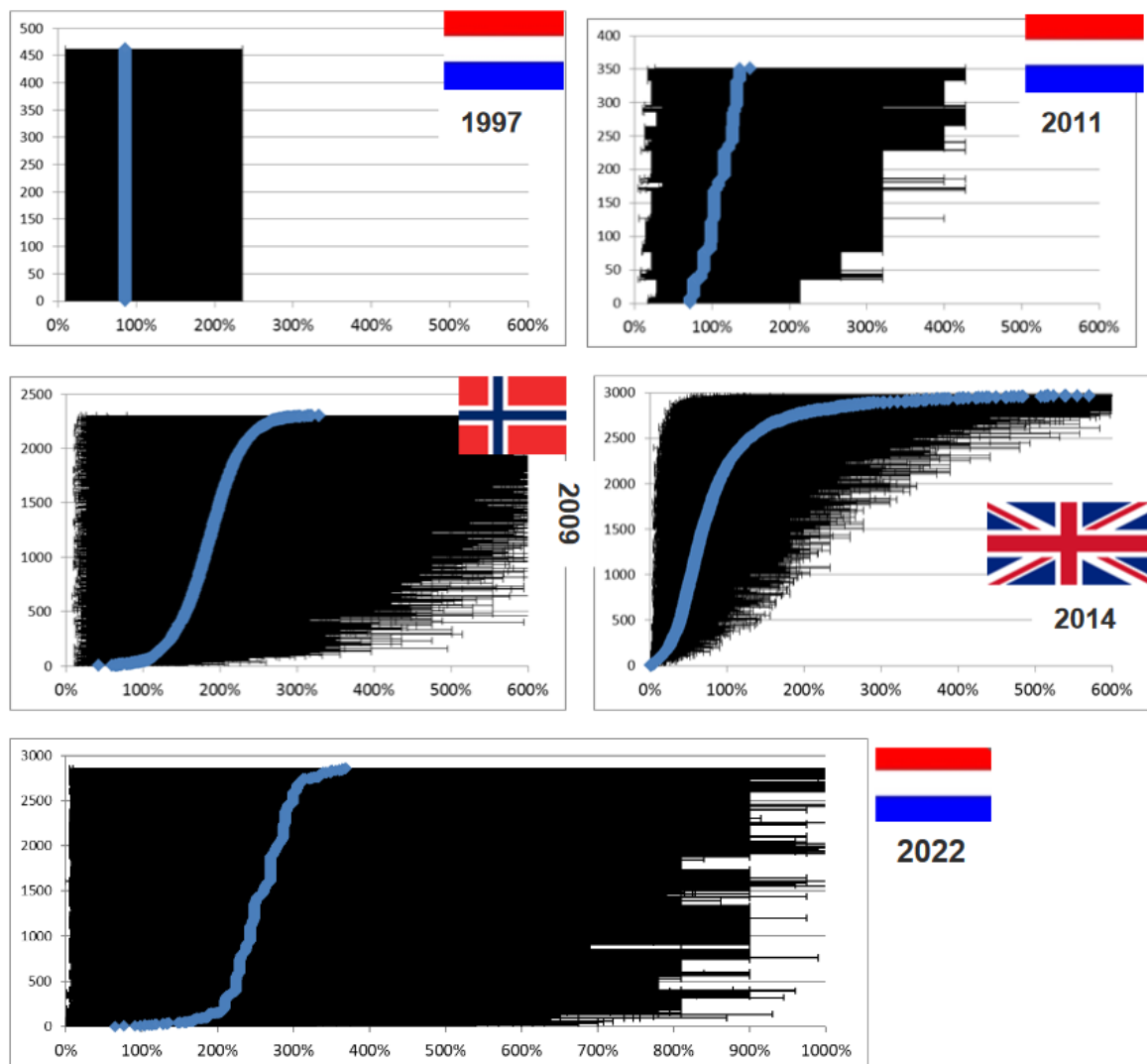


Figure 6: Overview of boundary VTT ranges presented to respondents in national VTT studies. The vertical axis represents the different respondents ordered by the average boundary VTT presented (blue line) with the ones at the top being presented with the highest boundary VTT. The horizontal black lines denote the variation in the presented boundary VTTs within a respondent. The horizontal axis is normalised where 100% represents the final VTT (Significance, 2023)

Another argument used to describe the reluctance to the use of RP in VTT work focused on a lack of real-world choice situations involving time/money trade-offs. This, however, seems to be largely due

to the interpretation that VTT work should make use of within-mode data, as per the above standard presentation of time vs cost trade-offs on route choice in most national VTT studies.

The continued reliance on within-mode settings in ARUP et al. (2015a) has potentially hampered the efforts to validate the SP results. For car travel, within-mode time/money trade-offs do not exist in the UK except for a handful of toll roads/tunnels/bridges, and for a minority trade-offs between slow and fast EV charging exist. This is in contrast with other countries where many real-world time/money trade-offs arise, opening opportunities for such within mode RP studies. For example, in Norway toll data is more readily available but even there the coverage of toll roads is geographically dispersed (Flügel et al., 2022). As mentioned before, in the 2014/2015 study, the use of an operator choice setting for RP rail journeys between Birmingham and London was seen as a viable solution, but ultimately, the model results were not satisfactory. With hindsight, the dataset used was too small and lacked richness to produce reliable results, in contrast with new emerging RP data.

The situation in relation to a lack of real-world settings involving a time/money trade-off for car travel remains, and, even for public transport, options are limited. There are some routes with a choice between slower and more expensive services, but price differences arise primarily in the context of departure time choice, which adds an additional level of complexity.

Another point that is often not mentioned is that the lack of real-world time/money trade-offs for car travel also creates a potential issue for within-mode SP scenarios. Indeed, it means that the SP scenarios faced in SP1 type games can be considered unrealistic by some respondents, putting the external validity of SP1 results in doubt. If mode choice is seen as a viable option for VTT work, then this removes one of the key perceived disadvantages of RP data.

The above discussion then poses the question whether a move away from within-mode settings should be considered, as this would remove the issue of a lack of settings with time/money trade-offs. An issue which will also be flagged in **Chapter 4** if the separation of mode and user effects is judged to be necessary. The general argument in favour of within-mode choices as opposed to mode choice data has centred around an ability to estimate VTT measures net of modal preferences. Numerous academic mode choice studies have, however, shown that modal preferences can be relatively easily disentangled from VTT differences.

SP surveys allow analysts to collect large samples quickly and easily apply quotas to ensure an ability to produce representative results. This is especially true with online data collection which has become the norm for SP surveys. The notion that RP data has disadvantages in this context goes back to a time where RP surveys were primarily either intercept or household surveys. These issues no longer apply with either smartphone-based GPS surveys (where respondents can be sampled in the same way as for SP) or big data (although these have already been identified as less relevant of VTT purposes in this report). Moreover, sampling protocols for SP studies (and by extension smartphone-based GPS surveys) can still be subject to selection biases (see Section 3.3).

Policy work requires not just a single VTT measure, but segmentation by mode, type of time (e.g. congested vs uncontested) and reliability. As argued above, mode specific VTT measures can just as easily be obtained from RP data as from SP data. Segmenting travel time by quality and reliability, however, raises different issues with RP data. These issues are not trivial, and especially travel time variability likely remains the area where SP has an advantage. One possible approach in this context is to gain insights from the study of time of travel (i.e. scheduling) and mode choice. If travel time quality and reliability vary by time of day (as is clearly the case) and do so differently for different modes, then insights into their impact on mode and departure time choices can again be obtained. The ability to



study departure time choices using RP data is a positive case in point (Bwambale et al., 2019). Another approach, which is consistent with Flügel et al. (2022), is to combine RP data collection efforts with SP, such that RP can be used for eliciting headline VTT values, and SP to derive multipliers pivoting of these values.

Overall, the trend in academia clearly points to a resurgence in RP work, driven not just by the availability of new data sources, but also by ongoing concerns about the reliability of SP data. It seems clear that VTT work will in due time follow this trend, with many countries now exploring the use of RP data, at the very least alongside SP data. A promising avenue in this context would be to use RP data and SP data jointly. Most of the findings would come from RP data (in terms of real-world ground truth) with SP data contributing measures that are more difficult to obtain using RP data (such as travel time reliability or specific travel time components). In this context, RP data is used to calibrate the overall VTT measures. Irrespective of the use of additional SP data, it is highly recommended to supplement RP data with generic survey data such that relevant background information is obtained allowing the estimation of richer behavioural models.

A promising avenue is an approach explored in the most recent German and Swiss studies. Most SP VTT studies already record data on a recent trip and use that data to customise the SP scenarios through a pivoting approach. On paper, a simple additional step is to enrich the data by adding in information on the non-chosen options and estimating an RP model on the resulting data. Joint RP/SP estimation then becomes possible, especially if the choice settings can be aligned between the RP and SP data, which would again be the case with mode choice.

The preferred method put forward by Flügel et al. (2022), i.e. through the use of GPS tracking data, increases the reliance on RP data relative to SP, but does not seem to come at a significant cost, unless recruitment issues are turning out to be challenging.

Purely passive data, such as call detail record data, is cheaper and provides large volumes of data, potentially dealing with the issue of time/cost correlation due to sample size, while also yielding near-representative data at the national level. A key remaining issue is background data, in terms of socio-demographics as well as inferring modes, routes, and purposes. For this reason, any use of such data in national VTT studies would first need a focus on the errors related to (incorrectly) inferring such information.

### 3.2 State-of-the-art design of VTT SP studies

Even though RP may take a more prominent role in future national VTT studies, it is very unlikely that SP will be entirely removed. If it were not for its ability to value specific dimensions of the VTT not feasible to capture using RP, it would be for consistency purposes such that new studies can be compared to previous ones. The purpose of this subsection is to look into lessons learned regarding the design of SP studies since ARUP et al. (2015a).

#### 3.2.1 Simple time vs. cost trade-offs

As discussed in **Chapter 2**, the most recent national VTT studies in Europe are characterised by using a series of different SP games. The most basic SP format, SP1 in the 2014/15 study, only includes the attributes of travel time and travel cost. This format is also used in Norway and the Netherlands. Subsequent SP games then introduce additional attributes in the form of travel time components, comfort factors such as crowding, and (or) travel time reliability.

Although travel time and travel cost are included in all different SP formats, a prevalent challenge is that the VTT estimated across different SP games is not necessarily consistent (e.g. Hess et al., 2020).

This poses challenges for producing values for appraisal purposes in relation to how to unify these results.

Hess et al. (2020) argue for the use of complex, or as complex as possible choice scenarios, because they are more in line with their real-life experiences, where a journey is described by many attributes. The fact that the SP1 format may be considered as too simplistic is illustrated by the follow-up work on congestion (see **Section 6.1**), where it became apparent that the conditions associated with the travel time in SP1 are described are insufficiently clear both to the analyst as to the respondent.

Which attributes are ideally included in such a complex SP format should be determined by forming a balance of i) priority VTT components articulated by the relevant authority, in our case the Department for Transport; and ii) respondents' interpretation of the most relevant journey components and their ability to process multiple attributes characterising the journey. Notably, Australia has a richer experience with VTT studies where respondents are faced with surveys that successfully capture multiple components at the same time. The Swiss and German national VTT studies also make use of a wider range of attributes across all the included SP formats.

Hess et al. (2020) do explicitly recognise that the use of the simple SP1 time vs. cost format has spurred a significant amount of high-quality methodological work, specifically focusing on the heterogeneity of the VTT, the potential presence of reference dependence, and non-linearities such as cost damping. Their conclusion does not invalidate this body of work. Analysts should, however, realise the potential for this SP format to unintentionally stimulate such behaviour.

From an analyst perspective, the binary SP1 choice format entirely isolates the time vs cost trade-off and enables the use of the notion of the 'boundary value of travel time' in the analysis. This makes it very convenient to analyse the range of BVTTs presented across and within respondents (e.g. **Figure 6**), study the degree of consistency in the choices made by respondents, and determine the extent to which the tail of the VTT distribution can be captured. When moving to a multi-attribute setting, these convenient properties are unfortunately lost and the notion of the BVTT becomes meaningless because the observed increase in cost can, for example, be spend on time and safety improvements.

### 3.2.2 Experimental designs

Appendix A.3 of Significance (2023) contrasts the different design strategies implemented in the main European national VTT studies for SP1. In this context, it is typical to describe four types of choice situations contrasting changes in travel time ( $\Delta t$ ) and travel cost ( $\Delta c$ ), see **Figure 7**:

1. WTP-based choices present respondents with reductions in travel time and increases in costs relative to the reference trip.
2. WTA-based choices present respondents with increases in travel time and reductions in travel costs relative to the reference trip.
3. Equivalent gains (EG) based choices present respondents with a choice contrasting a reduction in travel time at constant travel costs against a reduction in travel costs at constant travel time.
4. Equivalent losses (EL) based choices present respondents with a choice contrasting an increase in travel time at constant travel costs against an increase in travel cost at constant travel time.

The way recent national VTT studies have dealt with the four types of choice situations and related experimental design settings has varied significantly as revealed by **Figure 8**. The so-called 'Bradley' design, as used in the previous Dutch national VTT study (Kouwenhoven et al., 2014), is a kind of orthogonal design which prevents the occurrence of dominant alternatives. The study also presented a choice out of each of the four quadrants to respondents. The 'Scandinavian' designs originate from

the 2008 Swedish study (Börjesson et al., 2012b) and presents two choices out of each of the four quadrants to each respondent, making eight choice tasks in total. Notably, the presented levels of travel time and travel cost differences relative to the reference travel time and costs are randomly drawn from different strata thereby ensuring a wide range of variation in presented attribute levels. This design was also adopted in the subsequent Norwegian VTT studies.

Where the Dutch and Scandinavian designs always present the reference trip relative to another trip in one of the four quadrants, ARUP et al. (2015a) did not replicate the reference trip in every choice task and was accordingly not classified as ‘reference-based’ by Significance (2023). Moreover, ARUP et al. (2015a) did not include any restrictions on the number of choice tasks occurring in each of the four quadrants.

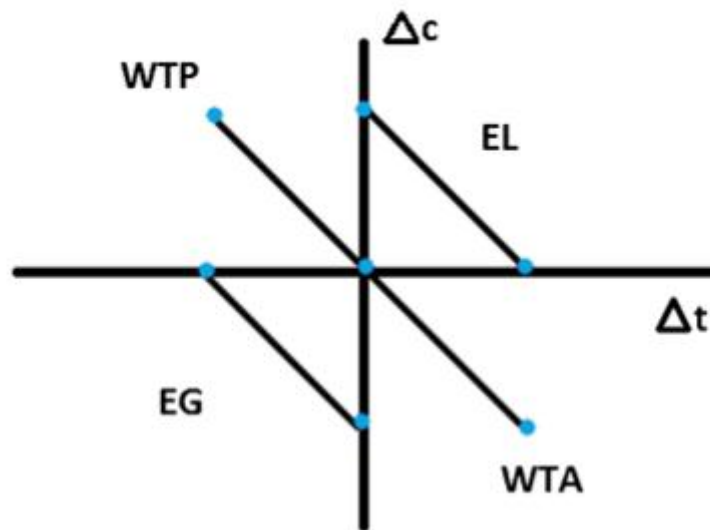


Figure 7: Four types of choices presented in most European VTT studies (Ojeda-Cabral et al., 2018)

Positive and negative deviations in travel time and travel costs from the reference trip were presented, but in order to prevent dominant choices from occurring (Bayesian) d-efficient designs were used (Bliemer and Rose, 2024). This was considered less restrictive than always presenting the reference alternative. The most recent Dutch VTT study adopts a combination of the Scandinavian and the UK-based approach.

Coming back to **Figure 6**, it becomes clear that irrespective of the design strategy adopted, variation in the (average) presented boundary VTT is observed across respondents, but that this pattern was more substantive for the 2014/15 study. Two reasons can be identified for this. First, the Dutch and Norwegian study both make use of eight instead of five SP1 choice tasks per respondent limiting the variability in presented BVTs in the UK. Second, the Scandinavian design strategy makes use of pre-defined strata to draw time differences and VTT bids to generate the BVTs for their respective designs. This makes the range of BVTs presented more comparable across respondents than in the UK where the implemented d-efficient design are not subject to such range restrictions within respondents as explained above.

The design approach adopted by Significance (2023) deals with two challenges. On the one hand, it presents a wide variation (and even larger ranges than the other studies) to capture the tail of the VTT distribution and encourage trading (i.e. prevent respondents always selecting the cheapest or quickest option) by means of the Scandinavian approach. Notably the non-trading in the Dutch study reduces

to 7% as opposed to 26% in the 2014/15 study. On the other hand, it learns more about the VTT around the selected prior VTT values by making partial use of d-efficient designs. Again, the use of more than five choice tasks really enables both benefits to materialise.

Ojeda-Cabral et al. (2018) remind us that different design settings may affect the obtained VTT estimates, this is not only the result of induced behavioural responses, but also by different levels of informational content comprised by the given design as already articulated in **Section 3.1**. Unfortunately, the referred paper does not provide specific guidelines for generating good experimental design.

What should, however, be remembered is that as sample sizes get larger the impact of the experimental designs should theoretically become smaller. As such, one of the key recommendations remain that plenty of attribute levels should be included in the design to spur variation in the presented BVTT values across the sample. This requires offering big and small travel time (and cost) savings to travellers on reference trips of all lengths. The use of pivot designs (which tend to work in percentage changes around the reference trip) assists creating variation in  $\Delta t$  and  $\Delta c$  across respondents. This may, however, conflate  $\Delta t$  with distance and hence some balance will need to be struck between the use of absolute and percentage values to determine  $\Delta t$  and  $\Delta c$ . Additionally, the use of attribute level balance ensures that different combinations are presented roughly equally across the design blocks. In this context, the use of priors, often argued to be central to efficient designs, only ensure that irrelevant (or dominant) choice tasks do not tend to occur.

|             | <b>Design strategy</b>                      | <b>Reference-based?</b>  |
|-------------|---------------------------------------------|--------------------------|
| NL-1988     | Fixed                                       | 6 out of 12 choice pairs |
| UK-1994     | Bradley                                     | Yes                      |
| NL-1997     | Fixed                                       | 6 out of 12 choice pairs |
| Sweden-2008 | Scandinavian                                | Yes                      |
| NL-2009/11  | Bradley                                     | Yes                      |
| Norway-2009 | Scandinavian                                | 8 out of 9 choice pairs  |
| UK-2014     | D-efficient                                 | No                       |
| Norway-2019 | Scandinavian                                | Yes                      |
| NL-2022     | Combination of Scandinavian and D-efficient | 4 out of 8 choice pairs  |

*Figure 8: SP1 design strategies implemented across Europe (replicated from Significance, 2023)*

Whilst the design of SP1 has received extensive attention in the VTT literature, this is less the case for multi-attribute cases such as SP2 (reliability) and SP3 (congestion and crowding). This is a direct consequence of the concept of BVTT losing its relevance in the context of multiple attributes. Namely, increases in costs are now no longer only attributed to reductions in travel time, but also to increases in comfort due to lower levels of crowding. Note that this is also the case when potentially moving to a mode choice study design in future studies, where differences in travel time and cost are also associated with differences in modal preferences (and comfort whilst travelling).

This does not, however, mean that the generic lessons discussed above should not be taken on board. In this context efficient designs strategies are most relevant, as the four quadrants are no longer meaningful. Again, the goal here is to ensure sufficient (but relevant) variation in all attribute levels meaning the presentation of many levels across a wide range. The use of priors can again assist in the presentation of meaningful choices. Additional general principles from the experimental design literature can be adopted in this context (see Bliemer and Rose, 2024).

### 3.2.3 Short vs long term choices

The typical route and mode choice SP formats used in national VTT studies adopt a short-term perspective where the benefits of travel time savings occur at the trip level. The same short-term perspective on the VTT also applies to cross-sectional RP estimates, where the trade-offs between travel time and travel costs are studied in the context of observed residential locations, location of work etc. This partial equilibrium approach to appraisal is also central to most transport models, where typical modelling outputs measure traffic volumes on road links, mode shares, and journey times etc.

Changes in accessibility may nevertheless induce long-term changes in travel behaviour as travellers may consider relocating to other locations, or consider changing work locations. This perspective aligns with Metz (2008)'s observation that in the long run average travel times are conserved but that users experience the benefits by accessing more distant destinations. The incidence and the type of benefits from transport investments are also part of a recently started project by SYSTRA and ITS for the Department for Transport related to distributional impacts. In all, accounting for long-term changes in behaviour on the incidence and magnitude of travel time savings following from transport investment involves a significant change from existing transport appraisal practices and is considered outside of scope of the present VTT scoping study. Additional guidance is required before progressing this in the context of a national VTT study.

Recent work in the VTT literature does consider potential differences in the VTT between short- and long-term choices (e.g. Swärdh and Algers, 2016, Beck et al., 2017, Peer and Börjesson, 2018, Dubernet et al., 2020). Beck et al. (2017) find using data on making trade-offs between income and commuting time in Sweden, that long-run values of time are substantially higher than short-term VTT estimates. This result is consistent with those of Peer and Börjesson (2018) who find higher valuations of trip attributes in the long-run context. Dubernet et al. (2020) delve deeper into these relationships using the German national VTT studies as reviewed in **Section 2.4.1** in this report. Their results contrast those of Beck et al. (2017) in that when salary neutrality is imposed people's willingness to pay to pay for commuting travel time savings through travel costs are lower in the long-run than in the short run. Notably, Dubernet et al. (2020) experience difficulties isolating the VTT from other aspects of monetary compensation, such as income as a status or career progression symbol.

## 3.3 Survey administration

In this subsection we work under the assumption that a future national VTT study makes use of a combination of SP and RP research using short term travel decisions as per current practice. Insights provided in this section incorporate reflections from conversations held with the market research company Accent.

### 3.3.1 Representativeness

It is often argued that sampling of respondents for VTT (and other choice modelling purposes) should be done on a representative basis. This requirement is, however, less important than one may think. That is, the choice models used to estimate functional forms for the behavioural VTT aim to disentangle the influence of separate variables on the VTT. These factors include, amongst others, journey purpose; trip characteristics such as reference time, cost, and distance; and socio-economic characteristics such as income, education, gender etc. To isolate the impact of these variables on the VTT one would ideally sample such that a good coverage of the domain of a given variable is obtained, and that there are limited levels of correlation across these explanatory variables. Unfortunately, the latter is not always possible for variables such as travel time, cost, and distance of the reference trip. To obtain a good coverage of specific variables of interest this may mean that we may wish to oversample short or long

distance trips or specific population cohorts. Ultimately, this will enable better identification of the parameters of interest than when using a fully representative sample.

To put this into context, the sampling approach used in ARUP et al. (2015a) resulted in an SP sample for which the average commute distance was much longer than observed in the NTS. This was partially a direct result from the choice of intercepting car and rail passengers on inter-urban and long distance trips (see Tables 3.13 and 3.14 of ARUP et al., 2015a). Ultimately, the recruited sample should be a good reflection of the travelling population with some selected cohorts being oversampled to avoid empirical identification issues of key parameters of interest. It is not entirely clear whether in the previous national VTT study in the UK sufficient short distance trips were sampled to accurately represent the VTT for short distance trips. This should clearly be reconsidered in a new national VTT study.

Any over- and under-sampling can, to a certain extent, be accounted for in the sample enumeration part after having accurately estimated and derived the behavioural VTT for the sample of interest. The importance of having a representative sample is perhaps even less relevant for the ARUP et al. (2015a) study because for the sample enumeration it makes use of the NTS dataset as opposed of the estimation sample. The Dutch, German, and Norwegian VTT studies all make use of the estimation sample in sample enumeration, thereby making it more important that there is a good reflection of all kinds of trips and the travelling population in the survey sampling. A necessary condition for sample enumeration is that robust VTT estimates can be provided for all trips (and travellers) on which the sample enumeration is conducted.

### 3.3.2 Selection bias and methods of recruitment

The limited importance of representative sampling articulated in **Section 3.3.1** only applies under the assumption that there are no systematic biases in the respondents sampled for a given mode-purpose combination. Prelipean et al. (2018) highlight overall response rates to traditional transport diary collection methods have seen a decrease in most countries. At first sight this may not be considered problematic as it may be overcome by contacting a larger sample, but Halse et al. (2023) express a concern that a selection bias may be present between those who do and who do not respond to survey requests. Particularly internet panels seem to be associated with panellists with relatively low opportunity costs of time thereby tending to produce lower VTT estimates. These effects have been observed in the previous Dutch VTT study (Kouwenhoven et al., 2014) and the Norwegian VTT study (Flügel et al., 2020), but also in other fields of research (e.g. Sandorf et al., 2020, Sandorf et al., 2022).

Alongside long-standing worries about lack of respondent engagement (e.g. Peer et al., 2022), there is growing evidence of dishonest practices by some panel providers, with either “fake” respondents (bots) or “farmed” respondents (professional survey respondents often from third world countries using VPN to mask their ineligibility for the survey) (e.g. Douglas et al., 2023) .

Kouwenhoven et al. (2014) specifically resampled using intercept studies when the results from the internet panel were coming out remarkably low, whereas Flügel et al. (2020) deliberately sampled using a mixture of recruitment methods such that the results could be contrasted across recruitment methods. Moreover, panel respondents were given a lower weight in estimation to correct for this implicit bias. In the most recent Dutch national VTT study lessons were drawn from the preceding study. Somewhat remarkably, Significance (2023) made use of a combination of online (80%) and intercept studies (20%) and deliberately went for a high-quality panel provider, where most notably respondents cannot register themselves but are approached by the panel company. Moreover, study



incentives were increased from the previous study from €1,50 to €3 per completed response, and filtering procedure based on a list of predetermined conditions was applied to the respondent data.<sup>6</sup>

**Figure 9** provides an overview of the recruitment methods used in Flügel et al. (2020). Besides the referred lower VTTs associated with members of internet panels, Halse et al. (2023) also find evidence of higher VTT values for those answering the survey whilst reporting to be on a trip, and whilst answering the survey at the point of interception. The latter effect may be the result of the benefits of travel time savings being more salient to people whilst travelling. If the latter is the case than Halse et al. (2023) consider this as a concern of SP studies closely aligned to the argument of Hess et al. (2020) that certain SP formats, and in this case response locations, may induce particular types of behaviour. As argued in **Section 3.1**, RP data on its own is less affected by these concerns as the collected data relates to actual trips, giving analysts the ability to validate the data from, for example, the GPS traces. However, it cannot be guaranteed that any supplementary SP exercises are void of these concerns.

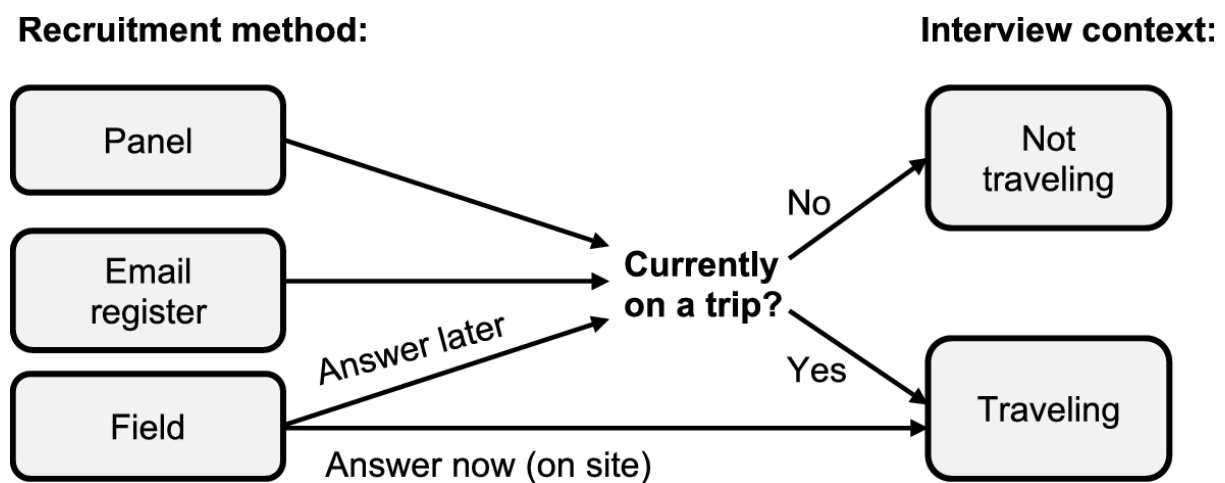


Figure 9: Mix of recruitment methods used in the Norwegian VTT study (Halse et al., 2023)

Where internet panels remain a cost-effective means to collect large samples, the discussion so far highlights that quality concerns remain. There are several ways in which the quality of responses can be improved. In the first place, clearly defined strata need to be defined such that effective quota are in place ensuring that the obtained sample reflects the intended properties of the sample. For example, Significance (2023) selected a stratified sample of respondents based on a pre-question where respondents reported the number of trips made by each mode-purpose combination in the past four weeks. Second, the quality of the sample can be improved by ensuring respondents read all the content presented and not allowing them to proceed until they have done so. Third, it needs to be attractive to be partaking in the survey. On the one hand, this can be achieved by ensuring appropriate incentives for completing the survey. Scheepers and Hoogendoorn-Lanser (2018) and Greaves et al. (2023) argue that increasing incentives indeed increases participation rates, but on the other hand they articulate that increased incentives do not necessarily address the intended biases in sampling. Accordingly, combinations of different incentives are required to ensure participation and buy-in from respondents and subsequent incentive bonuses are necessary to ensure completion of the survey and their continued use for longer term surveys. For example, by making the survey visually attractive and providing a good interactive and engaging experience the level of dropouts can be reduced.

<sup>6</sup> Intercepted travellers received a monetary compensation of €10 in 2022, as opposed to being included in a price draw for €500 used in the second data collection effort using intercepts in 2011.

Researchers often implement ex post quality checks and filter out responses which are considered of insufficient quality. Section 5.2 of Significance (2023) provides a detailed description of all filtering criteria used in the Dutch study. These include checks on survey errors, missing data, and eligibility, but also reported inconsistencies, and implausible reported levels of travel time and travel cost, implicit speeds, and costs per hour. Finally, lexicographic preferences – defined as respondents always selecting the left or the right alternative – are removed and the same applies with unrealistically short response times in the survey and the SP part. With respect to the latter, it should, however, be noted that Sandorf et al. (2020) do not find that response quality is directly correlated with time spend on the survey and the choice tasks.

An alternative to internet panels is to send out invitations through an obtained sample of postal addresses. An example of this in the UK is the Postcode Address File (PAF) which contains an up-to-date address database. This method uses a push to web approach (i.e. the letter includes a link to an online survey). The challenge with using the PAF is that response rates are typically lower than when using internet panels and quotas cannot be used. Moreover, if researchers are interested in non-response rates additional data collection efforts are needed as opposed to receiving this information from survey providers.

Note that high-quality internet panels are available in the UK making use of random probability address-based sampling thereby providing a more representative representation of the population than convenience-based panels. Due to the specific nature of the VTT research it, however, remains important to implement quotas in the survey platform, such that the sample provides a complete representation of the distinct travel patterns to be captured but at the same time avoid selection biases in recruitment.

Intercept samples are consistently used across national VTT studies to recruit respondents. Relative to internet panels, probability-based samples and address-file postal surveys, intercept sampling is more cost intensive, and it can be challenging to recruit respondents on shorter trips. Other methods of recruitment have included computer-assisted telephone and personal interviews. As Halse et al. (2023) note, it cannot be guaranteed that this data collection method is free from selection bias. For all data collection methods other than intercept sampling respondents will need to be allocated to specific mode and purpose segments based on reported recent trips.

Most sampling considerations discussed in this section apply to SP and traditional recall-type RP data collection efforts. We now add some reflection on sampling for RP methods making use of smartphone-based tracking apps. Sampling of respondents for smartphone GPS surveys can be done in the same way as for SP surveys or household travel surveys, and most of the population now have access to smartphones. Nevertheless, there may be some selection bias present across population cohorts in relation to their willingness and ability to install and operate these specific applications. Cellina et al. (2021), however, do not find evidence of self-selection effects in a study on using apps to promote public transport and active travel. This is, however, not the case for Lynch et al. (2019) who find increased participation of young, and low income groups for using an all-in-one app survey and GPS tracking setup relative to a pre-tracking questionnaire followed-up by the GPS tracking exercise. Since this type of survey format has not been implemented in large scale VTT studies it is highly recommended to conduct some tests in terms of selection bias. Note that the quality of the GPS data typically varies by phone type and hence this may correlate with specific user characteristics.

Numerous apps for such data collection exist, are easily customisable, and have the option of incorporating additional SP components. By nature, the number of trips recorded and the level of engagement (length of time tracking) will differ across respondents potentially resulting in a rather



unbalanced data structure. Molloy et al. (2023), for example, highlight that over a period of eight weeks 20% of their sample drop out, particularly in the last week. As such, it is highly important to test the optimal length of the recording period, especially when this exceeds more than one week. We are not aware of any research that have investigated the link between tracking app-engagement, selection biases and the provided incentives for study participation. A trade-off between the amount and richness of data, the need for higher incentives, and the risk of self-selection is therefore required and warrants further research.

Overall, this section concludes that irrespective of the data collection format selected significant effort needs to be put into sample recruitment to arrive at a high-quality sample enabling the analyst to make inferences regarding the travelling population of interest.

### 3.4 State-of-the-art estimation of VTT

The econometric methods used in ARUP et al. (2015a) to analyse the choices made in the SP games can be considered as state-of-the art at that moment in time. The key modelling features used in the respective study are summarised in Hess et al. (2017a) and builds on methods used in previous national VTT studies in Denmark, the Netherlands, Norway and Sweden. In this section we will go through some of the key features included in the modelling and discuss related developments since ARUP et al. (2015a).

#### 3.4.1 Random Utility vs. Random Valuation

In the VTT literature two alternative approaches exist with respect to the modelling of the route choices made in the SP games, respectively the Random Utility (RU) and the Random Valuation (RV) approach (Ojeda-Cabral et al., 2016).

RU assumes that individuals select the route generating the highest level of utility and that the deterministic part of utility of a given route is determined by its attributes (i.e. travel time, travel cost, etc.). The random part of the utility function is given by either an additive or multiplicative error term (see **Section 3.4.2**). The RU-formulation can easily be applied across the different SP-games because the utility function can easily be expanded for the inclusion of additional attributes, such as congestion, crowding or reliability, describing the conditions of travel. In RU, the VTT is determined by the ratio of marginal utilities for travel time and travel cost. In many empirical applications, the utility function is rewritten in such a way that the VTT is directly estimated such that its standard error is also directly obtained in estimation. This functional form is often denoted as 'WTP-space' (Train and Weeks, 2005).

RV only applies in the context of SP1, i.e. the simple time vs cost trade-offs, and makes extensive use of the underlying notion of the boundary value of travel time (BVTT). Respondents are assumed to select the fast and expensive option when their personal VTT is larger than the presented BVTT, and conversely the slow and cheap option when their personal VTT is lower. In contrast to RU, RV assumes that the BVTT and the VTT themselves are associated with a degree of randomness. The slight difference in interpretation of the VTT in RU and RV makes that the VTT concepts underlying both approaches are distinct and cannot be directly contrasted on a like for like basis. This forms part of the explanation why different VTT measures emerge from both approaches (see for example Table 14 in Significance, 2023).

The choice for RU and RV is therefore an important decision as part of the modelling strategy. Notably, the UK and Norway have adopted the RV specification whereas Germany and the Netherlands adopted the RU specification highlighting that the jury is out on what can be considered as the 'best' approach. In determining what is best different criteria can be used.

Econometric performance, or model fit, is often used as a key determinant in this context and the general conclusion is that RV is associated with a better model fit (e.g. Ojeda-Cabral et al., 2016, Significance, 2023). A second observation is that RV is generally associated with a lower VTT estimate than RU (ARUP et al., 2015a, Ojeda-Cabral et al., 2016, Ojeda-Cabral and Chorus, 2016, Significance, 2023). The most likely explanation for the difference in size is that when the RV model is translated back into RU, as set out by Ojeda-Cabral et al. (2016), the variance of the error term is proportional to the change in travel time. Not accounting for such heteroskedasticity in RU models may inflate the VTT estimates.

Whilst econometric performance has often been leading in the choice of model specification, some reconciliation is taking place in relation of the consistency of such model specifications (both in RU and RV) with welfare economics (see Box 1 and Batley and Dekker, 2019, Dekker et al., 2022, van Cranenburgh et al., 2023). This is especially relevant when the notion of the VTT has a strong rooting in economic theory (e.g. Becker, 1965, DeSerpa, 1971, Jara-Díaz, 2007) and its application forms a critical input in transport appraisal in most Western-European countries. The fact that the VTT difference between RV-approach and the RU-approach are not sufficiently understood was one of the reasons why Significance (2023) did not depart from the RU specification adopted in (Kouwenhoven et al., 2014).

The choice between RU and RV therefore remains a relevant area of research. Alternative model specifications can easily be contrasted once the data have been collected and hence this has limited impact on study design. Ojeda-Cabral and Chorus (2016), however, do present some relevant conclusions which need to be kept in mind when linking modelling results with experimental designs. They put particular emphasis on the role of the time differences, i.e.  $\Delta t$ , presented in SP1. In **Section 3.2.2** we already established that high degrees of variation in the attribute levels, including  $\Delta t$ , presented are important for accurately estimating the relationships of interest. This is particularly the case for non-linear relationships and for catching the tail of any distributions associated with random parameters (**Section 3.4.4**). In this context, Ojeda-Cabral and Chorus (2016) show using simulations using respectively RU and RV as the ‘true’ decision rule that the model with the model best fit is always the one corresponding to the ‘true’ decision rule. Whilst this is reassuring, the more worrying conclusion is that with large variations in  $\Delta t$  the incorrectly specified model, e.g. using RU when the true behaviour is RV, incorrect estimates of the VTT are obtained.

Indeed, the variation in  $\Delta t$  implemented in national VTT studies can be seen as an explanation for the differences in the VTT observed between RU and RV models. Moreover, using model fit as the decision criteria will support selecting the (most) correct behavioural mechanism underlying the choices.<sup>7</sup> If, however, the decision is made that RU models need to be used for consistency with welfare economics, this may result in substantial differences in recommended VTTs. Ultimately, these decisions need to be made in conjunction with the Department for Transport or other relevant authority. We do, however, recommend that these discussions and associated implications are explicitly considered.

### 3.4.2 Additive vs Multiplicative error terms

Where **Section 3.4.1** discussed the specification of decision rule (RU vs RV), this section focuses on the specification of the error term in RU and RV models. Traditionally most choice models include an additive error term in the utility function and assume it follows an independently and identically distributed extreme value type-I density such that the familiar logit choice probabilities emerge

---

<sup>7</sup> We will never know the true decision rule.

describing the probability that a respondent selects a particular route in RU, or that a respondents VTT is higher or lower than the presented BVTT in RV (Train, 2009).

Fosgerau and Bierlaire (2009) introduced an alternative model specification where the error terms are no longer treated as additive, but as multiplicative. One of the benefits of the multiplicative error terms specification is that the utility variance increases with utility. This is particularly convenient when long- and short-distances are included in the same model because with increasing travel times and travel costs the additive error term would assume that the model would become increasingly deterministic over longer distances. To operationalise the multiplicative error term specification a log-log form is adopted such that the transformed additive and deterministic component remain additive. Assuming that the underlying error term is log-extreme value distributed one again arrives at the well-known logit probabilities (Hess et al., 2017a).

Due to the log-transformation, Significance (2023) refer to the multiplicative error term as the RU-log and RV-log (vs RU-lin and RV-lin as additive) specifications. One of the challenges in estimating RU- and RV-log models is that negative arguments in the logarithm are not allowed, which is not always straightforward with more advanced model specifications. In line with ARUP et al. (2015a), Significance (2023) find significant improvements in model fit when switching from additive to multiplicative error terms both in the context of RU and RV with RV-log providing the best model fit.

The RV-log model results in further reductions in the VTT estimate in Significance (2023). ARUP et al. (2015a) acknowledge that the RV-log is likely to result in an underestimate of the true VTT and that the error term in this model specification is likely not to only capture noise, but also heterogeneity in the VTT, and that this is not easy to adjust for. The latter point is corroborated by Dubernet (2019) in the context of the German VTT study. A lot of variables included in the RV-lin specification are no longer significant when switching to the RU-log specification.

Significance (2023) end up using a RU-lin model specification (i.e. additive error terms), because they limit the discrepancy in the VTT across both alternative model specifications by controlling for relevant control variables, in particular size effects (travel time and cost) of the reference trip. This enables them to adopt a specification which is consistent with Kouwenhoven et al. (2014) whilst avoiding the numerical issues associated with estimating RU-log models.

The Norwegian study makes use of the RV-log model specification but no explicit discussion is included regarding the performance of alternative model specifications (Flügel et al., 2020). The same applies to the German study, where a RU-lin model specification is adopted with linear and log specifications of the key explanatory variables (Axhausen et al., 2014). It is only in the PhD-thesis of Dubernet (2019) that the multiplicative error terms are tested on the data. In this case the VTTs for the multiplicative error term are again tending to be lower, especially for business, but the differences are not as distinct.

As discussed in **Box 1**, the adoption of multiplicative error terms is also associated with the debate regarding model fit vs. theoretical consistency. In discrete choice models analysts estimate indirect utility functions, which represent the optimal decision for rational decision makers maximising their direct utility function subject to their budget constraint. Batley and Dekker (2019) highlight that the estimated indirect utility function is a combination of the utility derived from the selected route and the outside (or numeraire good). In an additive utility specification, it is feasible to separate these two components and attribute the error term to the selected good, but using the assumption of a multiplicative error term again the nature of the error term becomes rather ambiguous unless very strong interaction effects are assumed between the selected route (or mode) and the outside good (Dekker et al., 2022). Alike the RU vs RV debate, the model structure which tends to provide a better

model fit is therefore not necessarily consistent with welfare economic theory. Again, we recommend that informed decisions need to be held as part of the analysis and note that RU- and RV-lin are more consistent with economic theory than RU- an RV-log.

#### **Box 1: Consistency with economic theory**

The assumption that travellers are utility maximisers by default assumes that they are maximising their utility subject to a budget constraint and a time constraint, where utility emerges from the consumption of goods and services and performing activities. In this context travelling is associated with two forms of opportunity costs, respectively money and time which could have been spent on other goods, services and activities. The VTT in its simplest form is then defined by the ratio of the marginal utility of time and the marginal utility of income, i.e. the ratio of the Lagrangean multipliers for the two constraints.

In additive RU choice models these two constraints are implicitly modelled in the indirect utility function, such that the utility of route  $j$  is given by  $V_j = \Psi_j + \lambda(I - p_j) + \theta(24 - t_j) + \epsilon_j$ , where  $\Psi_j$  represents the utility of the route,  $I$  income,  $p_j$  travel cost, and  $t_j$  travel time (in hrs), with  $\lambda$  and  $\theta$  capturing the marginal utility of income and time respectively. Finally,  $\epsilon_j$

This model specification makes clear that indirect utility is a composite utility function of the utility of the route itself and utility derived from spending money on other things ('residual income') and spending time on other things ('residual time').

By working in the context of discrete choice, Marshallian and Hicksian demand levels need to be restricted to one (i.e. unity), and as such Roy's identity imposes restrictions on the shape of the modelled indirect utility

with respect to income and prices, such that  $-\frac{\frac{\partial V_j}{\partial p_j}}{\frac{\partial V_j}{\partial I}} = 1$ . For this to hold indirect utility needs to be specified

in 'residual income' terms. Furthermore, Shepherd's lemma requires  $\Psi_j$ , i.e. the utility of the route, to be independent of prices and income. Any price and income effects can thus only operate through the numeraire good, i.e. money spend on other goods and services.

A shift from linear RU to linear RV is not problematic for consistency with economic theory since the implicitly imposed heteroskedasticity of the error term of the indirect utility function only relates to time and not to income (or prices). However, in both cases when shifting from a linear RU/RV specification to a logarithmic RU/RV specification – i.e. imposing a multiplicative instead of an additive error term) things quickly become increasingly tricky in terms of satisfying the economic conditions for valid demand functions, and hence calculation of welfare effects.

Dekker et al. (2022) highlight that for consistency with economic theory in log RU/RV models the indirect utility function needs to be multiplicative between  $\Psi_j$  and residual income, i.e. very strong interactions effects are imposed which are not intuitive, and due to this multiplicate nature of indirect utility reformulating RV models based on the BVTT into RU is non-trivial.

NB: The shape of the indirect utility function is much less restrictive with respect to time, because the disutility can come from both the opportunity cost element (VoL) and the (dis)comfort whilst travelling (VTAT), whereas for income such effects can only operate through the numeraire as explained above.

#### **3.4.3 Reference dependence or size and sign effects**

The third modelling element implemented in ARUP et al. (2015a) is reference dependence. This is also known as 'sign' and 'size' effects in the literature (De Borger and Fosgerau, 2008, Börjesson and Eliasson, 2014) and originates from the loss aversion dimension embodied in Kahneman and Tversky (2012)'s prospect theory. Reference dependence considers the possibility that people have an aversion to increases (or losses) in travel time and travel cost relative to the reference trip. The sign effect thus

considers that losses are weighted more than travel time and cost gains. Size effects assume that the impacts of gains and losses are non-linear. Hess et al. (2017a) find that the marginal disutilities decrease and that time and cost damping appears. Specifically, this effect is larger for cost than time such that the VTT is increasing with costs (Rich and Mabit, 2016).

De Borger and Fosgerau (2008) and ARUP et al. (2015a) articulate that a reference free VTT can be recovered from a model estimated on reference-dependent choices. This can be achieved by taking a geometric mean such that the reference-free VTT is sign neutral and can be applied to travel time savings and losses. This reference free VTT then forms the basis for the calculation of the appraisal VTT in the sample enumeration process. The non-linear size effects can, however, not be circumvented such that in appraisal an arbitrary decision for the value of  $\Delta t$  needs to be selected. **Figure 10** shows the impact of  $\Delta t$  on the recommended VTT and that this impact is non-trivial. The selected value of 10 minutes is consistent with those selected in other countries such as Denmark and Sweden (Hess et al., 2017a). In similar vein, Flügel et al. (2020) make use of a  $\Delta t$  of 10min for short distance trips and 15min for long distance trips (see Table 7.4 for the respective impact on the VTT).

Notably, the German and Dutch national VTT study do not make use of reference dependence in their modelling specification (Axhausen et al., 2014, Significance, 2023). The Dutch study did control for size and sign effects but did not find differences in the VTT. Outside of the national VTT studies, we have only identified the use of reference dependence (and multiplicative error terms) in Lu et al. (2021), where again no significant size and sign effects were identified in more advanced model specifications, perhaps due to the use of smaller time and cost differences in their empirical applications.

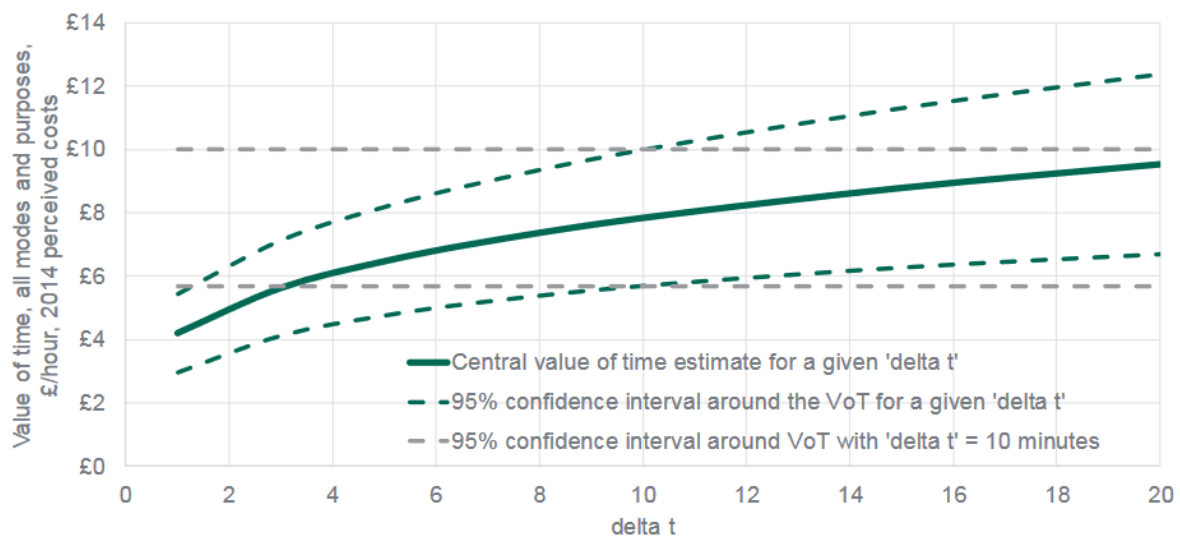


Figure 10: Impact of  $\Delta t$  on the recommended appraisal VTT. Source: (Department for Transport, 2015a)

Reference dependence by default refers to a behavioural phenomenon and Hess et al. (2020) highlight that such effects may be amplified in the simple time vs cost setup of SP1.<sup>8</sup> Aside from the impracticalities and ad hoc nature of selecting appropriate values for  $\Delta t$  in application, the use of reference dependence is also at odds with welfare economics. The notion of reference dependence by default falls in the domain of behavioural economics which is at odds with the assumption of the rational agent fundamental to the discussion in **Box 1**, and the framework of Kaldor-Hicks compensation tests underlying CBA in general (see also Sunstein, 2020). Adopting such behavioural

<sup>8</sup> Significance (2023, p73) argue that this is not entirely consistent with the previous Dutch VTT study.

traits in policy evaluation, including the VTT, is a choice and discussions need to take place when developing the model specification during a new national VTT study whether this is desirable. These discussions are similar to the adoption of RV or RU, and additive or multiplicative errors.

#### 3.4.4 Unobserved preference heterogeneity

Nowadays the default approach in choice modelling seems to be to account for unobserved heterogeneity in the form of latent class or mixed logit models (Fosgerau, 2024, Hess, 2024). In the context of VTT studies the default option is the use of mixed logit models where the unobserved heterogeneity is captured by means of a continuous distribution. The only exception in this context is the German VTT study which relies on standard multinomial logit models in the main study report (Axhausen et al., 2014), but Dubernet (2019) has conducted analyses using a latent class structure on the same data.

ARUP et al. (2015a), Flügel et al. (2020), and Significance (2023) all make use of the mixed logit specification, but the choice of distribution varies across the three recent national VTT studies. ARUP et al. (2015a) specified the baseline VTT parameter as random and made use of the log-uniform distribution. The choice for the log-uniform density was informed by two aspects. First, the 'log' element ensures that the VTT is always positive as per the expectation of economic theory. Second, the uniform instead of the more conventional normal density was adopted to avoid the long tail associated with the latter density (Börjesson et al., 2012b, Hess et al., 2017a). In contrast, both the Norwegian and the Dutch VTT study both adopt the lognormal density on the VTT. Significance (2023 Section 6.4) make the decision for the lognormal using an extensive non-parametric analysis on the SP1 data, highlighting that the differences between the top ranked distributions are relatively small. In addition to the VTT, Significance (2023) also specify the value of travel time reliability (VTTR) as random and introduce correlation between the respective distributions and find positive levels of correlation.

A specific development in the context of VTT work have been made by van Cranenburgh and Kouwenhoven (2021), who demonstrate that an artificial neural network (ANN) – a specific form of machine learning methods - can be used to uncover the VTT distribution from simple time/money trade-offs (i.e. SP1). This work remains, however, limited to such simple trade-offs at this point in time, and therefore its implementation in national VTT studies is perhaps limited to inform the shape of the selected parametric distribution at this stage.

Where the referred ANN approach adopts a fully non-parametric approach to unobserved preference heterogeneity, Train (2016) introduced an alternative semi-parametric approach to flexible mixing densities, the logit-mixed-logit model. Scarpa et al. (2021), however, highlight that this approach is only beneficial at sample sizes larger than usual and with some nuances. Overall, the use of parametric distributions in the final specification is preferable moving forward and fully non-parametric (ANN) and semi-parametric approaches can be used to inform the choice of density used. Fosgerau and Mabit (2013) provide a simple method using polynomial transformations of base distributions which can assist in this preference. Ultimately, the choice of distribution and the impact it has on the VTT remains an empirical matter.

#### 3.4.5 Observed preference heterogeneity

ARUP et al. (2015a) adopted a very distinct approach to including explanatory variables. Namely, all variations in the VTT by observable characteristics were estimated using multipliers. Note that this functional form makes the rather strong assumption that multiple explanatory variables enhance each other's effect through the multiplication of their effects. In contrast, the German and Norwegian studies adopted a more standard additive functional form, whereas the Dutch models followed the multiplicative form proposed by ARUP et al. (2015a).

In any forthcoming study, the decision about the functional form and which explanatory variables will be included will largely be an empirical matter. Machine Learning (ML) techniques are ideally suited to support this task, with a great ability to find patterns in behaviour. They can in this way be used to help inform the specification of econometric models, building on existing work on ML assisted specification of choice models (Ortelli et al., 2021). It is, however, not recommended to shift to ML based specifications for several reasons.

Ultimately, the final models need to remain parsimonious, policy relevant and based on variables which can easily be accessed through, for example, the NTS. Moreover, ideally restrictions are imposed on the functional form to adhere with the assumptions of rational economic agents. ML techniques do not (yet) work well with such restrictions. Recent efforts show that economic insights can be obtained from ML models, such as in the work of Wang et al. (2020). However, the measures obtained are not context independent, which causes problems comparable to the introduction of reference dependence as discussed in **Section 3.4.3**.

### 3.5 Summary of development in surveying and estimation methods

To bring this section to a close, we provide a quick summary of the main findings identified. Firstly, the comprehensive review by Flügel et al. (2022) highlights that there is increasing potential for the use of RP methods in future VTT studies. The most preferred RP option would be the use of GPS tracking combined with a background survey, and potentially supplemented with a limited set of SP choice tasks to capture components unlikely to be captured by the RP component. This option is preferred over more passive data sources, such as mobile phone data, because of its accuracy and its ability to capture a range of explanatory variables and supplementary SP choice tasks expected to be included in behavioural VTT models. The collection of such information is not feasible with mobile phone data and similar data sources.

Moving to an RP-based data collection method would mean a significant departure from existing practice, especially when shifting from a within mode route choice to a mode choice approach. It is therefore not expected that SP will be entirely abandoned soon for consistency purposes, and extensive testing will need to be done, also in terms of potential sampling selection and attrition biases associated with the method.<sup>9</sup> Other national VTT studies do not offer precedent for this at this stage.

In the context of SP research, progress has been somewhat limited since the publication of ARUP et al. (2015a). Primary concerns regarding the quality of SP data remain, and particular attention needs to be paid to the quality of any online panel to be used, as many internet panels are associated with a selection bias. It is becoming increasingly harder to include travellers with high values of time in such panels (and surveys in general) and therefore internet panels tend to be associated with lower values of time. With appropriate quality controls these effects can be mitigated, see for example the most recent Dutch VTT study. Preferred methods of surveying remain intercept surveying and the use of postal addresses, where the latter method requires assigning travellers to specific mode-purpose segments based on a selection of recent trips.

In terms of sample sizes, there is no need to increase sample sizes. The sample sizes used in the last UK national VTT study are comparable with those in other countries. However, in these other surveys respondents typically are not presented with all SP games, as opposed to ARUP et al. (2015a), resulting in a lower numbers of choice tasks per respondent per SP game. This affects the precision by which specific parameters can be estimated and this is something which cannot be controlled for by the

---

<sup>9</sup> Any move to RP data, and especially large-scale RP data, will create a need for extensive data analysis work prior to modelling analyses. ML techniques are ideally suited for this task, with a great ability to mine data.



experimental design. Overall, a priori definition of relevant strata to be sampled is of vital importance to avoid unintended correlations between socio-economic and trip characteristics although this cannot be entirely avoided due to inherent correlations between travel time, costs, and distance. We argue that in this context representativeness of the sample is not essential.

Most research on SP design and analysis has focussed on the simple travel time vs travel costs presented in SP1. We cannot ignore the fact that despite their dominance in national VTT studies, this format is associated with a certain degree of critique due to its simplistic nature. In generating new experimental designs a balance needs to be struck between behavioural relevance to the respondents and statistical information. Where ARUP et al. (2015a) presented realistic choices to respondents, Significance (2023) induce some extreme trade-offs to force respondents to not always select a similar type of journey. Some useful lessons can be drawn from how the latter SP design was generated. If a future design moves away from the SP1 format, D-efficient designs are the go-to option and a wide literature is available outside the VTT literature to draw valuable lessons from.

In terms of data analysis, ARUP et al. (2015a) raised the bar for other national VTT studies and, as reflected by **Table 17**, there is a large variation in the adopted functional forms since. In determining the functional form two key principles should be balanced, respectively model fit and theoretical consistency. We highlight that many of the econometric features implemented in ARUP et al. (2015a) are not consistent with economic theory (see **Box 1**) but widely applied in the literature and typically associated with a better model fit. Unfortunately, no clear solution to this conundrum is available and additional research and open discussions are necessary in this context.

The good news is that apart from options to identify more flexible distributional forms, and the ability of machine learning techniques to assist in model specification (both in observed and unobserved preference heterogeneity), no significant developments have been taken place calling for a radically different approach to the analysis of collected SP data.

*Table 17: Model specification characteristics in recent national VTT studies*

|                            | UK 2014/15                   | Germany                | Norway                       | Netherlands            |
|----------------------------|------------------------------|------------------------|------------------------------|------------------------|
| RU vs RV                   | RV                           | RU                     | RV                           | RU                     |
| Additive vs multiplicative | Multiplicative (i.e. RV-log) | Additive (i.e. RU-lin) | Multiplicative (i.e. RV-log) | Additive (i.e. RU-lin) |
| Reference dependence       | Yes                          | No                     | Yes                          | No                     |
| Unobserved heterogeneity   | Loguniform density           | No                     | Lognormal density            | Lognormal density      |
| Observed heterogeneity     | Multiplicative               | Additive               | Additive                     | Multiplicative         |



## 4. Segmentation of the Value of Travel Time

VTT for UK appraisal is, at present, typically segmented by:

- Trip purpose: business, commute, and non-work other
- Mode: only for business trips
- Distance: only for business trips by car and rail
- Trip stage and quality: for access/egress to public transport by active travel modes a weight factor of 2 is applied; multipliers for wait time, lateness and crowding are also included in guidance (weight factors relative to the VTT)

Source: Department for Transport (2022c) and Department for Transport (2024j)

In transport appraisal worldwide, these are indeed some of the most common dimensions of variation for VTTs. Some dimensions are user-specific (e.g. income), whereas the ones presented above are more trip-specific. Throughout this chapter, the focus is placed largely on the above dimensions (purpose, mode, distance) and income plus a few others, such as travel time and travel cost. The segmentation challenge comprises theoretical and empirical issues, and often the main problem is arguably the complex interrelations that exist among dimensions, especially due to empirical confounding of user- and trip-effects.

**Sections 4.1 to 4.3** review the overall current state of play with respect to segmentation in UK modelling and appraisal practice, focusing on the main user groups categories in use and the current and outstanding issues. The discussion is mainly framed within the context of relevant policy areas for the near future. **Sections 4.4 and 4.5** expand the discussion on two particularly contentious and important areas for segmentation: mode (**Section 4.4**) and income (**Section 4.5**). **Section 4.6** draws conclusions for an upcoming VTT study.

### 4.1 User group and trip purpose segmentation in appraisal

This subsection reviews whether fewer or additional segmentations by trip purposes or user groups are desirable. We evaluate the appraisal implications and the alignment of more granular user group segmentations with widely applied transport models (at different geographical scales).

#### 4.1.1 UK-appraisal guidance with respect to segmentation by trip characteristics

The segmentation adopted in TAG follows the empirical findings from the 2014/15 study reasonably close (Batley et al., 2019). This is particularly the case for segmentation by trip purpose (discussed here), and for mode (discussed in **Section 4.4**). Only some of the original recommendations regarding segmentation by distance from Batley et al. (2019) found their way into appraisal guidance, and a similar situation occurred with respect to income, which is discussed in **Section 4.5**.

Regarding trip purpose, the 2014/15 study found significant differences in the VTT across the three types of trip purposes considered, respectively being business, commute, and non-work. This segmentation mirrored existing practices in TAG (pre-2015) and therefore no change in trip purpose segmentation, for appraisal guidance, was recommended. The new values by trip purpose were ultimately adopted. The recommended values were derived through a process of sample enumeration on the NTS dataset as described in **Section 2.2**.<sup>10</sup>

---

<sup>10</sup> The sample enumeration process can be summarised as: (i) Data from the NTS for 2010-2012 were used to obtain a representative sample of journeys in Britain; (ii) For each trip in the sample the behavioural VTT was estimated using the appropriate model and co-variates from the NTS trip and person data; (iii) to growth from

For both business and non-work trips significant elasticities of VTT with respect to distance were found in the behavioural modelling. At the same time, significant elasticities of VTT to reference time and reference cost were also found (Hess et al., 2017a). These findings, though not entirely surprising, led to a debate about the interpretation of the results (e.g. whether a 'pure' distance effect is captured) and the role of distance segmentation in appraisal. More precisely, most of the 'apparent' variation with distance stems from the elasticities for reference time and reference cost (for all modes except for bus) due to high levels of correlation across these three trip characteristics.

Ultimately, after further modelling, the study recommended distance segmentation (VTT increasing with distance), but arguably this was not a straightforward matter. Using the sample enumeration process on the NTS data gives individual VTTs for every trip. These were then banded by distance and trip purpose and averaged using distance weights. These were considered robust enough to recommend adoption into TAG for all trip purposes: business, commute and other non-work (see Sanders et al. (2015 Table 7.28 p245). Much of the debate at the time of the study is encapsulated in Section 7.6.4 "*Further considerations relating to income and distance*" of ARUP et al. (2015a).

The Department recommended adoption of disaggregation by distance for the business values, but not for the commute and other non-work values. The grounds for not adopting disaggregation by distance for commute and other non-work were not made explicit by the DfT in its consultation document, where it identified further work was required to segment by distance (Department for Transport, 2015b pp27-28). One potential explanation is the Department's decision not to segment the VTT by mode for commute and other non-work (see **Section 4.4**) possibly making the decision to further segment by distance across all modes contentious. Although significant variation in the VTT across all modes was found for commute and other non-work purposes, the results were interpreted as problematic for appraisal practice due to the likely confounding between pure mode effects and remaining user effects (e.g. income, self-selection) - in particular, the main reason was the concerns around equity due to remaining income effects. Another potential explanation refers to the uncertainty introduced into the values via the sample enumeration process that indicated a 95% confidence interval of +/-60% for other non-work, though the commute values had the same level of confidence as the business at +/-25%, which when applied at the different distance bands led to overlapping confidence intervals between bands.

Segmentation of VTT by distance/trip length is common in some countries (e.g. Norway, Sweden, Switzerland), where typically VTT increases with distance (with a few exceptions, though). Often, distance segmentation is implemented by bands. For instance, below/above 100km in Sweden, and three bands in Norway (thresholds at 70km and 200km). No appraisal guidance, to the best of our knowledge, segments explicitly with the reference time and cost of a trip. Arguably, distance segmentation may be more practical – although there are potential issues which are discussed throughout this chapter.

We need to ask ourselves what the behavioural reasons are why people may be willing to pay more to save time on a long-distance journey relative to a short distance one. Mackie et al. (2003 p.27) make a distinction between so-called 'budget effects' and 'relative effects', where the former identify that the marginal disutility of travel time and travel costs go up as the opportunity costs of travelling are increasing with distance. Since the two marginal utilities have opposite effects on the VTT, increases in the VTT with distance under pure budget effects imply that the impact of distance on the marginal

---

the sample to a national level trip weights were applied; (iv) distance weights were then applied to get a distance weighted average; (v) to convert to appraisal values the business values were factored by the market price conversion factor.

(dis)utility of income is smaller than on the marginal (dis)utility of time. The latter, relative effects, are not consistent with economic theory and acknowledge that small increases in travel time and costs are of less concern when these are relative to already large amounts of travel time and cost. On the one hand, the VTT may go up for higher cost trips due to “cost damping” (the sensitivity towards £1 goes down as total cost increases). On the other hand, the VTT may go down if reference time is high, due to the analogous idea of “time damping” (see Ojeda-Cabral et al., 2018). However, as noted in the Section 7.6.4 of ARUP et al. (2015a), we do not observe variation for the same individual, but across different travellers. Therefore, it was concluded that the reference time and cost elasticities can pick up self-selection effects: we observe some individuals on long distance trips, and other – potentially different – individuals on short distance trips. The findings reported significant coefficients of similar magnitude but opposite signs (and highly correlated). Similar results were found on a more recent examination of the 2003 data (Ojeda-Cabral et al., 2018). Whether the VTT increases with distance depends on the magnitudes of these two sources of variation, combined with any additional variation due to distance itself.

Note also that, additionally, there will be other factors that also support a higher VTT for longer journeys, such as the 24h/day time constraint and its interplay with other costs (e.g. the possibility to avoid the costs of an overnight stay), although these are more relevant for business trips.

Although the observed correlation across distance, time and cost of the reference trip limits the extent to which their influence on the VTT can be estimated empirically, in terms of implementation all three effects are accounted for simultaneously in obtaining behavioural values for the VTT for specific trips present in the NTS data *before* sample enumeration. As such, the above discussion may not be too much of a concern as the net effect of distance, time, and cost on the VTT is captured. Distance, however, has a separate role in the sample enumeration process where higher weights are attributed to longer distance trips, primarily for the increased chance of being exposed and therefore the opportunity to benefit from travel time savings on such trips.

#### 4.1.2 User group and trip purpose segmentation in appraisal – a look forward.

From an appraisal perspective, ideally we would prefer to disaggregate VTT by any statistically significant covariates to ensure that the willingness to pay for travel time savings is robustly estimated. If the key appraisal parameter values are biased, then project selection will also be biased. In practice this idealistic position must be tempered to: i) ensure that the appraisal is proportionate to the task or decision in hand and, ii) align to the desired position on equity. Thus, we only wish to vary VTT if it makes a difference in an appraisal, and we need to be wary of equity implications.

The point of proportionality is best illustrated with an example. Think about a city centre urban realm project with active travel measures. We speed up walking and cycling journeys and impose some delays on car users. Here average reference time, reference cost and trip lengths by car are short. Therefore, the VTT for car users would be less than the standard ‘average distance’ value. For pedestrians and cyclists reference time, and reference cost are much larger, though distance remains short. Here our standard values of non-work time are possibly too low, but certainly higher than those of car users. Thus, we have a situation where for urban realm projects the dis-benefits to road users are possibly over-valued and the benefits to active travel users are possibly under-valued (all else equal, and in particular omitting any ‘positive’ utility of walking or cycling per se) – See also **Chapter 7** of this report. For similar reasons projects which primarily benefit short distance trips (e.g. roads projects in cities), are likely to experience a systematic overestimation in benefits using a standard ‘all distance’ VTT. Conversely, the benefits of improvements to transport links that carry a large proportion of long-distance travel, whether by train or car, will be systematically underestimated if using a

standard 'all distance' VTT. However, note also that this problem is linked to having a single VTT across all modes. The problem might go away if there was robust evidence on walking and cycling VTTs and their corresponding segmentation with distance. Nonetheless, what this suggests is that considerations of distance effects must consider the role of varying speeds across modes, which affect the overall correspondence of distance-time-cost in the absence of mode-specific values.

Thus, we can see that a lack of disaggregation by distance in VTT for non-work trips risks biasing project selection following appraisal. Furthermore, and as Börjesson and Eliasson (2019) point out, the true effect of income on VTT is much smaller than would be suggested if other covariates, such as distance and trip purpose, are not controlled for. Whilst this is something we will discuss further in **Section 4.5**, it emphasises the point that disaggregation of the VTT values improves appraisal and avoids systematic biases. Though, of course, there may be practical limits to the number of covariates that an appraisal system could cover simultaneously.

In looking forward we should also consider the types of projects that appraisal may need to consider and whether new or amended user group or trip purposes may be needed. To an extent this could be viewed as crystal ball gazing, but we already know that there are certain pressures that the transport system is responding to and that these will continue. An obvious change, partly technologically driven, is the work from home culture and only part-time attendance at the office. This has created a different profile of peak demand on the transport system, something the appraisal system needs to accommodate. Climate change is also an important underlying policy driver, that combined with the increasing role of cities in our economy lead to urban mass transit, active travel and urban realm projects being very much in the fore in most city plans. Furthermore, potentially there may be more emphasis on improving journey quality by road rather than road capacity. Taking these in turn.

**Work from home.** Those who choose to work from home may *ceteris paribus* have higher VTTs. They therefore choose to work from home as they dislike the commute. On the other hand, others who also choose to work from home may do so for affordability reasons (low income), and may thus present a lower VTT. This could be driven by existing covariates such as reference time, reference cost, distance, income, family status, etc. Future demand models will be predicting the impacts of work from home, but it will be important to understand if from an appraisal perspective these workers have the same commute VTTs as those who go into the office or if they differ *ceteris paribus*. This could be tested in a future VTT study. There is also an important connection with transport modelling, especially if differences in work from home opportunities relate to the VTT and thereby systematically affect the types of users (and frequency by which) use is made of transport infrastructure provisions.

**Investment in urban mass transit.** Here the journey quality multipliers on the base VTT, such as walk, wait, headways, crowding, etc., are likely to become more important than they are now. These will vary by user group: young, old, families, etc. Currently, appraisal trip purposes do not distinguish by these demographics, but the sample enumeration process would be able to account for such variations. Thus, if significant variations in journey quality multipliers by demographic were found these could be aggregated into trip purpose averages during the sample enumeration. Also, large investments that introduce a new form of mass transit (e.g. new tram systems or routes) will attract demand from other modes (bus, car, etc.). Segmentation by mode that picks up mode VTT differences (*ceteris paribus* for a given user group) would become essential to understand the potential for mode shift and its benefits. Indeed, a key omission in previous studies has been data for the same group of individuals about *different* modes. This was the missing piece of the jigsaw that would have enabled the separation of mode effects for more refined mode segmentation (**Section 4.4** covers this).

**Urban realm and active travel.** As illustrated by our introductory example above, a key issue is the potential overestimation of disbenefits to short distance car trips and an underestimation of the benefits to walking and cycling trips which have higher reference times. Disaggregation by distance for non-work trips is therefore required. This also interrelates with the modal discussions in the following section. It must also be noted that VTTs for walking and cycling as main modes of travel were not estimated successfully in the 2014/15 study, due to well-known difficulties for SP in those contexts (e.g. setting a realistic monetary cost for time-cost trade-offs) – see **Chapter 7**.

**Improvements in road journey quality.** These are likely more trip related rather than user related, as they refer to the valuation of reliability, surface quality, congestion, queuing, platooning on single carriageways, of which some are covered in **Chapter 6**.

**Automated vehicles (AVs).** Looking further into the future, automated vehicles (AVs) may become important. These are likely to be of little immediate policy relevance but given the length of the appraisal period any systematic lowering of VTT due to increased quality of journey time in an AV will feed through into biasing the appraisal. There is a lot of uncertainty associated with AVs, and as a mode that does not yet exist, they would be very hard to elicit a VTT for, creating further uncertainty. It may therefore be best to address the impact of AVs by expanding uncertainty analysis on future VTTs. See also **Section 5.10**.

Further disaggregation of VTT values to meet policy needs during appraisal could, however, be onerous. We need statistically reliable values and if the differences between different segments are not significant then there is an argument that they should be combined. Thus, a highly disaggregated VTT dataset may require a very large sample. We also note the above set of possible topics is only indicative and by no means exhaustive. There may be other policy priorities, such as addressing transport inequalities (e.g. considering ethnicity, gender or disabilities) that may benefit from exploration through the lens of VTT research.

## 4.2 Segmentation in UK modelling practice

Most UK transport models fit into two broad categories when it comes to segmentation of journey purposes and user groups. We can refer to these as ‘minimum TAG’ and ‘bespoke’, with the former by far the more common. This excludes agent-based and activity-based models which are discussed in the next subsection.

Segmentation in ‘minimum TAG’ models is very much driven by the minimum requirements set out in TAG Units M2.1, M3.1 and M3.2. Such models are typically severely constrained by budget and programme during development and application. They are often developed to support the business case for a specific scheme (for example to secure DfT funding), rather than as long-term strategic planning tools.

Segmentation in these models is typically along the following lines:

- *Highway assignment models:* segmented by journey purpose (for cars) and vehicle type to give five user classes: car employer’s business (EB), car commute, car other, LGV, HGV. Different VTTs apply for each user class. Note that the VTT does NOT vary by distance (which would be difficult and time consuming to implement in commercial software<sup>11</sup>).

---

<sup>11</sup> Commercial assignment software typically requires a fixed VTT for each user class, rather than one that varies by distance. As a workaround it is, in principle, possible to split existing user classes by distance band (e.g. ‘car commute 0-5km’) and allocate a different VTT to each. We are not aware of this ever having been

- *Public Transport (PT) assignment models*: if the model has been developed for a highway scheme there will often not be a PT assignment model. If there is one, then it is quite common to have no segmentation at all. This is typically the case where PT fares are not explicitly represented in the assignment and therefore removing the needs to consider the VTT, perhaps because the assignment is unimodal (e.g. just bus), or PT submode choice (e.g. bus vs rail) is handled in the demand model rather than the assignment (in other words, the impact of fares, and the use of VTT, is only considered within the demand model (see below); the simplifying assumption is almost always made that fares do not affect route choice for a particular submode and are therefore ignored in the assignment). In a few cases, the PT assignment is multimodal and handles submode choice, in which case segmentation by trip purpose may be included, with VTT explicitly included in the assignment. VTTs may then be needed for each component of PT travel time (i.e. walk, wait, transfer, IVT, interchange penalty).
- *Demand model*:
  - ‘Standard’ journey purposes are split by home-based/non-home based (HB/NHB) for business and other non-work trips, but not commute.
  - Traveller types are split by car available/no car available (CA/NCA). For some applications (usually highway schemes) the NCA segment is ignored, i.e. only CA trips are included in the model.
  - Since VTTs are imported from TAG, VTTs for business trips will vary by mode (with rail>car>bus), but for other journey purposes they will not. This may not accurately reflect mode preferences for a given individual or segment because, for modelling VTTs, we only want to include user effects, whereas the TAG values include mode *and* user effects (see **Section 4.4**).
  - VTTs may vary by distance, but usually do not. Many demand models require the use of ‘cost damping’ to reduce the demand response for longer distance trips. Although varying VTT by distance is one of the cost damping methods suggested by TAG, in practice one of the simpler alternatives is more commonly used, often ‘damping generalised cost by a function of distance’. These use of this function could be interpreted as approximating the effect on travel behaviour of explicitly modelling distance-dependent VTT<sup>12</sup>.
  - VTTs are not usually used at all in PDFH-based rail demand models, which instead specify fare and Generalised Journey Time (GJT) as separate variables.

Calibration of ‘minimum TAG’ models is often based on importing as many parameters as possible from TAG (VTTs, vehicle operating costs, demand model sensitivity parameters etc.) and making adjustments to achieve TAG calibration standards. For example, demand model parameters may be taken from TAG Unit M2.1 and these are then adjusted to meet the realism testing standards (e.g. ensuring that the elasticity of vehicle kms with respect to fuel costs is within the expected range).

In contrast, ‘bespoke’ models are likely to have much more detailed segmentation that goes beyond the minimum requirements set out in TAG. Such models tend to have a broader strategic planning role, beyond simply supporting the business case for a specific scheme. Rather than relying on importing parameters from TAG they are more likely to involve a formal estimation based on local data (such as

---

done in practice. It would make the model more complicated and increase run times, plus there are concerns about the use of distance bands (see para 1.10 onwards of DEPARTMENT FOR TRANSPORT 2016. Understanding and valuing impacts of transport investment: Value of travel time savings consultation response.).

<sup>12</sup> Alternative cost damping methods are set out in section 3.3 of TAG Unit M2.1. One of these is an explicit dependence of VTT on distance, but this seems to be relatively rarely used compared to the other, simpler, methods.

household travel diaries). This may mean that VTTs themselves are estimated from local data instead of imported from TAG.

Examples of where additional segmentation appears (compared to ‘minimum TAG’ models) include:

- For models involving some kind of road user charging (including workplace parking levies) additional segmentation by income (low/medium/high) is usually included in demand and highway assignment models (as required by TAG).
- Much greater segmentation in the demand model by socio-economic variables, such as age, gender, employment status, household car ownership and household income. However, many of these segments are represented as dummy variables in the utility function and are not associated with segment-specific VTTs.
- More detailed journey purpose segmentation. For example, other non-work may be split further into shopping, personal business, leisure, visiting friends etc.
- VTTs may vary by mode for some or all journey purposes and not just for business trips as per ‘minimum TAG’ models.

Segmentation by (bus) concessionary travel pass ownership is surprisingly rare, even though it is likely to be an important determinant of travel behaviour. Some ‘bespoke’ models use a ‘linear plus log cost’ formulation, which implicitly introduces a dependency of VTT on distance.<sup>13</sup>

Some ‘bespoke’ models lack the data to locally estimate VTTs and are therefore constrained to the values and segmentation in TAG. Even where VTTs are locally estimated, it does not necessarily mean that there is a different VTT for every single demand segment. Segmentation of VTT is often more aggregate than the most detailed segmentation used in the model (which is typically seen within the demand model).

Any discussion of segmentation in modelling needs to consider the data used in model development. A typical ‘minimum TAG’ model will make extensive use of mobile phone data (MPD) in building base year trip matrices. MPD currently has no person type segmentation and limited journey purpose segmentation (home-based work/education, home-based other, non-home based), which is often seen as unreliable when benchmarked against other data sources. It therefore needs to be blended with other data sources (such as synthetic trip matrices, NTS, NTEM) to produce the minimum segmentation required by TAG. This tends to result in considerable uncertainty in the segmented data used in the base model. While new data processing techniques are likely to improve matters in the future, it is important that future versions of TAG do not impose a requirement for increased segmentation in transport models that is greater than can be supported by the most commonly used data sources. These observations are consistent with those made in **Section 3.1** regarding the applicability of MPD and other passive MPD for the estimation of the VTT using revealed preferences.

Overall, we do not believe there is a significant demand from modellers for increased journey purpose or person type segmentation in TAG VTTs, with most models using the minimum segmentation required by TAG. Any move to increased segmentation, based on evidence from the next UK VTT study, needs to be balanced against the limitations of the data used to build transport models. One solution might be for TAG to present VTTs for different levels of aggregation, with practitioners picking whatever level is appropriate for their application, i.e. there would be the ability to drill down to a greater level

---

<sup>13</sup> See section 3.3 of TAG Unit M2.1 for details. Strictly speaking, this leads to VTT being dependent on monetary cost (such as fuel, PT fare), but this is usually strongly correlated with distance.

of segmentation, below the current ‘headline’ figures, by journey purpose, traveller type, or other dimensions.

### 4.3 Agent-based modelling

In this section we consider both the modelling and appraisal demands of Agent Based Modelling (ABM) on VTTs by trip purpose and user group. ABMs are quite a diverse group of models (for a brief overview see for example Department for Transport (2024I)). A unifying feature is that they make use of a synthetic population. Every individual in the study area is represented as an agent. They have attributes allocated to them which are constrained to overall demographics. Such attributes include household type, age, income, qualifications, etc. If needed it could also include ethnicity and other indicators which are useful for distributional impact analysis.

The economic agents then make travel decisions including where to travel and how to trip chain based on generalised cost-based accessibility measures. It is conceptually feasible that an individual VTT could be assigned to every potential trip based on the attributes of the agent (age, income, household type, etc.) and the trip (purpose, distance, mode, etc.). This would be akin to the allocation of a VTT to every trip in the sample enumeration process. How feasible that is from a demand modelling perspective will very much depend on the software and modelling system. However, the behavioural VTT models from the 2014/15 study, and any behavioural VTT models from a future VTT study would be highly beneficial for this allocation.

Even if it is not possible to incorporate individual VTTs in the modelling, it would still be possible in the appraisal to allocate individual VTTs to each trip, each agent makes based on their and the trip’s attributes. For the reasons advanced earlier, this would improve appraisals, particularly for ‘non-average’ projects. Such an approach does of course move away from the standard VTT approach and has many similarities to a behavioural WTP approach. We discuss such approaches further in **Section 4.5** on income.

In terms of the implications of ABM models for a future VTT study this would require the behavioural models to be made available for use. Possibly this could be in some variant of the ‘tool’ used to for sample enumeration of the NTS data in the 2014/15 study. It goes without saying that a future VTT study would be expected to test for and control for the effect of personal and trip making characteristics on the VTT as part of the analytical process. It is the results of that process that would be most useful to ABM modellers should they wish to utilise it.

### 4.4 Mode segmentation: mode vs. user effects

This section focuses in more depth on the question of segmentation by mode. It reviews the fundamental issue of disentangling mode and user effects – including existing international evidence and challenges – and the related quantification of welfare effects associated with modal shifts in larger transport infrastructure projects (KAI3). We discuss practical considerations, potential implications, and possible viable methods for implementation in a national study context.

#### 4.4.1 Context and problem statement

The acceleration of modal shift to public and active transport is a priority area for DfT in the context of decarbonisation plans. Being able to capture ‘within-person’ differences in VTT across modes would enrich the appraisal of large modal shift interventions. This is at present limited by using a single VTT across modes for commute and other non-work trips or, perhaps more importantly, by very limited UK-based evidence on ‘within-person’ modal differences.



The initial question is: what makes VTT segmentation by mode particularly challenging, for instance more so than the counterpart segmentation by purpose? One answer would be that estimated VTT varies by mode for multiple reasons, which can be divided into ‘mode effects’ (ME) and ‘user effects’ (UE). Starting from the current practice of targeting respondents by their main mode of transport and purpose, the resulting mode-specific VTT by mode is going to differ across modes for two reasons: i) differences in the *perceived qualities of the modes*, such as comfort or ability to use time in certain ways; ii) differences in the *user groups* currently travelling by one mode or another, such as income differences or self-selection (see Flügel, 2014).

Going back to the 2014/15 VTT study, there was a clear desire to achieve mode segmentation for all trip purposes. Although the final estimates revealed significant variation in the VTT across modes for commute and other non-work trips, equity values were adopted representing a uniform VTT for these two trip purposes across modes. This decision was largely driven by concerns about confounding of ME and UE in the 2014/15 VTT study. Notably, other countries which have made use of largely similar study designs did segment the VTT by mode and trip purpose and made little to no reference to the separation of ME and UE.

From an appraisal perspective, the analyst is interested in disentangling ME from UE and observing the mode effects – i.e. the within-person (or within-group) differences in VTT across modes, exclusively due to the characteristics of the modes. This segmentation would seem to require information *from each group of users* on the VTT for different modes. In other words, car users should be able to reveal their VTT by car and by *other mode(s)* (e.g. rail or bus); rail users about their VTT by rail and other mode(s) (e.g. car or bus); etc. This additional information was missing in the last UK study, compromising the task of disentangling the ‘mode effects’ for each group.

It should be noted that the 2014/15 VTT study did include income controls, but unfortunately it was judged that the resulting VTTs by modes still reflected some remaining of income effects (and thus that income was not fully controlled for). This was a disappointing outcome, but understandable. In essence, there was an expectation that VTTs by mode could be estimated in a way that the ranking by mode was in line with considerations of mode quality/comfort (i.e. bus having the highest VTT). However, this was an unrealistic expectation, because data was only collected on the main mode used per respondent. Thus, if one group has lower income (e.g. bus users), regardless of how well the models control for income, it is likely that a lower income group will reveal a lower VTT. In other words, user and mode-specific effects could not be disentangled. Some key considerations around confounding mode effects with income effects will be addressed in this section, while **Section 4.5** provides a fuller discussion of the fundamental role of income in the segmentation debate.

#### 4.4.2 Review of evidence

Wardman (2004)’s meta-analysis of VTT studies was the first, to the best of our knowledge, to raise awareness of the important distinction between ‘user effects’ and ‘mode effects’ in the VTT. Wardman’s review concluded that VTT studies generally failed to disentangle the two. The reason was that, in national VTT studies, car users tend to be asked only about car VTT, bus users about bus VTT, and so on. The only study which had investigated this at the time was the first Dutch national VTT study (Gunn and Rohr, 1996). Wardman (2004) concluded that self-selection was arguably a main driver of mode differences, which is part of the ‘user effects’; and, in general, argued that there was evidence of a stronger role of user effects over mode effects.

Batley et al. (2019)’s own meta-analysis however managed to provide estimates of mode effects for the sample of car users. The samples of bus and rail users were too small to allow estimation. The findings suggest the following relationship: Rail VTT < Car VTT < Bus VTT, which implies rail offers the

highest quality of time on average (*ceteris paribus*), followed by car. The relationship holds for various distance bands, varying in magnitude.

Mackie et al. (2003), for the former UK national VTT study, relied on Wardman's meta-analysis to cover the issue of mode effects. However, the estimates were not subsequently used in appraisal guidance. The statistical significance (only significant at the 90% level) of the bus vs. car VTT findings was seen as problematic, and the authors recommended corroborating evidence to support modal VTT in appraisal. They also argued that the variation with distance could be due to the greater disutility associated with distance for a car *driver* compared to a passenger in other modes (e.g. due to fatigue).

The Swedish and Norwegian studies in the mid-90s seemingly collected evidence that would have enabled the isolation of mode effects, but analyses on this issue were not reported (Wardman, 2004).

Fosgerau et al. (2010) attempted to disentangle these in practice in the Danish context. The study tested the hypothesis of whether the differences between modes were caused by self-selection or strategic behaviour. They concluded that the evidence pointed to self-selection. In contrast with studies in the Great Britain context, the author's theoretical expectations were different, expecting car to be the more pleasant mode (thereby associated with a lower VTT). The results confirmed this expectation for the samples of car and rail users but found no statistically significant differences for the bus sample. Likewise, bus VTT was not found to be significantly different from car and rail VTT in any of the samples.

Flügel (2014) was, to our knowledge, the first paper providing an in-depth analysis of the potential use of mode effects estimates in appraisal. The paper provides estimates of the mode effects (and user type effects) for the Norwegian context – although focusing only on long-distance (>100km) trips. The results suggest a pattern of VTT across modes where rail and bus have values not significantly different from each other, and both have lower VTT than car (suggesting higher quality of time in public transport). Air travel is found to have the highest VTT by a wide margin.

The paper also presents and illustrates several approaches to estimate and exploit variations in VTT across modes in appraisal, comparing them to current approaches in practice (see **Figure 11** for a diagrammatic illustration of how the approaches may fare, relatively, in terms of providing 'precise' behavioural WTPs versus providing 'equity' values).

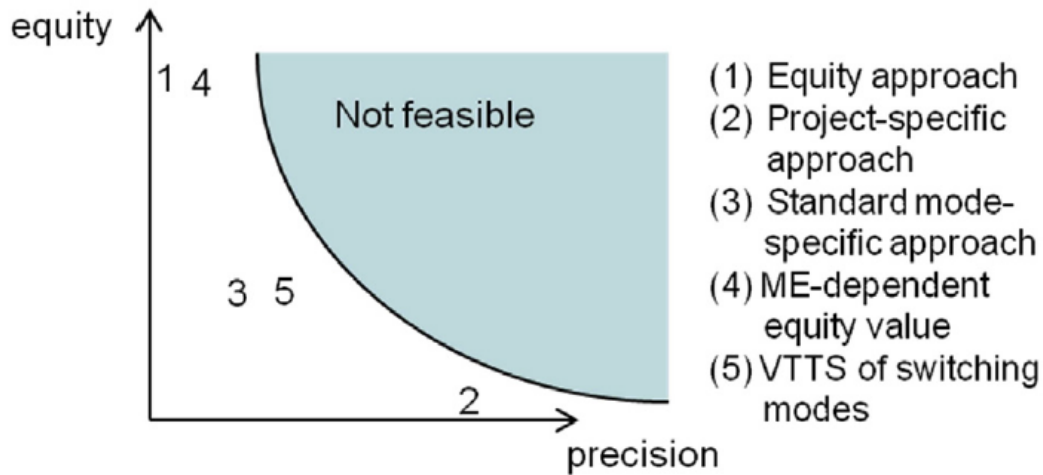
Flügel (2014) concluded that the *ME-dependent equity VTT* is preferred over a single equity VTT, except for schemes leading to mode shift, where the *VTT for switching modes* is preferable. The key, in both cases, is to be able to separate mode effects (ME) from user effects (UE), and in turn the key to that goal is to collect data on people's VTT of *at least* two different modes. In that way, it is possible to compute the ME as the VTT difference between the two modes for a given group of users (likewise, the UE would be the difference between two user groups for a given mode).

The Mode-Effect-dependent equity<sup>14</sup> VTT is seen as an improvement in the equity dimension over the standard mode-specific approach – as applied in, for example, e.g. Sweden, Netherlands, or Norway. This ME-dependent approach calculates a "weighted average of the VTT in a given transport mode over all user groups in the population" (Flügel, 2014 pp.55), thereby removing user effects (UE) from the standard mode-specific VTT values, so that the resultant VTTs across modes differ exclusively by the intrinsic mode characteristics (the ME). In other words, the ME-dependent equity values are effectively "representative for the whole population in a given transport mode" (*ibid*). When compared

---

<sup>14</sup> Technically, this refers to the approach where one single VTT is applied across modes (but there are different views around whether it is an 'equity' value or not).

with approach (3) below, the ME-dependent equity values provide an improvement on the equity dimension because they are controlling for self-selection (since they include all possible groups in the calculation, instead of only the users of a given mode). Note that this has substantive implications for adopted sampling strategies.



**Fig. 2.** Illustration of VTTs-approaches (including the two proposed approaches) on the equity and precision dimension.

*Figure 11: Different VTT approaches as presented in Flügel (2014)*

The proposed ‘VTT of switching modes’ approach is more complex. This is because it requires the definition (and thus observation) of sub-groups *within* the user groups. User groups (by current mode used) would need to be subdivided into sub-groups based on a *previous mode choice* (i.e. switchers from mode  $j$  to mode  $k$ ). As noted by Flügel (2014), in the context of national VTT studies “it seems difficult (practically impossible) to keep track of the composition of different user groups” (p.55). Flügel argues that the VTT for switching modes “is in principle preferable to (standard) mode specific VTT in situations where projects lead to changes in mode choice. This is because it considers the fact that there are user type effects between “switchers” and “non-switchers”. In principle, it therefore provides a more detailed representation of actual WTP. However, when project specific valuation studies are not feasible, one must rely on proxies for the actual (project specific) group of switchers. As this may involve a notable inaccuracy, it is not certain that one would gain precision in every project appraisal” (p.59).

The implementation of these alternative approaches rests on the ability to successfully disentangle ME and UE across the relevant user groups (and, in the case of the VTT for switching modes, across sub-groups). However, designing the surveys to gather this evidence is not without difficulties, as became apparent in the recent Norwegian study (Flügel et al., 2020). They followed a similar approach to the earlier study reported in Flügel (2014), but encountered difficulties unpacking ME and UE successfully for some segments. This is probably due to additional challenges arising in SP settings when targeting VTT in modes used less often by the individual, and when the goal is to compare these VTT with the VTT of the user’s main mode (e.g. extra care is needed in SP designs given variations with distance or reference costs and time). Indeed, although the Norwegian study did contain some estimates that remove the UE in the computation of mode-specific VTTs (which could have enabled

the application of ME-dependent equity values), these were not adopted in appraisal guidance – which still makes use of approach (3) as per the figure above (i.e. standard mode-specific VTT).

The topic was again one of many issues addressed in a more recent VTT meta-analysis by Wardman et al. (2016), covering nearly 400 European studies between 1963 and 2011. Table 7 in Wardman et al. (2016) summarises the average ratios of VTT between pairs of modes as observed within studies, separating the estimates by segments (e.g. car users, rail users, etc.) to limit confounding. The evidence from the studies themselves allowed the authors to make the following generalisation: *“train is regarded to be superior quality given its lower VTT even amongst car users, followed by Metro then car with bus regarded as the most inferior quality”* (p.11), which they found to be largely in line with expectations. Air travel was again found to have the highest VTT, in comparison to rail. The meta-analysis model, however, was not able to derive statistically significant differences across modes other than between air and rail. The authors were surprised with this outcome given the clear pattern of variation observed within studies, but it is possible that the model simply could not disentangle the mode effects from the many other variations being estimated simultaneously. They also pointed out to possible differences between short and long-distance trips, which could not be untangled.

Batley et al. (2019), in the latest Great Britain national VTT study, expected ‘comfort effects’ to be reflected through VTTs that are lowest for rail and highest for bus (leaving car and other PT somewhat in the middle). In models that control for income effects, they found this relationship only for business trips (noting no bus trips were sampled in business by design). However, for commuting trips, bus had the lowest VTT (followed by rail); and for other non-work trips, car had the lowest VTT. They concluded that the observed modal differences could not be explained solely by comfort, and thus their recommended values were averaged out across modes (i.e. no mode segmentation for commute and other purposes).

Börjesson and Eliasson (2019)’s notable contribution to the question of VTT segmentation – previously mentioned – also includes empirical evidence from the Swedish VTT study. They present VTTs both controlled and not controlled for income, for the different segments (1: Car, Stockholm; 2: Car, other regions; 3: Car, all regions; 4: Bus; 5: Train) split by distance groups (a: short distance, commute; b: short distance, other purpose; c: long distance, all purposes), alongside the after-tax monthly income (in k,€) for each segment group. Their results show the following relationship for each distance-purpose group: Bus VTT < Rail VTT < Car VTT, regardless of whether income is controlled for or not. This suggests that the marginal disutility of time (i.e. the level of comfort / quality) for bus is consistently higher than for rail and that it is lowest for car journeys across income groups.

The recommendations from the Swedish and UK studies are, however, different and highlight that there is an element of judgement embedded in the debate on VTT differentiation. In the UK study, the patterns were interpreted as a failure to fully control for income. This was not the case in the Swedish study, where the conclusion was that *“it is therefore clear that the main driver of VTT variation in these dimensions is not income differences”* (Börjesson and Eliasson, 2019 pp.370). This is because controlling for income does not have a huge impact on the VTTs for a given group (which they attribute to the small income differences in Sweden as measured by the Gini coefficient). However, we argue that this is debatable by looking at the bigger picture of values across groups. Below we plot the VTTs from all groups (15 groups) alongside their income (Table 1 in Börjesson and Eliasson, 2019). These tables show that there is a clear relationship between VTT and income that remains, both with behavioural VTTs and with the income-controlled VTTs. We believe this is probably a similar situation to the UK study, where similar modal differences were interpreted as a sign that income still played a role (across groups), e.g. due to self-selection. While the Swedish felt comfortable recommending their obtained mode variation in VTT for use in appraisal, the British did not.

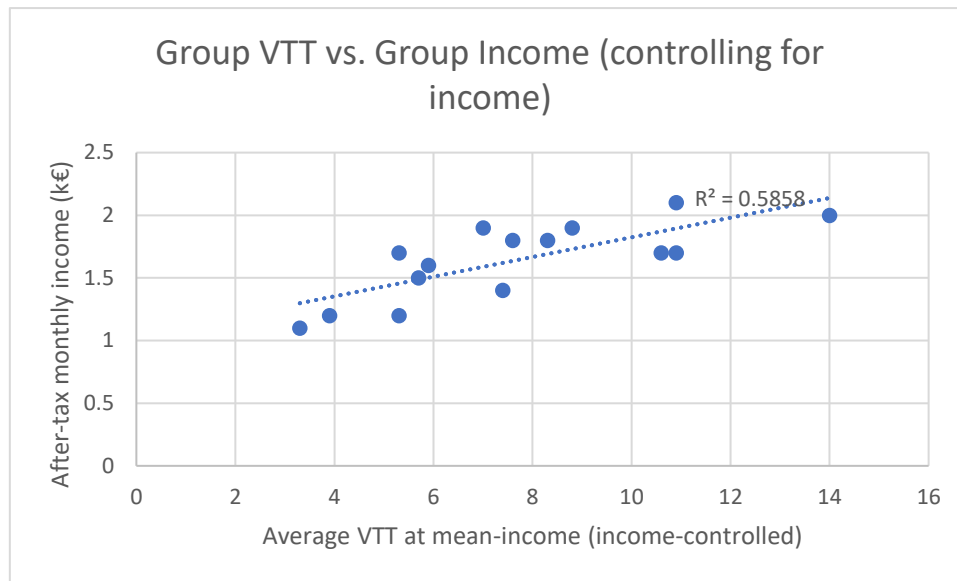


Figure 12. Source: own work based on values from Table 1 in Börjesson and Eliasson (2019, p.370)

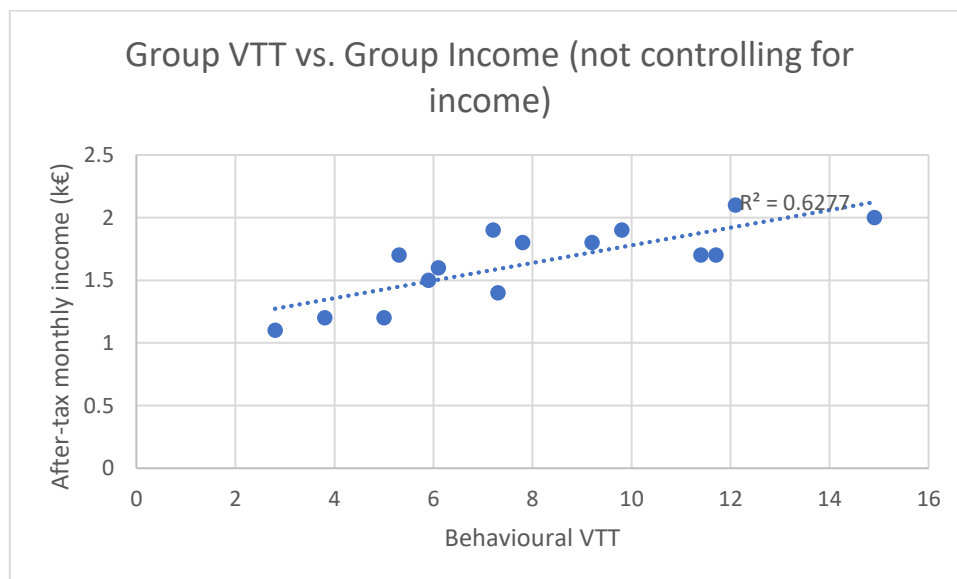


Figure 13 Source: own work based on values from Table 1 in Börjesson and Eliasson (2019, p.370)

A more recent series of studies deployed new techniques to unpack mode effects. These studies are based on an Austrian innovative survey using a Mobility-Activity-Expenditure Diary (MAED), first reported in Aschauer et al. (2019). This novel procedure was designed to overcome the limitations of simpler datasets (e.g. just SP-based) for the estimation of the theoretically well-established subcomponents of the VTT, namely where  $VTT = VoL - VTAT$  (where  $VoL$  is the value of leisure and  $VTAT$  is the value of time assigned to travel). It is the  $VTAT$  what precisely differs between modes and travel conditions, and thus drives mode effects. Of course, the  $VTAT$  is highly subjective too and thus can vary among people/user groups – thus,  $VTAT$  can have both ME and UE components, whereas  $VoL$  is in principle free of ME as it only refers to the opportunity costs of time (thus, the separation of  $VoL$  and  $VTAT$  is not the same thing as the ME vs. UE identification, although it certainly can shed light on it). The MAED procedure is, however, data hungry. In this first application, it collects a diary and questionnaire from 748 Austrian workers covering 1 week. This includes information on 10 activity

types, 14 expenditure groups and 4 transport modes (walk, bike, car, PT). The final sample covered information from over 17k trips, with an average of 23 trips per person in the reporting week.

Using this dataset, Schmid et al. (2019) provide VTT estimates of mode and user type effects in the Austrian context. Effectively, they used a discrete choice model that combined RP/SP data and different experiment types (mode, route and shopping destination choice) – see also Jokubauskaitė et al. (2019). 74% of the 748 MAED respondents completed the follow-up SP survey. Their results, which included active modes (a novelty relative to past studies that addressed mode effects), revealed the following pattern of VTT across modes: Car > Bike > Walk > PT. Noting that PT encompasses bus, tram, light rail and heavy rail, their results show PT VTTs are markedly lower than Car VTT (8.1 vs 12.3 €/h). The large difference is attributable to mode effects (e.g. PT being perceived as providing a better travel quality experience such as higher comfort or increased opportunity to use time effectively, for a given length of trip) and persists regardless of what mode and user effects the model controls for. The authors note the contrast to previous European studies (where user effects dominated). The within PT variation was explored in less depth (only available through the RP data but not SP data), but indicated bus had the highest VTT of the subset (9.6 €/h vs. 8.1, 7.9 and 8.5 €/h respectively). It must be noted that for the RP component, the information on alternative modes was inputted from other sources to be combined with respondent's self-reported information on chosen modes. This input process required for alternative modes, typical for RP procedures, can play a role in the VTT estimates. Without an in-depth understanding of the data, it would seem adequate to be cautious when comparing the modal differences results from these newer studies with those relying solely on SP such as Flügel (2014).

Hössinger et al. (2020) provided for the first time mode-specific VTAT as part of their wider endeavour of estimating the full modelling framework based on Jara-Díaz et al. (2008), covering simultaneously time-use and expenditure allocation choices. However, a notable limitation of this study is that they do not estimate their own discrete choice model, but instead used Schmid et al. (2019)'s results as inputs – as flagged by Jokubauskaitė et al. (2019), this separate estimation is problematic as it fails to account for possible sources of correlation among different decisions and error terms. This is precisely one of the gaps that Jokubauskaitė et al. (2019) try to fill by estimating the full framework through a joint estimation process. Accordingly, they also provide estimates of VoL and mode-specific VTAT and VTTs. They claim their results are comparable to those from Schmid et al. (2019), although their headline average results seem to show some differences. For instance, the VTT ordering across modes is slightly different as follows: Walk > Car > Bike > PT. With Car and Bike VTT very close at 9.9 €/h and 9.25€/h, the PT VTT is much lower (€3.9€/h); whereas Walk VTT is also much higher (17.5 €/h). They also find that the VTAT is only positive for PT (at 5.38€/h), arguably “*reflecting the favourable PT conditions in Austria*” (p.397).

A similar pattern was found for the case of Zurich (Switzerland) by Schmid et al. (2021), with PT VTT being half (48%) of Car VTT. While the focus was not placed on the isolation of user and mode effects, the authors argued that these observed differences are largely attributable to mode effects due to most respondents being familiar with both modes, the correlation patterns in the data and the advanced model specification used in estimation. Like the Austrian results, they also find a positive VTAT for PT and for bike.

In the latest study on this topic, Hartwig et al. (2024) take stock of the evidence to date, highlighting that there is a systematic pattern of VTT being lower for PT users than for car drivers. Their headline VTTs are 4.96€/h for PT users and 15.1 €/h for car users, respectively. Hartwig et al. (2024) use a more recent diary-based mobility-activity survey dataset collected between 2019 and 2020, also from Austria. Their focus is on decomposing mode-specific VTT with respect to different levels of travel-based multitasking, using RP data. Part of their explanation the lower VTT for PT can be found in the

good opportunities for multitasking. However, they note they do not thoroughly control for user-type vs. mode-type effects, which is an important limitation. Relative to Schmid et al. (2019) (with PT/car ratios of 0.66 and 0.48) and Jokubauskaitė et al. (2019) (PT/car ratio of 0.39), the obtained ratio of 0.33 is somewhat on the low end.

It should be noted that some of these MAED studies also report results in relation to VTT by trip stage (e.g. access, headway, etc.), another area where the MAED survey approach could provide valuable evidence.

#### 4.4.3 Taking stock, challenges, and opportunities

In the context of a vast body of literature on the VTT, the issue of isolating mode effects from user effects has received relatively little attention. The key take-aways from our review are as follows:

- The evidence is scarce and limited to a few studies in Europe, but the topic seems to be gathering momentum in recent years.
- Evidence, but also a priori expectations, can vary by country. This suggests that the comfort characteristics of each mode can vary by country to the extent that it may affect the expected ranking of mode VTTs in the eyes of the analyst.
- The key to separate mode effects (ME) from user effects (UE) is to collect data on people's VTT of at least two different modes. This can be done in the context of multiple within-mode route choice experiments varying the mode of transport, or in the form of mode choice experiments. In that way, it is possible to compute the ME as the VTT difference between the two modes for a given group of users (see Flügel, 2014).
- The only UK-based study (Wardman, 2004) pointed towards the following ranking, for a given group of car users: Rail VTT < Car VTT < Bus VTT.
- In most cases, we find that Car VTT is at the high-end of the modal VTT 'ranking', suggesting perception of lower quality of time spent in the car compared to PT. This is especially the case in recent studies published since 2019 – however, the countries where these studies were conducted are also notorious for having good public transport services, which would influence the relative perceptions. The only exception throughout the review, both in terms of expectations and results is the Danish study (Fosgerau et al., 2010), where Car VTT < Rail VTT.
- Distance has been observed to play a role in the modal differences. For example, Wardman (2004) found that rail to be even more favoured over other modes for longer distance trips, while Flügel (2014)'s findings only apply to the long-distance segment (over 100km).
- Since 2019, a new data approach has been deployed to address this issue and unpack VTT more than in previous studies, combining SP with diary-based surveys (see series of Austrian and Swiss studies using the 'Mobility-Activity-Expenditure Diary').
- These more recent studies also cover active modes (walk and cycling VTT) and new forms such as carsharing or carpool. Walk and cycle VTT have also been found to be lower than Car VTT in some cases, but not consistently.
- These studies also show a larger gap between PT and car VTT, with PT VTT being between much lower than the Car VTT. However, due to methodological differences and contextual differences (they cover Austria and Switzerland), it is not possible to ascertain whether this is a trend over time.

Going forward, we identify some opportunities but also some challenges:

- It would be beneficial to gather the necessary data to explore the prevalent mode effects in UK contexts, given the very limited UK evidence available. This should cover at least the most

common modes (car, rail, bus) but ideally also other forms of public transport (metro, tram) and active modes (walk, cycling).

- The key missing piece in previous studies was data on the VTT for various modes for each person surveyed (rather than only for their chosen mode in the intercepted journey) - that is, multiple modal VTT per user group. This gap can be easily addressed in a future study.
- It would seem practical to use SP methods for this purpose – aligning with the wider VTT study.
- Further work would be necessary to explore the avenues for implementation in practice of sets of VTT that identify ME and UE, building upon Flügel (2014) and the Norwegian appraisal experience.
- Consideration should be given to Flügel's recommended ME-dependent equity approach, which is – in theory – an improvement upon both most common forms of current practice (specifically equity VTT – which do not allow appraisals to account for mode effects – and standard mode-specific VTT – which do not remove user effects). This would require collecting VTT data on at least two modes for every respondent.
- For appraisal, it is relevant to think through the different contexts in which ME matter most, e.g. mode-switching contexts. This could inform the development of suitable data collection and approaches (e.g. determining the optimal combination of modes and user groups for a VTT study). One example of a relevant context is the appraisal of new modes, where the DfT has recently published a new unit and recommends – for that context – the use of VTT multipliers that pick-up mode effects based on Mackie et al. (2003)<sup>15</sup>.
- If a study is successful in unpacking ME in a usable way for appraisal segmentation (e.g. via Flügel's ME-dependent approach), this can open the door for broader segmentation than in current practice, avoiding some of the issues from the past that were linked to concerns about user effects and, in turn, improving the approaches also in relation to the other variables such as distance.
- A word of caution comes from Flügel et al. (2020), where results suggest that unexpected results might occur, and thus extra care must be taken in the design and surveying process.
- Potentially, it might also be interesting to undertake more exploratory research that follows the MAED approach used in Austria and Switzerland. This is however data-intensive and, from the point of view of a national study, it is not clear what the additional benefits would be relative to the simpler SP-based Norwegian approach as per Flügel (2014).

#### 4.5 Treatment of income

Income is one of the more prominent user effects associated with the VTT. The relationship is most clearly illustrated by revisiting the simplest definitions of the VTT, i.e. where the VTT is defined as the ratio of the marginal utility of time over the marginal utility of income. In economic theory, it is commonly assumed that the marginal utility of income is decreasing in income due to decreasing marginal utility of consumption. As a result, the VTT is expected to increase in income and there is ample evidence available supporting this assumption. Wardman et al. (2016) estimate an income elasticity of around 1.0 in their meta-analysis of VTT values across Europe, and this is also the elasticity used by the Department to uprate the VTT with income growth over time (ARUP and ITS, 2017).

##### 4.5.1 Estimation of income effects and elasticities in behavioural VTT models

Mackie et al. (2003) and Hess et al. (2017a) directly estimate income elasticities in their choice (i.e. behavioural VTT) models in the two most recent UK national VTT studies. In both instances the

---

<sup>15</sup> <https://www.gov.uk/government/publications/tag-appraisal-of-new-modes>



emerging income elasticities are significantly smaller than 1. This is also the case for the most recent Dutch VTT study (Significance, 2023) and other national studies estimating direct income elasticities. A meta-analysis on the cross-sectional value of time by Binsuwadan et al. (2023) further confirms this picture, and Rich and Vandet (2019) find that there are differences in how the inter-temporal and intra-temporal VTT relate to income. Their estimates suggest that cross-section income elasticities are smaller than inter-temporal income elasticities. Re-analysis of the 2014/15 UK study by Tjiong et al. (2022), however, does reveal that by redefining the income variable in estimation – specifically by accounting for the progressive nature of income tax and social benefits – produces cross-sectional income elasticities closer to 1.

Before discussing the measurement of income in more detail, it should be noted that Batley and Dekker (2019) implies that many empirical VTT specifications accounting for the role of income are inconsistent with economic theory (see also **Box 1** in **Section 3.4**). First, most empirical specifications separate the role of prices and income in their functional form. Batley and Dekker (2019), however, highlight that the theoretically correct specification should make use of ‘residual income’, where any remaining budget (or change in budget due to a price change) affects the consumption of the outside good (or numeraire). In the simple discrete choice settings underlying most VTT studies, respondents have no opportunity to express how any money saved will be spend on other goods and thereby generate utility. Second, when relying on pure discrete choices respondents cannot express non-linearities in their marginal utility of the outside good, i.e. the marginal utility of income. Consequently, for empirically estimated models of VTT based on discrete choice to be consistent with economic theory they can only make use of a constant marginal (dis)utility of income (and price). In effect, this results in the well-known additive-income random utility maximisation model (e.g. Fosgerau et al., 2013), which describes the behaviour of a representative consumer, where income has no material impact on the model and the VTT does not change with income. In other words, income effects can in theory not occur. Such specification would significantly simplify estimation and aggregation of welfare effects. The only way to overcome this shortcoming of discrete choice data is to make use of alternative data sources, such as for example the MAED discussed in the previous section can consume discrete-continuous amounts of trips (Small and Rosen, 1981).

The empirical application of discrete choice models, including national VTT studies, is in sharp contrast with the above theoretical recommendations and separate estimation of income and price effects are prominent. Indeed, empirical identification of such effects is entirely possible and increasingly flexible software to estimate discrete choice models enables researchers to do so. Where price information is readily available and generated by the experimental design in the VTT stated choice surveys, or inferred from external sources in revealed preference, the measurement of income is significantly more challenging in surveys or even lacking in RP studies.

The first challenge is to determine the relevant income metric, or budget, used by households to determine their expenditure on travel. There is large variation in the collection of personal and household income, gross vs net, weekly, monthly, or annual income. As noted by Tjiong et al. (2022), little account is taken of any additional social benefits individuals and households may receive. Particularly, since these social benefits primarily are targeting lower income households the true distribution of income in the sample may be much smaller than in direct measures of net (or even gross) income.

The second challenge is that it is nearly impossible to elicit accurate estimates of the relevant income metric. In general, respondents are given the opportunity to indicate to which income category they belong, but these are generally widely specified. Moreover, to avoid high levels of drop-out

respondents are typically not forced to report their income, resulting in large shares of 'don't know' or 'prefer not to say' responses.

Together, these two challenges mean that in traditional VTT-based stated choice surveys income effects are likely to be associated with large degrees of uncertainty and most likely associated with a bias due to their imprecise measurement, particularly towards the upper end of the income distribution where the income categories tend to become increasingly wide.

As such, direct estimation of income elasticities will pick up prevalent patterns of the relationship between VTT and income, but less emphasis (or faith) should be placed on the exact elasticities estimated. This is an issue which is rather hard to solve because of the reluctance of respondents to provide precise income information, but some gain can perhaps be made with more consistently specifying income categories and identification of relevant characteristics determining social benefits. The latter can be done based on background questions such that metrics like equivalised household income can be derived ex post. Alternatively, modelling attempts could move away from estimating income elasticities and instead use dummy variables as relevant interactions between the VTT and income.

The issues related to identifying the relationship between income and the VTT are not independent from the sampling concerns expressed by Halse et al. (2023), where respondents with higher VTTs (i.e. often those with higher incomes) are harder to survey. Börjesson et al. (2023) return to this point by studying repeated national VTT studies in the Netherlands and Sweden to study intertemporal changes in the VTT. They experience significant difficulties in explaining the variation in the VTT over time. They either conclude that the VTT has declined for given income levels for many modes and purposes, or that the issue can be attributed to sampling procedures getting increasingly selective with respect to VTT because busier people drop out first thereby lowering the VTT.

Overall, this highlights that there are significant challenges in accurately capturing the relationship between income and the VTT. This may particularly be associated with the use large national VTT studies spread over wide time intervals (typically more than 10 years in between) and associated with changes in the study design making it increasingly difficult to attribute variations in the VTT to specific sources. The suggestion put forward in ARUP and ITS (2018) to re-sample at regular time intervals using the same study design may offer new insights.

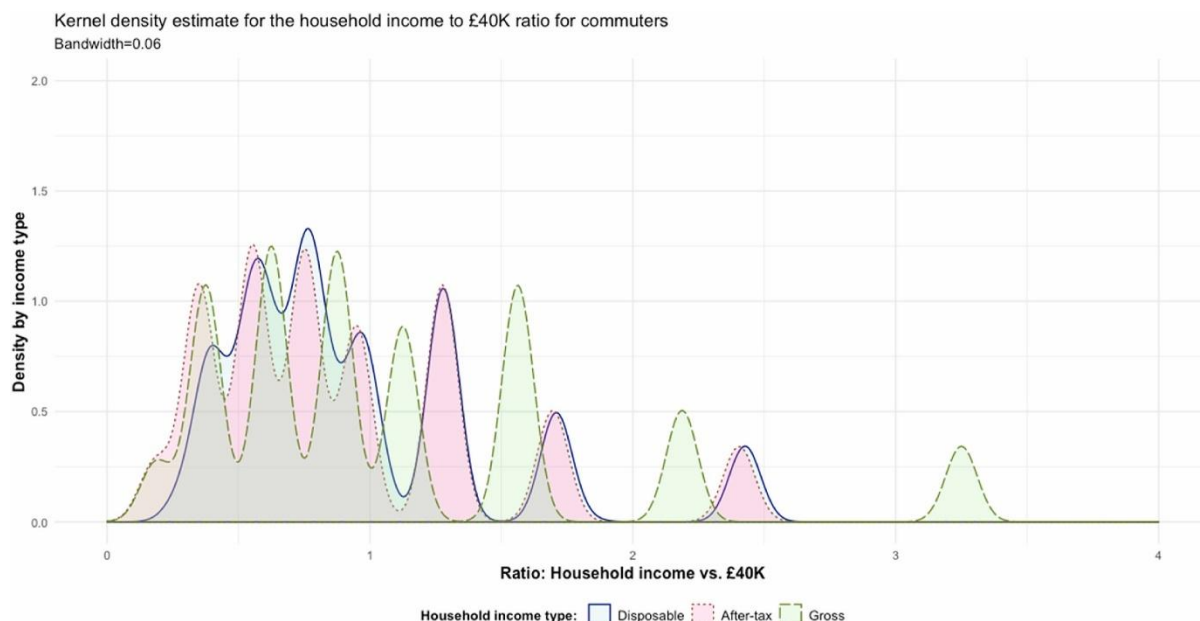


Figure 14: Impact of alternative assumptions on income on the income distribution in the 2014/15 UK VTT study. Source: Tjong et al. (2022).

#### 4.5.2 The role of income in appraisal practices

The treatment of income in the VTTs used for appraisal has not substantively changed since the first national value of travel times savings study in 1987 when stated preference values were first adopted to the appraisal VTT. This is despite recommendations in the 2003 study (Mackie et al., 2003) and the 2015 study (ARUP et al., 2015a) that VTT should differentiate by income particularly for large projects or for projects that involve pricing. Effectively, the recommendations in 2015 did not differ from those in 2003.

There is a slight difference for modelling purposes. In the main 'standard' VTT values are used for modelling, but where a road pricing policy is analysed or within the model (e.g. London Congestion Charge, low emission zones, etc.) then the demand segments in the model are disaggregated by income with differing VTTs by income segment. VTT values for three income groups are provided in TAG in Tables M2.1 for non-work and M2.2 for employers' business. However, even for projects where the modelling segments by income, the appraisal uses the standard values (i.e. non-income segmented VTT values).

Conceptually, the income of travellers influences their WTP for VTT, with higher income travellers typically willing to pay more than lower income travellers for the same time saving. Given that traveller income varies systematically by region and by mode, then varying VTT with income in appraisal may be expected to lead to project investment being directed towards high income modes (e.g. rail) and regions (e.g. London), with projects on low-income modes (e.g. bus) and in low-income regions (e.g. North of England) losing out unless the number of users affected is significantly larger. Such distributional matters can be addressed by using distributive weights, whereby benefits to low-income households are weighted more highly than benefits to high income households. The Green Book provides weights to make such calculations, but their translation into practice is not straightforward and ongoing research efforts are exploring these issues. From that perspective, it would be useful to future-proof a new VTT study to ensure compatibility of emerging values with the adopted approach for distributional weights – for instance, consideration should be given to obtaining robust and compatible data on passengers' incomes and consumption.

Distributional weights should ideally not be implemented at the level of specific benefits, such as travel time savings, but at the project level where the relative distributional impacts of projects across income groups can be quantified and adjusted for using distributional weights (i.e. for any impacts). Framing this in the context of the social welfare function, one would aim to identify the impact of the project on an individual's utility and recognise those benefits may vary across individuals and specifically income groups, and it is in the aggregation of those changes in utility that distributional weights come into play a role.

The risk of not differentiating VTT by income is that the wrong project may get developed. Consider two mutually exclusive projects: a grant subsidy for the infrastructure for a tram service or a grant subsidy of bus priority infrastructure. Both would serve the same low-income housing estate. The tram tickets cost substantially more than the bus tickets. Low-income households in the housing estate have a high value for money and prefer financial savings over time savings which is expressed through a low VTT. The use of a standard VTT would therefore overestimate the value of the tram service vis a vis the bus service for these households. Whether the tram or the bus infrastructure is preferred would depend on their relative costs, the differences in time changes between the two projects and very importantly the impact on the household money budget of the two projects. It would only be by using the actual VTT values that the latter point could be picked up in a CBA of the two projects.

On the face of it therefore we would ideally undertake appraisals that use pure WTP for VTT changes combined with a distributional analysis. A 'Green Book' style distributional analysis of a transport project would be very demanding as the transport user benefits are invariably passed through to landowners as changes in rents, or to households via changes in prices for goods and services. This pass through and who the real income gains accrue to would need to be modelled. To our knowledge such an analysis has never been attempted, though the tools for modelling the pass through, land-use transport interaction models and spatial general equilibrium models are more available. In addition to modelling this pass-through to final benefits, Mackie et al. (2003) identified the following additional challenges: obtaining the relevant data on the pattern of usage by income and social group; treating the non-monetised elements in the appraisal consistently with the monetised elements; and agreeing a set of weights. The latter point has been addressed by the Green Book, but the other challenges remain.

The distributional analyses undertaken in appraisal, and certainly UK appraisals, focus on the first round incidence of benefits and how they impact on certain demographics, in either positive or negative ways (Department for Transport, 2024g). This distributional analysis is presented alongside the 'Value for Money' statement (Department for Transport, 2024m) and the 'Spending Objective Analysis' (Department for Transport, 2023d). As such the distributional analysis does not influence the value for money statement itself but can influence the economic dimension through its alignment with the strategic objective analysis. That is, if a spending objective is the provision of improved accessibility to low-income households, then a project that achieves this objective is highlighted to the decision-maker as the most relevant economic analysis, where relevant relates to the alignment with the strategic objectives. In other words, the economic analysis is not adjusted as in the context of using distributional weights, but a project achieving such objectives is given more weight in the project shortlisting stage.

In the absence of a full weighted distributional analysis we return to the question as to whether and how VTT should be disaggregated into standard values. ARUP et al. (2015a pp 206) set out four options:

- Option (1): Include the observed variations in income by person and trip in the calculation of average VTT;
- Option (2): When calculating average values across the NTS data, treat all respondents as being at 'average' income.
- Option (3): Remove the income covariate from the behavioural model, letting the effect be picked up by other covariates.
- Option (4): Include variations with income as in option (1), but then re-weight according to the process set out in the Green Book.

The current TAG approach for non-work VTT is Option (1). Option (1) was recommended by ARUP et al. (2015a) if VTT was only to be disaggregated by trip purpose, as it currently is. This choice corresponds to Börjesson and Eliasson (2019)'s preferred approach. If also disaggregating VTT by distance and region (e.g. London) then in the absence of distributive weights, and given the systematic variation in income within these segmentations, ARUP et al. (2015a) favoured using an average income – i.e. Option (2). ARUP et al. (2015a) soundly rejected Option (3) as a mis-specification error that would bias other important behavioural parameters in the econometric model. Option (4) weighting the VTT by Green Book weights was felt to be impractical at the time but is now given greater data availability and modelling capabilities being actively explored by the DfT through a series of research studies.

With respect to business values the same options exist for the treatment of income. There is however a significant difference between business values and non-work values. The non-work values accrue to a household, whilst business values are social values that in part reflect increased productivity that will have broad social benefits. The differences in employers' business values by mode therefore reflect real societal differences on the value of additional output from the productivity increase. On this argument Option (1) is most relevant.

Overall, there exist theoretical and empirical arguments supporting disaggregating VTT by income, but these do need tempering by the practical realities of undertaking an appraisal. The income of travellers is rarely accurately observed and is difficult to impute, as highlighted at the start of this section and discussed by Tjong et al. (2022). These realities mean there will always be a role for some standardisation of income in appraisals – even in the Level 1 appraisal of Mackie et al. (2003) and ARUP et al. (2015a). Larger projects or those that involve pricing on either public transport or on the road network may justify the additional effort required to disaggregate VTT by income in the appraisal, particularly for projects where the modelling already segments by income (e.g. models which include some form of road charging scheme). Though it is worth emphasising that the evidence shows that to obtain a good estimate of WTP in the appraisal it is equally important to disaggregate the VTT by co-variables that explain the differences in marginal utilities of time alongside the disaggregation by income. This conclusion is slightly distinct from Börjesson and Eliasson (2019) who argue that these co-variables are more important than controlling for income. The reason for this is that the ability to separate mode and user effects in the VTT was not discussed in detail in Börjesson and Eliasson (2019), whereas our analysis clearly shows this issue has not been resolved (yet) in the wider VTT literature.

The Department's response to the 2015 research and recommendations identified a future research stream that would attempt to pool the results from the different modal SP surveys and attempt to separate out income and distance effects for non-work trips Department for Transport (2015a) (pp 51). To our knowledge this research has not taken place but aligns with the necessity to improve our ability to separate mode, user, and trip effects on the VTT.

#### 4.6 Conclusions on segmentation

From an appraisal perspective, VTT segmentation was a challenging area in the last UK VTT study in 2014/15, mainly due to the complex interrelations among key dimensions (e.g. income, mode and distance). Indeed, some questions remained unresolved, and some pragmatic decisions were made based on judgement. This was not unexpected, as it has for decades been a topic of debate in the academic literature. It does, however, appear that these discussions are more prominent in the context of UK appraisal practices than in other Western-European countries, such as Norway and the Netherlands, where for instance segmentation of mode and user effects were not considered of sufficient concern despite large similarities in study design and analysis.

The initial issue is the distinction between ‘behavioural VTT’ and ‘appraisal VTT’. This is a problem for appraisal but not for modelling, since demand prediction is concerned with obtaining the best possible predictions, for which in theory behavioural values are most suitable. However, appraisal is concerned with theoretical issues of welfare and income distribution, posing additional constraints for the selection of values. Thus, the translation of behavioural values into ‘acceptable’ *appraisal values* has been common since early VTT studies in the UK. This distinction, in turn, gives relevance to another: the distinction between ‘trip effects’ (e.g. mode comfort, distance, etc.) and ‘user effects’ (e.g. income, self-selection, etc.) as part of estimated values. However, as noted above, different countries have adopted different positions and the debate is ongoing (Börjesson and Eliasson, 2019).

This chapter has reviewed the state of play in this area nearly 10 years later, paying attention also to the modelling perspective on segmentation issues. Since 2015, the state of the art has not moved on significantly, with only a few direct contributions in this area identified – arguably Börjesson and Eliasson (2019) being the most relevant. On the other hand, the evolution of the policy context might suggest that some segmentation issues can become more prominent. There are certainly some lessons to be learned for a future UK VTT study.

Ultimately, there will always be choices associated with coming up with a representative VTT for each desired segment. The available choices can be influenced by the approach to data collection (e.g. more information from each user group will likely favour better identification of user effects). Also, there are two important perspectives to consider (modelling vs. appraisal), which may differ in terms of segmentation needs – or the perception of these needs by the respective analysts (modellers vs. appraisal practitioners). Even when segmentation problems are more prominent for appraisal, there is a value of consistency between modelling and appraisal methods.

From a modelling perspective, it should be kept in mind that most models use the minimum segmentation required by TAG and that any increased segmentation increases model run times and complexity. Any move to increased segmentation, based on evidence from the next UK VTT study, needs to be balanced against the limitations of the data used to build transport models. Models, which goal is prediction, have their workarounds to navigate heterogeneity (e.g. alternative specific constants) and do not necessarily have to worry about appraisal considerations. One solution might be for TAG to present VTTs for different levels of aggregation, with practitioners picking whatever level is appropriate for their application.

In contrast, from an appraisal perspective we may welcome more segmentation (e.g. by mode or distance), if issues from the past VTT studies can be resolved. Of course, as long as more VTT segmentation is expected to make a difference in appraisal (the proportionality principle). Apart from not segmenting the VTT by mode for commute and other non-work trips the current segmentation used in the UK is not too dissimilar from other countries where only Norway makes a more prominent use of segmentation by distance.

When revisiting the UK current position on VTT segmentation (including the issues exposed in the 2014/15 study) considering the more recent literature, the following conclusions are extracted:

- Segmentation in UK is broadly aligned with other countries with advanced appraisal guidelines. Trip purpose segmentation is well established and largely consistent with other countries (with a split between commute, leisure/other and business trips). Where the UK stands out is on the lack of segmentation by mode and distance (for commute and other trip purposes, i.e. effectively all trips except business). This should be considered as an area of priority.
- Also, for some reason, in the UK more consideration has been given to the identification/confounding issues around user vs. trip effects than in other countries. For example, it may be argued that the Swedish VTT study (Börjesson and Eliasson, 2019) faced the same segmentation challenges, but a view was taken that segmentation by mode and distance was justifiable. This comparison can shed some light and inform future judgements and trade-offs for a future study. We may hypothesise whether the different positions are driven mainly by different philosophical standpoints, or whether there are fundamental differences between countries around income inequality and transport costs across modes (e.g. rail fares) which accentuate the importance of equity considerations in UK vs. Sweden.
- Notable advances have been made in the identification of mode effects, and their separation from user effects, particularly the work by Flügel (2014) has been central to these developments. This requires the observation on the VTT in at least two modes for a given user group (i.e. a survey would need to intercept a person on mode A to ask about their VTT, but also ask about their VTT for mode B or C). Such data enables estimation of a weighted average VTT in a mode over all relevant user groups – referred to as the ‘Mode-Effect-dependent equity value’ approach. This would help address some of the challenges faced in the 2014/15 study, facilitating mode segmentation while retaining the prevailing position on distributional concerns (i.e. by controlling for user effects)
- Other studies have explored identifying mode effects using expenditure diary data (a more costly endeavour than a SP study), providing evidence on mode VTT variation that suggests this source of segmentation is likely to make a difference in appraisal. This form of data has not yet been collected in the UK and presents an interesting avenue for future research but likely falls outside the scope of a future VTT study due to the amount of personal data which needs collecting.
- Other important issues at the time, such as the ‘distance effect’ (and the strong relationship to reference cost and time) do not seem to have received much attention in the last decade. It is likely that similar issues would arise in a new study, and therefore we have provided some further reflections upon the 2014/15 challenges on this front. The empirical separation of distance, reference time and reference cost effects is arguably impossible due to the very high correlation between these three variables. However, this may not be necessary. The goal should be to cater for substantive variability among different types of journey, including short vs. long distance trips.
- As it turns out, these ‘distance issues’ come to life particularly in cross-mode comparisons, due to mode differences in speed, especially for active modes of transport. A first implication is that distance effects may be better explored and estimated on a mode by mode basis – of course, once the mode segmentation question is resolved. A second point to consider is

sampling: the 2014/15 study struggled to sample sufficient short-distance trips, consequently affecting the ability to estimate distance-relationships in the analysis. Thus, it is arguably the case that problems of mode and distance segmentation are inter-related, and both will benefit from paying increased attention to the sampling strategy.

- Provided that the sampling strategy enables a sound segmentation of VTT by modes and distance, in principle a continuous function for the VTT with distance would be preferred from a modelling point of view (and this would not cause any real issues from an appraisal perspective – e.g. current business VTT are already segmented by distance using a continuous function)
- So far, we have treated the issue of ‘small time savings’ as out of scope. However, we would like to note a potential interdependency with the distance issue. Namely, seeking realism in the SP designs means that, often, small changes are more often shown on short trips (and large changes on long trips). The SP designs should be designed carefully to ensure that ‘delta T’ heterogeneity can be clearly disentangled from distance related heterogeneity.
- Like mode and distance impacts, the treatment of income remains a challenging subject in estimation and in appraisal. Regarding the former, accurate measurement of income and relevant travel budgets remains an issue, especially due to the use of income categories in most data collection exercises. This issue is unlikely to be solved due to inherent increases in dropout rates from respondents. As such, less emphasis should be placed on the relatively low estimates of cross-sectional income elasticities widely observed. Regarding the latter, deriving appraisal VTTs based on behavioural VTTs for relevant segments appears the preferred approach and is consistent with current practice.
- Segmentation of the VTT for a given mode and journey purpose by income group – as an intermediate between appraisal VTTs and behavioural VTTs - may represent one way by which differences in the distributional impacts from transport investments can be articulated. This approach, however, risks prioritising schemes targeting higher income groups due to increased Value for Money (or efficiency) levels associated with such schemes. Indeed, the continued use of a uniform equity-weighted VTT – as derived through the sample enumeration process – circumvents these problems but loses part of the distributional aspect of the analysis. An alternative approach to correct for the higher monetary benefits associated with wealthier parts of the population is the implementation of distributional weights at the level of the scheme, not only at the level of travel time savings. Distributional weights shift the focus away from a value for money perspective to a social welfare function perspective.
- Where a new national VTT study is inherently focused on articulating the benefits from transport investments on time savings (and related components) in isolation from other benefits and not the appraisal of given schemes, our recommendation is that distributional weights should not play a role a new national VTT study, but that distributional analyses can be facilitated by improving our understanding of the relationship between income and the VTT both within and across modes of transport.

The above identifies various avenues for improvements in a future VTT study, with some relatively low-hanging fruits achievable through a refined survey and sampling strategy. Going forward, it is useful to think through policy relevant cases where the appraisals are suspected to be biased due to the segmentation approach chosen for VTT (some cases discussed in the chapter include urban realm or active modes, given the potential role of distance and mode effects). What are the policies/projects



that we are going to have to appraise? And are the available VTTs fit for purpose? For instance, Börjesson and Eliasson (2019)'s conclusions about VTT segmentation are context-specific. Their recommendations vary depending on whether transport prices are under public control and can be modified by the intervention, and on whether benefits are transferred from travellers.

Finally, connected to all the above, is the underlying policy position. A move towards a more WTP-driven approach can increase precision in the appraisals, but this would have to be balanced against the possibilities/impacts in modelling and against equity concerns as articulated above.

## 5. Valuing savings in business value of travel time and worthwhile use of time

This Chapter reviews the valuation of business travel time (VBTT), where the productive use of travel time has received significant attention over the past decades. Building on the productive use of travel time, the second half of the chapter delves into the worthwhile use of time across different modes of transport and journey purposes other than business. In this context specific attention is paid to the VTT associated with autonomous vehicles where one of the main perceived benefits is that users can benefit from doing different activities whilst being driven to their destination.

### 5.1 The history of valuing savings in business travel time

As the basis for official valuations of travel time savings, the Department for Transport has previously commissioned three major national studies, as reviewed in **Section 2.1**. All three studies considered, in some shape or form, the value of business travel time savings (VBTT), where the latter is defined as time spent travelling during an employer's business activities. In contrast to non-work (i.e. commute and leisure) travel, VBTT will be largely driven by the benefits to the company on whose behalf the journey is being made, as opposed to being a matter of personal preference and willingness to pay.

### 5.2 The alternative methodologies to valuing savings in business travel time

In the academic and policy literature, essentially three methodologies have been proposed for estimating VBTT, namely:

- Cost-Saving Approach (CSA)
- Hensher Equation (HE)
- Willingness-To-Pay (WTP)

#### 5.2.1 The Cost-Saving Approach

The CSA was the formative approach (Harrison, 1974), and continues to be the recommended approach in the official appraisal practice of most countries. It rests upon a set of assumptions which give rise to a VBTT closely related to the wage rate, namely:

- a) productivity is a function only of work time (thus implying a short run perspective, where capital is fixed, and labour is the sole variable input);
- b) perfectly competitive goods and labour markets (implying that the price of the final good is equal to the value of the marginal productivity of labour producing that good);
- c) no indivisibilities in the use of work time for production (such that every additional minute allocated to work derives a productive benefit, and the marginal product of labour is well-defined);
- d) all released time goes into work, not leisure (implying that any time saved in another activity, such as travel, is re-assigned to work and thereby generates productive benefit);
- e) travel time in the course of work is itself unproductive (although there must of course be a productive benefit from undertaking such journeys);
- f) travel time in the course of work generates neither utility nor disutility for the employee.

In practice, a wage plus approach is implemented, whereby the resource cost of time spent travelling in the course of work (i.e. excluding commuting) is calculated as being equal to the gross wage rate plus non-wage labour costs (e.g. national insurance, pensions, etc.), thus:

$$VBTT = w + c \quad (1)$$

where:

- $w$  is the gross wage

- $c$  is non-wage labour costs.

The inference of some of these assumptions, notably that travel time is ‘wasted time’, has provoked considerable debate in the literature (e.g. Cornet et al., 2022).

To a greater or lesser degree, the HE and WTP approaches purport to relax these assumptions, although the theoretical underpinnings of both approaches have hitherto remained somewhat opaque. The contributions of Kato (2013), Batley (2015) and Batley et al. (2023) have however made progress towards better understanding the mathematical derivation of the HE. In essence, these derivations relax assumptions d), e) and f) of the CSA and, more generally, substantiate Wardman et al. (2015)’s assertion that the HE gives rise to a generalised framework, from which the CSA and WTP can be derived as reduced forms.

### 5.2.2 The Hensher Equation

The HE evolved from a body of theoretical, conceptual and empirical work, which was developed by David Hensher whilst sub-contracting to consultants Travers Morgan, and subsequently documented in Carruthers and Hensher (1976) and more definitively in Hensher (1977). Some years later, Fowkes et al. (1986) proposed the following mathematical formulation based on selective elements of Hensher’s approach – and it is this formulation which is now commonly referred to as the ‘Hensher Equation’<sup>16</sup>:

$$VBTT = (1 - p \cdot q - r) \cdot MPL + MPF + (1 - r) \cdot VW + r \cdot VL \quad (2)$$

where:

- $p$  is the proportion of business travel time saved that would have been spent working;
- $q$  is the productivity of working whilst travelling relative to working at the workplace;
- $r$  is the proportion of business travel time saved that is allocated to leisure;
- $MPL$  is the value of the marginal product of labour;
- $MPF$  is the value of extra output due to reduced travel fatigue;
- $VW$  is the difference between the employee’s valuations<sup>17</sup> of ‘contracted’ work time and business travel time;
- $VL$  is the difference between the employee’s valuations of leisure time<sup>18</sup> and business travel time.

The  $p$  and  $r$  parameters are distinguished from their counterpart  $p^*$  and  $r^*$  parameters, where the latter refer to overall business travel time and the former refer to savings in business travel time.

It should be acknowledged that the full HE formulation is more a theoretical construct than an empirical one. As discussed in Mackie et al. (2003), the  $MPF$  term is usually omitted because of the difficulty of measuring, whilst some researchers have introduced other restrictions. Wardman et al. (2015) summarised three reduced forms which have been most actively used in practice, referred to as HE1-3.

#### HE1

<sup>16</sup> This chronology and attribution has been previously verified in private communication with David Hensher.

<sup>17</sup> In utility terms, excluding any wage income.

<sup>18</sup> Again in utility terms, pertaining to the residual of the time budget after subtracting contracted work.

Advocated by Fowkes (2001) and employed for a time in Norway, HE1 assumes that employees are indifferent between travelling in the course of work and working at the workplace<sup>19</sup>, such that  $VW = 0$ , in which case:

$$VBTT = (1 - p \cdot q - r) \cdot MPL + r \cdot VL \quad (3)$$

Since empirical evidence is lacking,  $VL$  is commonly approximated by the value of leisure travel time savings ( $VLTT$ ) when applying (2) in practice.

#### HE2

Currently employed in Sweden for rail Trafikverket (2024), HE2 assumes that the majority of business travel is undertaken during the ‘standard’ working day, such that  $r = 0$ , in which case:

$$VBTT = (1 - p \cdot q) \cdot MPL \quad (4)$$

#### HE3

Previously employed in the Netherlands (Hague Consulting Group, 1990, Hague Consulting Group, 1998, Significance et al., 2013) and Sweden (Algers et al., 1995), HE3 assumes that both  $VW$  and  $VL$  are based on employees’ valuations of travel time savings for their own benefit, such that  $VW = VL = VLTT$ , in which case:

$$VBTT = (1 - p \cdot q - r) \cdot MPL + VLTT \quad (5)$$

Another permutation is to assume  $p = r = 0$  and  $VW = 0$ , in which case (1) reduces to the CSA. Wardman et al. (2015) commented that: *‘Of the three commonly used restricted HE forms, conceptually our preference is for HE1. This is on the grounds that it is least restrictive, albeit with the most significant challenges of parameter estimation. HE3 is only valid where it can be demonstrated that  $VW=VL$ ’* (p203).

#### HE4

More recently, Fosgerau (2019) proposed a further formulation where  $r = 0$ , here referred to as HE4:

$$VBTT = (1 - p \cdot q) \cdot MPL + VW \quad (6)$$

It perhaps should be qualified that Fosgerau did not demonstrate, nor assert, that HE4 is a reduced form of the HE, but **Section 5.4** will elaborate upon this further. A very similar formulation is currently used in Norway Flügel et al. (2020).

#### 5.2.3 The Willingness-to-pay approach

In principle, the willingness-to-pay approach simply involves inviting the relevant economic agents to declare/reveal their WTP for saving journey time in the course of business, i.e.

$$VBTT = WTP \quad (7)$$

where WTP equals the willingness-to-pay to save journey time in the course of business.

On the face of it, the approach is intuitively simple. In principle, it captures all of the relevant benefits of saving travel time and avoids the complications of estimating the numerous distinct terms of the

---

<sup>19</sup> It is worth highlighting that, following the Covid pandemic, there has been a marked shift in homeworking (Office for National Statistics, 2022), such that the ‘workplace’ may increasingly be the home rather than the employer’s premises.

Hensher equation. In practice however, implementation of the approach encounters several questions and challenges. The most fundamental question is that of identifying the relevant economic agent(s) to survey – should it be the employer or employee? Theory suggests that, in this context, the employer's interests predominate, since it is the employer who decides how many labour units to employ, and how to utilise those units. Indeed, it is sometimes argued that, provided business travel time and work are explicitly linked (in the form of overtime for any travel encroaching upon leisure) then, apart from a tax adjustment, all costs or benefits will fall ultimately to the employer.

### 5.3 Application of the WTP approach in the 2014/15 study

Much of the background thinking to the current recommended approach to valuing business travel time savings was undertaken in the scoping study (ITS, 2013) which preceded the 2014-15 study (ARUP et al., 2015a). In this scoping study, reservations were expressed regarding the CSA traditionally employed by the Department, essentially reiterating long-standing and well-rehearsed concerns that not all travel time is unproductive and not all the time savings would be converted into productive use to the benefit of the company. It was highlighted that the digital revolution has increased the potential for using travel time productively. Other arguments against the CSA were cited, concerning difficulties in estimating the value of the marginal productivity of labour (which underpins the approach), the benefits of spending more time at the destination (say with a client or at a sales pitch), and the benefits of avoiding overnight accommodation and travel in unsocial hours. It was argued that, by contrast, these effects should in principle be captured by a WTP-based valuation, thereby eliciting a reliable representation of what the company would pay (ITS, 2013).

ITS (2013) considered that an intuitively appealing approach would be to survey employers about how much they would be prepared to pay to reduce their employees' travel time – bearing in mind that it would be employers who would be purchasing the time savings. At the same time, the difficulties of, and uncertainties surrounding, a valuation approach based on surveying employers were recognised. For example, the data collection costs are typically higher and there are challenges involved in identifying the appropriate employer agent – making it difficult to achieve a representative sample of travel-using employers.

The 2013 study recognised that a potentially complementary approach would be to undertake employee surveys, couched within an awareness of company travel policy. Compared to collecting employers' surveys, obtaining large samples of employees travelling on company business is relatively straightforward. The potential downside is whether employees can make choices in response to hypothetical scenarios that accurately represent the company's interests and willingness-to-pay.

All things considered, a preference was expressed for WTP-based approaches alongside methods for corroboration and interpretation. This reflected a proposition that well designed and conducted quantitative research could provide a coherent 'story' as to how business travel time savings are valued, or better still, can elicit direct estimates of WTP that lend themselves to comparison against the CSA.

Against this background, the business travel component of the 2014-15 study was informed by three sources of survey evidence, namely employer SP, employee SP, and employee RP. The information collected on income and working hours during the survey also allowed comparison with the CSA. To reconcile these different perspectives on business VTT, ARUP et al. (2015a) pursued two lines of enquiry. First, the degree of similarity between SP-based estimates of VTT and the CSA were

explored.<sup>20</sup> Second, the degree of consistency between various properties of the VTTs emanating from the different surveys was explored.

Generally speaking, ARUP et al. (2015a) found similar values for the two different SP analyses (employer and employee), and for some occupational types similar values were elicited from SP and the CSA. This was particularly so for blue collar workers, who would be expected to have relatively low productivity whilst travelling. For briefcase travellers, who are more likely to be productive, SP-based values appeared to be lower than the CSA. Moreover, the degree of similarity between the SP-based VTTs and the CSA was partly dictated by the trip length distribution and did not hold all over all distances. The self-employed values were lower than those for employees; whilst we cannot substantiate empirically, this result would seem plausible if the time saved is taken as leisure.

Turning to the properties of the VTT estimates, the theoretically-driven CSA embodies an income elasticity of one (i.e. implying that VTT increases in direct proportion to income) and applies a constant unit value to all trips (e.g. irrespective of time, cost, distance, travel conditions, productivity, etc.). By contrast, the SP-based VTTs exhibited income elasticities within the range 0.3 to 0.4 (and significantly less than one), and significant variability by several of the aforementioned dimensions (and notably by distance). Thus, whilst there was some correspondence between the actual estimates of VTT from the CSA and SP analyses, this correspondence did not extend to key properties of those estimates.<sup>21</sup>

Having reconciled the various sources of evidence on business VTT through the lines of enquiry summarised above, Arup et al. and the DfT decided that the employee SP survey should be the definitive source of evidence taken forward to the Implementation Tool. This was because it generated – with some qualifications – similar values to the employer SP survey, but offered a considerably more substantial dataset, amenable to generating statistically robust values for a range of trip and traveller segments. Furthermore, the employee dataset was more comparable to the NTS data used as the basis for the sample enumeration. That is to say, the Tool applied the choice model from the employee SP to business trips in the NTS, to derive an average value over specified segmentations, as shown in **Table 18** (see **Section 2.2** for a discussion on the role of the implementation tool in the sample enumeration process).

The average distance-weighted personal income across the NTS sample is £46,615 (2014 prices and values). This gives a business VTT in 2014 perceived prices of £28.27 using the CSA (second row, third column). This compares to the previous CSA-based TAG values which had an all modes value of £25.47 (first row, third column). If we also compare these values to the SP-based values re-weighted for NTS (third row), then we see that the VTT for employees' business across all modes is £18.23. This is 72% of the TAG value. We also find substantial variation by mode with 'other PT' lowest at £8.33 and rail highest at £27.61 for the all-distance values. As proportions of the TAG values, these range from 34% ('other PT') to 92% (rail).

**Table 18** shows that the SP-based values from the 2014-15 study are substantially less than the previous CSA-based TAG values, particularly for short distances, but as trip distances increase the SP-based values increase to be close to the previous TAG values at long distances (>50 miles).

---

<sup>20</sup> The RP-based estimates were largely used as a corroborative device and will not be discussed in any detail here.

<sup>21</sup> **Section 4.5** includes a discussion on the relatively low income elasticities found in most cross-sectional VTT studies using SP research.

Table 18: BVTT by method of calculation, mode, and distance (2014 perceived prices and values (£/hr)

| Source/method                                                                | Distance      | All modes | Car   | Bus   | 'Other PT' | Rail  |
|------------------------------------------------------------------------------|---------------|-----------|-------|-------|------------|-------|
| Previous TAG (2014 prices and values)                                        | All distances | 25.47     | 24.43 | 15.64 | 24.72      | 30.07 |
| CSA estimate from NTS 2010 to 2012 data (2014 prices and values)             | All distances | 28.27     | 27.05 | 13.13 | 26.33      | 36.46 |
| Employees' business SP re-weighted to NTS 2010-2012 (2014 prices and values) | All distances | 18.23     | 16.74 | N/A   | 8.33       | 27.61 |
|                                                                              | <5 miles      | 5.39      | 5.27  | N/A   | 8.33       | n/a   |
|                                                                              | 5-20 miles    | 8.84      | 8.79  | N/A   | 8.28       | 10.19 |
|                                                                              | >=20 miles    | 21.14     | 19.51 | N/A   | N/A        | 28.99 |
|                                                                              | >=50 miles    | 24.55     | 22.53 | N/A   | N/A        | 32.56 |
|                                                                              | >=100 miles   | 28.62     | 25.74 | N/A   | N/A        | N/A   |

Notes: All modes. Distance-weighted, income option 1, SP1  $\Delta t=10$ ; Tool version 1.1. PT cost is imputed for a trip with a zero cost, and employers paying for EB trips. TAG 'Other PT' is Underground passenger, TAG car EB is weighted average of driver and passenger (vehicle occupancy of 1.2).

#### 5.4 Application of the Hensher Equation to the 2014-15 data

Batley et al. (2023) revisit the Hensher Equation and applies it to the 2014-15 dataset. More specifically, the paper combines theory and empirics to elicit new insights on the Hensher Equation for estimating the value of business travel time savings. Several variants of the Hensher Equation are derived from first principles and their properties demonstrated; a formulation is recommended which assumes equilibrium in the allocation of work and leisure time, since this maximises the social welfare benefits of travel time savings.

Using data from ARUP et al. (2015a), a modelling exercise explores variations in the Hensher parameters for car and rail journeys, demonstrating that they can vary considerably depending on the features of the journey and traveller. Finally, the theoretical and empirical contributions of the paper are combined to elicit new estimates of the value of business travel time savings, reporting both nationally representative average values for the UK as well as variations across different journeys and travellers.

The paper concludes that:

- The HE can be derived theoretically from a reasonably standard social planner model.
- Its underpinning properties and assumptions are reasonable and arguably more intuitive than those applying to the competing CSA and WTP approaches – and especially so for the case of briefcase travellers.
- The theoretical distinctions between alternative reduced forms of the HE can be clarified; such that definitive guidance can be given to practitioners on what variants are applicable to what practical contexts. In particular, the paper advocates the HE4 formulation, although the absence of empirical evidence on the  $VW$  term means that, in practice, the analyst must defer to HE2 as a lower bound approximation to HE4.
- Sufficient empirical evidence exists or can be readily obtained to support practical application of relevant HE variants, including insights on variations in VBTT by features of the business traveller and journey.

Batley et al. (2023) report the following findings in respect of the  $p^*$  and  $r^*$  parameters:

- i) First and foremost, the HE parameters exhibited considerable (and generally plausible) variability across contexts. Key covariates were travel time, income and whether travel costs were reimbursed – plus various ad hoc features of the traveller and/or journey. This highlights the need to better understand the key dimensions of variability, since this will strengthen the spatial and temporal transferability of the parameters and improve the accuracy of the resulting estimates of VBTT.
- ii) One of the factors accounted for by Batley et al. (2023) is journey time, which in line with the results for WTP reveals that trips of shorter duration are associated with a lower VBTT than trips with a longer duration.
- iii) The  $p^*$  and  $p$  values were found to be larger for rail than car (and often somewhat larger, especially for  $p^*$ ), increasing with hours worked and income, greater for outward than return travel and for solus than group travel, whilst also exhibiting some variation by journey duration.
- iv) The  $r^*$  and  $r$  values were consistently found to exceed zero – indicating disequilibrium in the allocation of work and leisure time. The values were somewhat larger for rail, larger for employees working longer hours (although this was offset to some extent by a reimbursement effect), whilst also exhibiting some variation by journey duration and leg.
- v) In line with expectations, the  $p^*$  value almost always exceeded the corresponding  $p$ , and  $r$  always exceeded  $r^*$ .

The paper then goes on to report new HE-based estimates of VBTT for the car and rail modes from the 2014-15 dataset. This encompasses both ‘starred’ (i.e. referring to total travel time) and ‘unstarred’ (i.e. referring to travel time savings) versions of the HE1-3 formulations, in combination with the empirical variations in the HE parameters. Two sets of values are reported, firstly average values for a representative sample of travellers (**Table 19** and **Table 20**), and secondly variations across a wide range of traveller/journey scenarios. It is found that:

- vi) VBTT exhibited considerable variability, not only across different business travellers and journeys, but also across the different formulations of the HE. This highlights the importance of employing the most appropriate formulation, as well as adapting that formulation to the practical context of interest.
- vii) In terms of travel time savings, HE2 gave the highest average values for both car and rail, respectively 91% and 80% of the MPL. Analogously, in terms of total travel time, HE2\* gave the highest average values for both car and rail, respectively 90% and 55% of the MPL.
- viii) Except for HE2, the average VBTT was somewhat lower for rail than car. It was also considerably lower for reimbursed employees than non-reimbursed.
- ix) Since both  $p^*$  and  $r^*$  were found to increase with business travel time for both rail and car, VBTTs exhibited a decreasing relationship with travel time. It should be qualified that this relationship highlights a potential endogeneity within the Hensher framework that warrants further investigation.
- x) For car, the VBTTs pertaining to total travel time (i.e. starred) was generally higher than that pertaining to travel time savings (i.e. unstarred), both in absolute terms and relative to the



MPL; the reverse was true for rail. Arguably, the unstarred version should be adopted in appraisal guidance since policy is typically focussed upon travel time savings.

- xi) Since some of the HE2 and HE2\* values fell just short of the MPL, it is quite conceivable that, if  $VW$  could be estimated, then the (theoretically preferred) HE4 and HE4\* values could in some scenarios exceed the MPL.

Table 19: Representative average VBTT for car (£/hr at 2014 market prices)

|       | HE1    | HE1/MPL | HE1*  | HE1*/MPL | HE2   | HE2/MPL | HE2*  | HE2*/MPL | HE3   | HE3/MPL | HE3*  | HE3*/MPL |
|-------|--------|---------|-------|----------|-------|---------|-------|----------|-------|---------|-------|----------|
| VBTT  | 19.33  | 0.53    | 27.24 | 0.75     | 33.03 | 0.91    | 32.58 | 0.90     | 22.66 | 0.62    | 32.28 | 0.89     |
| $n$   | 19,809 |         |       |          |       |         |       |          |       |         |       |          |
| $p$   | 0.10   |         |       |          |       |         |       |          |       |         |       |          |
| $p^*$ | 0.11   |         |       |          |       |         |       |          |       |         |       |          |
| $q$   | 0.90   |         |       |          |       |         |       |          |       |         |       |          |
| $r$   | 0.45   |         |       |          |       |         |       |          |       |         |       |          |
| $r^*$ | 0.17   |         |       |          |       |         |       |          |       |         |       |          |
| MPL   | 36.27  |         |       |          |       |         |       |          |       |         |       |          |
| TAG   | 17.58  |         |       |          |       |         |       |          |       |         |       |          |
| VLTT  | 6.10   |         |       |          |       |         |       |          |       |         |       |          |

Notes: <sup>a</sup>All values are expressed in UK £/hr at 2014 market prices. <sup>b</sup>TAG refers to the UK Department for Transport's official VBTT given in TAG.

Table 20: Representative average VBTT for rail (£/hr at 2014 market prices)

|       | HE1   | HE1/MPL | HE1*  | HE1*/MPL | HE2   | HE2/MPL | HE2*  | HE2*/MPL | HE3   | HE3/MPL | HE3*  | HE3*/MPL |
|-------|-------|---------|-------|----------|-------|---------|-------|----------|-------|---------|-------|----------|
| VBTT  | 17.74 | 0.35    | 14.30 | 0.29     | 40.00 | 0.80    | 27.75 | 0.55     | 22.41 | 0.45    | 21.41 | 0.43     |
| $n$   | 1119  |         |       |          |       |         |       |          |       |         |       |          |
| $p$   | 0.20  |         |       |          |       |         |       |          |       |         |       |          |
| $p^*$ | 0.46  |         |       |          |       |         |       |          |       |         |       |          |
| $q$   | 0.90  |         |       |          |       |         |       |          |       |         |       |          |
| $r$   | 0.57  |         |       |          |       |         |       |          |       |         |       |          |
| $r^*$ | 0.34  |         |       |          |       |         |       |          |       |         |       |          |
| MPL   | 50.05 |         |       |          |       |         |       |          |       |         |       |          |
| TAG   | 29.46 |         |       |          |       |         |       |          |       |         |       |          |
| VLTT  | 10.79 |         |       |          |       |         |       |          |       |         |       |          |

Notes: <sup>a</sup>All values are expressed in UK £/hr at 2014 market prices. <sup>b</sup>TAG refers to the UK Department for Transport's official VBTT given in TAG.

## 5.5 Capturing of VBTT in other national VTT studies

**Section 2.4** already provided an overview of the VBTT values elicited and indicated the methods used in the most recent European national VTT studies. The German VTT study only adopted the WTP approach.

Significance (2023) implement a split sample approach in the Netherlands where half of the business travellers is allocated to the HE approach and half is allocated to the WTP approach. The reason to adopt a dual approach is that the recent literature, and other recent national VTT studies, tend to adopt the WTP approach, whilst the previous national VTT study relied on the HE for the employer component of the VBTT. For internal consistency purposes and aligning with the wider VTT literature the dual approach was taken forward. The recommended values are ultimately based on the WTP estimates, because both the outcomes of the HE and the WTP-based approach were considered plausible. Detailed estimates for the parameters of the Hensher equation are included in Appendix B of Significance (2023).

In practice, this meant that respondents were allocated to the SP part of the survey into four different groups each using a different description in the SP preamble as to who (self or employer) would benefit/pay for changes in travel time and travel costs.

When comparing the results of the Hensher Equation with those from the WTP-based approach, the WTP estimates are roughly 30% lower than those of the Hensher equation for all modes (Car, rail, and local PT) but air. For the latter, the WTP value is roughly 40% higher. When contrasting the results from the Hensher equation with those from the previous national VTT study, the newly elicited values have decreased by 7 (car) to 21 (rail) per cent. For reference, the Hensher Equation specification adopted in the Netherlands is HE2.

When Norway still made use of the Hensher Equation it relied on HE1, but they moved to the CSA approach in 2008-2010 (Samstad et al., 2010). In the most recent study, a WTP approach is adopted, and the presented evidence is consistent that the VBTT is increasing in journey length, which is consistent with Batley et al. (2023)'s observed correlation of the VBTT with travel time. Note, however, that Table 3.2 in Flügel et al. (2020) reports results for the p and q parameters of the Hensher equation, but not the resulting HE-based VBTT. The German VTT study entirely makes use of the WTP approach and does not estimate the components of the Hensher Equation (Dubernet and Axhausen, 2020).

### 5.6 Conclusions on the VBTT

Based on the material presented on the VBTT, it can be concluded that the state of practice in national VTT studies is the use of the WTP approach. The challenge with the method, however, is in defining and approaching the relevant agent to survey. Where it is relatively easy to contact employees, it is rather challenging to survey employers and within those organisations to obtain access to the relevant staff responsible for business travel strategy and policy.

In terms of alternative approaches, the Hensher Equation is preferred over the Cost-Savings Approach because there is general acknowledgement that not all travel time is unproductive. Moreover, most national VTT studies still elicit the core parameters of the Hensher Equation highlighting its relevance in the literature.

Different functional forms exist for the Hensher Equation varying in their degree of practical applicability. Recent work by Batley et al. (2023) suggest that the most preferred approach from a theoretical perspective is HE4, but that from a pragmatic perspective HE2 is preferred and should be treated as a lower bound on HE4 due to the absence of empirical evidence on the VW in HE4.

### 5.7 Worthwhile use of travel time – general state of research

Productive use of travel time is not restricted to business travel and also applies to non-business travel. Research concerning the role of travel time use, its driving factors and VTT impacts has accelerated over the past two decades with increased use of information and communication technologies (ICT) for work and non-work activities.

The first systematic review presented by Keseru and Macharis (2018) covers 58 studies that look at travel-based multitasking. The review paper led to a summary cross-tabulation (**Figure 15**) of factors affecting participation in various activities whilst travelling. At the time of publication, they concluded that consistent evidence (at least 3 studies) existed concerning the effects of:

- a. Age
- b. Gender
- c. Trip duration
- d. Travel mode
- e. Trip purpose
- f. Time of the day
- g. Day of the week
- h. Presence of travel companion.

Another aspect discussed by the authors concerns the taxonomy surrounding the notion of 'travel-based multitasking'. This discussion fits a long-standing debate concerning how participation in activities while travelling should be named. Terms such as '(productive) travel time use' and 'travel time activities' were used in the early days of the discourse, driven by the critique of the conventional assumption of travel time being 'wasted' and to be minimised. The term 'travel-based multitasking' emerged as a competing notion to describe the same phenomenon in the early 2010s, driven by observing that conducting activities during travel is an example of the broader concept of multitasking, recognised and researched in time use studies, largely sociological, for decades.

The term 'travel-based multitasking' has been used, however, in a somewhat confusing manner. Specifically, its somewhat lax meaning does not specify whether travelling is an activity by itself, in which case any activity undertaken while travelling is, indeed, multitasking by definition, or whether travel represents more of a context or location (especially if not requiring any cognitive effort from the traveller) in which case it becomes a 'location'. The latter perspective has prevailed in time use data, where travel modes are customarily included amongst 'locations'. In this interpretation, participation in one activity whilst travelling can be interpreted as *monotasking*, leading to erroneous reporting.

The concept of multitasking can become even more obscure, when considering activities undertaken more passively, such as listening to music or monitoring e-mail, while doing something else and travelling. From the point of view of study into the VTT, introduction of the term in the data collection or reporting could lead to confusion and would not be recommended, apart from acknowledging existence of the nomenclatural dichotomy.

The review also pointed out the lack of standard categories of activities undertaken while travelling, throughout empirical research. The actual activity categorisation is typically ad-hoc, driven either by the method of data collection (e.g. for observational studies categorisation follows activities easy to observe, for questionnaire-based research it may be more detailed but cannot result in an overburdening complexity), based on time use surveys (resulting in broad categories) or based on particular objectives of a given study and in practice devised by the authors to fit the analytical (modelling) requirements.

The review also pointed out the lack of consensus in how to report on empirical findings concerning the worthwhile use of travel time. In particular, studies variously report on their findings, partially as a result of the nature of data analysed, including combinations of activity types, their durations, number of episodes, making it often infeasible to directly compare observations between the studies.

Another systematic review of interest by Pawlak (2020), covering 77 studies, though focused more narrowly on 'digital activities' and the role of internet connectivity in shaping the nature of the travel time. In this case, the motivation stemmed from firstly suggestions and subsequently the growing evidence concerning the role of internet connectivity in enabling travel time use and its productivity.

| Category    | Variable                                                | No relationship | Stat. sign. rel.* | Descriptive stat.** | Playing | Relaxing | Reading | Talking | Media use | Mobile phone use | Working/studying | ICT use | Eating/drinking | Other |
|-------------|---------------------------------------------------------|-----------------|-------------------|---------------------|---------|----------|---------|---------|-----------|------------------|------------------|---------|-----------------|-------|
| Attitudinal | Advance planning                                        | 0               | 1                 | 0                   |         |          |         |         |           |                  | 1                |         |                 |       |
|             | Desire or obligation to be available to others          | 0               | 1                 | 1                   |         | 1        | 1       |         |           | 1                | 1                | 1       |                 |       |
|             | Attitude to technology                                  | 0               | 1                 | 0                   |         |          | 1       |         | 1         |                  | 1                | 1       |                 |       |
|             | Propensity to use active modes                          | 0               | 1                 | 0                   |         | 1        |         |         |           |                  |                  |         |                 |       |
|             | Ownership of a car to impress others                    | 0               | 1                 | 0                   |         |          | 1       |         | 1         |                  |                  | 1       |                 |       |
|             | Frugality                                               | 0               | 1                 | 0                   |         |          |         |         |           |                  |                  | 1       |                 |       |
|             | Polichronicity                                          | 0               | 1                 | 0                   |         | 1        |         |         |           |                  | 1                | 1       |                 |       |
|             | Attitude to activities while waiting                    | 0               | 1                 | 0                   |         | 1        |         |         | 1         |                  |                  | 1       |                 |       |
|             | Desire to do recreational activities on the commute     | 0               | 1                 | 0                   | 1       | 1        | 1       |         | 1         |                  |                  |         |                 |       |
|             | Propensity to work during commute                       | 0               | 1                 | 0                   |         |          |         |         |           |                  | 1                |         |                 |       |
|             | Approval of other people                                | 0               | 1                 | 0                   |         |          |         |         |           | 1                |                  |         |                 |       |
|             | Risk of fines                                           | 0               | 1                 | 0                   |         |          |         |         |           | 1                |                  |         |                 |       |
|             | Perceived service value                                 | 0               | 1                 | 0                   |         |          | 1       |         |           | 1                |                  | 1       |                 |       |
|             | Norms of conduct (privacy)                              | 1               | 0                 | 0                   |         |          |         |         |           |                  |                  |         |                 |       |
|             | Organisational norms (having to be available for calls) | 1               | 0                 | 0                   |         |          |         |         |           |                  |                  |         |                 |       |
|             | Task type to be conducted                               | 0               | 0                 | 1                   |         |          | 1       |         |           | 1                | 1                | 1       |                 |       |
|             | Frequency of internet use                               | 0               | 1                 | 0                   |         |          |         |         |           |                  |                  | 1       |                 |       |
| Comfort     | Crowdness on vehicle                                    | 2               | 1                 | 0                   |         |          |         |         |           | 1                |                  |         |                 |       |
|             | Crowdness at stop                                       | 0               | 1                 | 0                   |         |          |         |         | 1         |                  |                  |         |                 |       |
|             | Wi-Fi connection available on vehicle                   | 1               | 1                 | 0                   |         |          |         |         |           |                  |                  | 1       |                 |       |
|             | Silent compartment                                      | 1               | 1                 | 0                   |         |          |         |         |           | 1                |                  |         |                 |       |
|             | Seated passenger                                        | 1               | 1                 | 0                   |         |          |         |         |           |                  |                  |         |                 | 1     |
|             | Jerkiness of the ride                                   | 1               | 1                 | 0                   |         |          | 1       |         |           |                  | 1                | 1       |                 |       |
|             | Noise level                                             | 2               | 0                 | 0                   |         |          |         |         |           |                  |                  |         |                 |       |
|             | Class of travel                                         | 1               | 0                 | 2                   |         |          | 1       |         |           |                  | 1                | 1       |                 |       |
|             | Connectivity (signal strength)                          | 0               | 0                 | 2                   |         |          |         |         |           | 2                |                  |         |                 |       |
|             | Space available                                         | 0               | 0                 | 2                   |         |          |         |         |           |                  |                  | 2       |                 |       |
|             | Seating comfort (availability of power socket, table)   | 0               | 0                 | 2                   |         |          |         |         |           |                  |                  | 2       |                 |       |
|             | Position in vehicle                                     | 0               | 1                 | 0                   |         |          |         |         | 1         |                  |                  |         |                 |       |
|             | Temperature                                             | 0               | 1                 | 0                   |         |          |         |         | 1         |                  |                  |         |                 |       |
|             | Position inside the vehicle                             | 0               | 1                 | 0                   |         |          |         |         |           |                  |                  |         |                 | 1     |
|             | Bus shelter available                                   | 1               | 0                 | 0                   |         |          |         |         |           |                  |                  |         |                 |       |
|             | Seats available at stop                                 | 1               | 0                 | 0                   |         |          |         |         |           |                  |                  |         |                 |       |
|             | Rainy weather                                           | 1               | 0                 | 0                   |         |          |         |         |           |                  |                  |         |                 |       |
|             | Windy weather                                           | 1               | 0                 | 0                   |         |          |         |         |           |                  |                  |         |                 |       |
|             | Location of seat (aisle, window)                        | 1               | 0                 | 0                   |         |          |         |         |           |                  |                  |         |                 |       |
| Equipment   | Equipped with laptop                                    | 0               | 1                 | 1                   |         |          |         |         |           |                  | 1                | 1       |                 |       |
|             | Passenger used iPad                                     | 0               | 1                 | 0                   |         |          |         |         |           |                  |                  | 1       |                 |       |
|             | Passenger used smartphone                               | 1               | 0                 | 0                   |         |          |         |         |           |                  |                  |         |                 |       |
|             | Passenger used iPod                                     | 1               | 0                 | 0                   |         |          |         |         |           |                  |                  |         |                 |       |
|             | Hands-free kit availability (in car)                    | 1               | 0                 | 0                   |         |          |         |         |           |                  |                  |         |                 |       |

Figure 15: Overview of variables investigated for the relationship with travel-based multitasking (travel time use). Numbers indicate the number of papers in which a specific variable was investigated and the number of cases it was found significant. Source: Keseru and Macharis (2018).

One of the key observations of the review concerned the need to unwrap the term ‘*use of ICT*’, customarily encountered as one of the activities reported in research related to the use of travel time. In this instance, reporting was related to nature of the equipment used to conduct the activity, i.e. Information and Communication Technology, rather than purpose of the activity. Towards this end, the paper discusses conventions of reporting ICT-related activities, distinguishing hardware-centric and function-centric, with the former focused on the equipment, e.g. use of laptop, use of smartphone, whilst the latter is oriented more towards specific functions, e.g. using e-mail, watching a movie.

Arguably, the function-centric convention offers a better ability to discern whether an activity conducted in the course of travel is an effective substitute of an activity that could have been conducted without the travel, a key consideration for understanding the impacts of travel time activities on value of time. Nonetheless, a fine balance is required to provide opportunities for respondents to provide reasonably complete information on their digital activities conducted while travelling, such as ‘communication’, ‘word processing’, ‘learning’ with an indication of purpose (professional or leisure) as opposed to attempting to record specific types of applications individuals used, given the sheer number of such.

Another aspect discussed extensively in the review concerns the role of connectivity in enabling a wider range of activities to be undertaken while travelling. The study observed (**Figure 16**) how the notion of ‘digital activities’ in the context of travel time became coupled with connectivity around year 2010, providing evidence for an increasing interdependence between the two, itself resulting from an increasing reliance of digital activities on remote (web-based) services and communications. Indeed, this proposition was the basis of the so-called ‘GJT\_Trend’ term adopted in v6 of the Passenger Demand Forecasting Handbook.

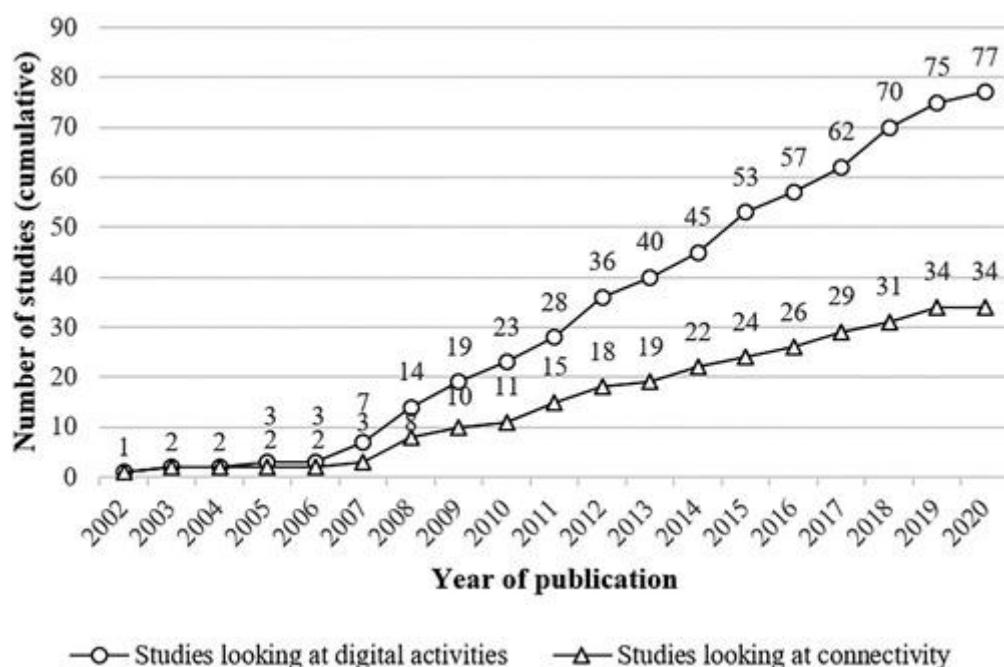


Figure 16: Studies looking at digital activities and the role of connectivity in the context of travel as reported in the review by Pawlak (2020).

The review observed that accounting for the presence of connectivity may need to reflect not just its nominal presence, but also attributes that make it a feasible resource to use while travelling, e.g. speed or reliability. This is an issue somewhat peculiar to transport contexts, where involvement of a moving

vehicle requires reliance on wireless technologies and their related limitations resulting from topography, proximity of ground infrastructure, speed of travel, density of demand or signal propagation and proneness to interference and attenuation. Without accounting for such factors, measuring and reflecting the effects of connectivity are likely to be incomplete or even biased.

Another important contribution, in fact preceding the above ones, summarising state of knowledge, came from Wardman and Lyons (2016). The authors, discussing the context of business travel, were very clear in re-stating the long-standing *'concern that the traditional Cost Savings Approach (prevalent across administrations that conduct cost-benefit analysis) overstates the benefits of travel time savings since it does not allow for the time savings encroaching upon productive use of travel time.'* (Wardman and Lyons, 2016 p.526). The authors provided the example of Sweden, where the value of time (savings) for rail was downrated to reflect the productive potential of travel time. They indicated that little research was undertaken by time of their publication concerning the impacts in the case of non-business travel time, though their expectation was the effects to be of similar nature.

As part of their meta-analysis, the authors summarised evidence concerning the Hensher parameters (**Figure 17**), specifically the  $p$ ,  $p^*$  and  $q$  terms. Their meta-analysis indicated that whilst the reported proportion of travel spent on working was comparatively low (typically below 30%), the relative productivity often resembled or even exceeded that of the office conditions, i.e.  $q > 1$ . Interestingly, even though the analysis included studies well prior to advent of modern mobile ICTs, the evidence of productive potential of travel time could already be observed.

Mokhtarian (2018) proposed the notion of travel becoming increasingly 'passengerised', which describes the impact of new products and services that enable travel time being used for activities. In particular, the author discusses how under various scenarios, proliferation of travel time activities can lead to changes in total travel time, depending on whether such activities are newly generated or whether they substitute activities undertaken at other times of day. The author acknowledges emerging evidence concerning the postulated impacts of multitasking, though also observing that studies indicated (in particular, the study by Malokin et al., 2017) that 'millennial commuters have lower values of travel time savings for both in-vehicle and out-of-vehicle travel time, even after controlling for personal attitudes and the propensity to multitask' (Mokhtarian, 2018 p.6).

Furthermore, the authors also suggested several possible research directions including:

- a. Despite presence of 'fully productive' activities while travelling, individuals are likely to have upper bounds concerning how long travel is. However, such bounds have not been looked at due to focus on, primarily, travel time reduction.
- b. To what extent the productive and hedonic use of travel time is a meaningful factor in travel decisions as opposed to being a by-product and reaction to circumstances? How is that going to vary across the population, depending on the profession, equipment at hand and skills to undertake activities while travelling.
- c. Equitability issues: are the benefits of enablement of travel time activities accruing equally across the population or will they benefit information jobs, managers and people whose professional duties or personal preferences allow them to participate in things while travelling more easily.
- d. How to quantify the benefits of activities that are seemingly 'unproductive' (boredom, idleness, daydreaming, sleeping) and yet can offer comfort or mental and physical benefits?

There could be instances in which enablement of travel time activities disbenefits travellers, such as resulting from the removal of the transition time between activities, contexts, and roles. How should that be accounted in the models?

| Reference                           | Study Date | Country     | Mode                | p    | p*   | q    |
|-------------------------------------|------------|-------------|---------------------|------|------|------|
| Hensher (1977)                      | 1977       | Australia   | Air-domestic        | -    | 0.36 | 0.61 |
|                                     |            |             | Air-international   | -    | 0.19 | 0.51 |
| Fowkes et al. (1986)                | 1986       | UK          | Car                 | -    | 0.03 | 1.01 |
|                                     |            |             | Train               | -    | 0.20 | 0.95 |
|                                     |            |             | Air                 | -    | 0.14 | 0.99 |
| Hague Consulting Group (1990)       | 1988       | Netherlands | Car                 | -    | 0.02 | 0.90 |
|                                     |            |             | Train               | -    | 0.11 | 0.89 |
|                                     |            |             | Bus                 | -    | 0.03 | 0.93 |
| Hague Consultin Group et al. (1999) | 1994       | UK          | Car                 | -    | 0.04 | 1.02 |
|                                     |            |             | 10-30 min           | -    | 0.03 | 1.11 |
|                                     |            |             | 31-60 min           | -    | 0.04 | 1.01 |
|                                     |            |             | 61-121 min          | -    | 0.05 | 1.00 |
|                                     |            |             | >120 min            | -    | 0.05 | 1.01 |
| Algers et al. (1995)                | 1994       | Sweden      | Car                 | -    | 0.14 | 1.01 |
|                                     |            |             | Company car         | -    | 0.19 | 1.11 |
|                                     |            |             | Car-Self employed   | -    | 0.10 | 1.20 |
|                                     |            |             | Train-inter urban   | -    | 0.28 | 1.03 |
|                                     |            |             | Train-Self employed | -    | 0.23 | 1.08 |
|                                     |            |             | Train-X2000         | -    | 0.28 | 1.04 |
|                                     |            |             | Train-urban         | -    | 0.18 | 1.15 |
|                                     |            |             | Bus-inter urban     | -    | 0.13 | 0.93 |
|                                     |            |             | Bus-suburban        | -    | 0.17 | 1.26 |
|                                     |            |             | Air                 | -    | 0.13 | 0.97 |
| Hage Consulting Group (1998)        | 1997       | Netherlands | Air-self employed   | -    | 0.11 | 1.02 |
|                                     |            |             | Car                 | -    | 0.04 | 0.93 |
|                                     |            |             | Train               | -    | 0.16 | 0.90 |
| Ramjerdi et al. (1997)              | 1997       | Norway      | Bus                 | -    | 0.03 | 0.89 |
|                                     |            |             | Car-inter urban     |      | 0.03 | 0.32 |
|                                     |            |             | Car-urban           |      | 0.21 | 0.02 |
|                                     |            |             | Train-inter urban   |      | 0.18 | 0.39 |
|                                     |            |             | Bus-inter urban     |      | 0.06 | 0.20 |
|                                     |            |             | PT-urban            |      | 0.30 | 0.07 |
|                                     |            |             | Air                 |      | 0.07 | 0.28 |
| Beca Carter Hollings et al. (2002)  | 2002       | New Zealand | Ferry               |      | 0.03 | 0.19 |
|                                     |            |             | Car                 | 0.03 | 0.22 | 0.93 |
| VSS (2009)                          | 2003       | Switzerland | Car                 | 0.64 |      | 0.98 |
| Mott MacDonald et al. (2009)        | 2008       | UK          | Train               | 0.41 |      | 0.97 |
|                                     |            |             | <45 min             | 0.60 | 0.46 | 0.98 |
|                                     |            |             | 45-89 min           | 0.35 |      | 0.97 |
|                                     |            |             | 90-149 min          | 0.28 |      | 0.98 |
|                                     |            |             | 150 min +           | 0.22 |      | 0.96 |
| Significance et al. (2012)          | 2011       | Netherlands | Car                 | -    | 0.04 | 0.91 |
|                                     |            |             | Train               | -    | 0.16 | 0.94 |
|                                     |            |             | Bus                 | -    | 0.06 | 0.83 |
|                                     |            |             | Air                 | -    | 0.14 | 1.00 |

**Table 1** Summary of p, p\* and q evidence

Figure 17: Hensher Equation parameters as reported in Wardman and Lyons (2016)

## 5.8 Worthwhile use of travel time – theoretical and conceptual developments

Several studies have sought to build theoretical underpinnings for the notion of travel time activities. The purpose of the current section is to demonstrate that travel time activities have been variously acknowledged and incorporated into theoretical frameworks underpinning travel behaviour models. In most instances, this is achieved by relaxing implicit assumptions, such that the models collapse to their original formulations under conditions of travel time activities being impossible, negligible, or not delivering any productive or utilitarian value to the travellers.

Banarjee and Kanafani (2008) extended the microeconomic time use theory, stemming from the works of Becker (1965), DeSerpa (1971) and Jara-Diaz and Guevara (2003), to incorporate the possibility of undertaking work during travel, characterised by relative efficiency factor (drawing upon Hensher's 'q' term). Their motivation was to capture the role of connectivity (Wi-Fi) on the ability to use time productively whilst travelling. By following the works of Train and McFadden (1978) in deriving appropriate indirect utility function associated with selecting a particular mode of transport, they demonstrated how their proposed extension leads to a logit formulation of mode choice where the role of on-board Wi-Fi and travel time activities is present in the modal utility functions.

Pawlak et al. (2015) proposed a microeconomic framework that conceptualises how people choose their departure time and activities while travelling, in the presence of ICT. Their framework draws upon Winston (1987)'s approach, relying on the concept of activity intensity and utility rate. With respect to the context of travel time activities, the contribution allows for the same activity to take place in the form of 'spells' during the travel, depending on the benefit to the user reflected in the utility and productivity. The contribution has also shown the means of establishing an indirect utility for a choice of a specific activity combination (including travel time activities), alongside departure time, considering variations in the duration of travel across different times of day. In terms of operationalisation provided in Pawlak et al. (2017), the model was based upon the hazard-based formulation for modelling travel time activities (work and non-work) as well as productivity (self-reported, in reference to office conditions), based on the Survey of Productive Use of Rail Travel-time (Fickling et al., 2009). It therefore demonstrated the feasibility of modelling multi-spell activities jointly with the experienced productivity using a paper-based questionnaire data, where participation in work activity was indicated by marking respective deciles of travel time and productivity reported in reference to standard (office) conditions. In an associated study, Pawlak et al. (2014) discussed the idea of an approach that focuses on the 'productive output' of travel as opposed to value of travel time savings, as a way of addressing the fallacy of investment appraisal contexts in which improved travel conditions is not well reflected in the VTT.

Pudāne et al. (2018) proposed a microeconomic time use model based focused on the context of autonomous vehicles (though extensible to other contexts). Their perspective stems from the observation that the, arguably simplified, postulated effect of travel time activities in the form of reductions in the size of the travel time are only a crude proxy for the variety of changes in time-use and travel patterns that are likely to occur at the advent of AVs. Their discussion is related to the question of whether the intended activities can be undertaken on-board but also whether the activity requires a designated location, and thereby travel at all. When implemented as a model, these considerations lead to potentially reductions in the disutility of travel. However, the reduced utility can in turn encourage longer travel to a destination with qualities superior to the destination that would have been selected without AV and travel time activities.

In 2018 International Transport Forum (ITF) hosted a workshop 'Zero Value of Time Roundtable' workshop (ITF, 2019). The roundtable featured 7 presentations which covered the broad set of



research that looked at the utility and disutility of time spent travelling, whether and how they should be measured and captured in the valuation of travel time. The specific presentations included:

- a. Fosgerau (2019) discussed implications for activity-travel behaviour analysis, such as trip scheduling, mode choice but also residential selection or congestion pricing, primarily in the context of vehicle automation. Based on crude assumptions concerning value of the work output delivered during lifetime of the AV use, he concluded that AVs may see low penetration rates for a long time given the substantial expected technology costs.
- b. Lugano (2018) shared, at that time, initial observations from the large scale MotiV project, highlighting the need to measure worthwhileness of travel time not only in terms of 'work-related' metrics, such as productivity, but also a wider set of needs that can be fulfilled by undertaking activities while travelling.
- c. Lee et al. (2020) provided evidence from Korea concerning the impact of improved connectivity on attractiveness of public transport. Such improvements were argued to be capable of attracting car users towards public transport, highlighting the modal shift potential of travel time use enablement.
- d. Meunier (2020) looked at how the approach towards use of VTT in transport appraisal practices in France has evolved since 1960s. He observed that the prevailing modelling methodologies may evolve in response to the changes brought by proliferation of mobile ICTs, suggesting amongst others, the possibility of investigating the use of activity-based models (ABM) for reflection of such effects in appraisal.
- e. Batley et al. (2020) discussed UK's experiences in treatment of the worthwhile use of travel time in the context of the UK 2014/15 national value of time study. They demonstrated novel attempts at incorporating the notion of travel time use in VTT estimation, observing and discussing somewhat counterintuitive findings, e.g. VTT on arguably most productive modes, such as long-distance rail, being higher than expected.
- f. de Jong and Kouwenhoven (2020b) presented experiences from the Dutch 2009/2011 national value of time study (see **Section 5.9** for a more detailed discussion of the specific results). In addition, they observed that whilst relatively slower growth (compared to income) in VTT may be expected, it is unlikely to end up 'zero VTT', for example as not all travel time is spent working (and unlikely to ever be). Moreover, adjustments to modelling methodologies
- g. Molin et al. (2020) presented initial findings concerning the SP study based in the context of rail, looking at measuring the value of being able to participate in an activity while travelling (see **Section 5.9** for a more detailed discussion of the results).

The workshop pointed out clearly that people continue to value travel time reductions, since recovery of travel time permits undertaking full range of activities which may not necessarily be possible during travel. Indeed, the proliferation in mobile ICTs and their increasing capabilities expand on the range of (digital) activities that can be done while travelling, such set of activities is not as complete as what can be done at the origin or destination. All in all, the conclusions supported the case for continued use of VTT due to reduced travel time in investment appraisal, but also acknowledge and emphasised the need for an increased reflection of the need to make travel time more comfortable and productive, so that quality of time spent travelling is improved (ITF, 2019).

Pudāne (2020) proposed a model, in the context of autonomous vehicles' use for commuting trips, in which travel time activities are related to the choice of departure time, congestion level as well as road capacity. The proposed framework implies that in absence of congestion, depending on whether the on-board environment supports more home or work activities (which is driven by people's preferences), the departure time would shift towards earlier or later times respectively. In the presence of congestion, the preferences for undertaking home or work activities will have may lead to shifting of congestion patterns towards earlier or later times respectively. The overall conclusion

stemming from the research was that travel time activities undertaken on-board AVs may lead to worsened congestion unless AVs capabilities lead to an increased road (bottleneck) capacity.

Pawlak et al. (2020) proposed a theoretical framework, based on the microeconomic time use theory, in which they relate consumption and time allocation decisions (time allocated to activities at home, out of home and while travelling) to energy consumption and costs. Their motivation stemmed from accounting for variable energy (electricity) pricing that can, in turn, affect decisions concerning travel and charging times by electric vehicle users. Whilst VTT would not be computed directly in the paper, the framework points out towards a possibility where VTT is jointly determined by the possibility of undertaking activities during travel as well as by variability of the electricity costs across the day, in turn affecting broader travel and activity scheduling decisions.

Calastri et al. (2022) proposed a means through which the capability of observing duration of time spent working online but the travellers can provide a means of obtaining insight into overall travel time allocated to work activities, which coincides with 'p' term in the Hensher's equation. The authors argued that one way such information could be gathered passively would be via interrogating patterns in networking data flowing via on-board networking infrastructure. Arguably, such patterns could be established on aggregate basis (i.e. without infringing people's privacy) by analysing patterns in encrypted traffic, though such approaches are largely experimental.

Cornet et al. (2022) noted that existing transport assessment tools rarely factor in the perspective of travellers, i.e. concerning their subjective travel experience while on the move. They proposed that in addition to the economic utility of the activity at the destination of the trip and the work-related productivity during the trip, a more complete perspective should reflect the idea that travel can be pleasant, meaningful, or worthwhile. In their proposed conceptual framework, experience factors (such as safety, comfort, noise, congestion) carry particular importance, can be assessed in terms of satisfaction levels, and are enabled (or not) by the availability of facilities and equipment. These factors affect travel activities (including the act of travelling itself), which is also related to wanted and unwanted efforts. Lastly, the combined experience results in the perceived value and worthwhileness, assessed in terms of enjoyment, productivity, and fitness. Based in data from 8 European countries, the authors demonstrated how such values vary across the modes of walking, public transport (short and long distance), private motorised and cycling and micromobility.

Abeille et al. (2024) proposed a theoretical model which looked more into the behavioural concerning motivations for undertaking activities while travelling. They proposed that activities undertaken during travel do not (always) happen ad-hoc but are an outcome of a decisions driven by ambitions and anticipation against the actual conditions. Their framework also links these aspects to the subsequent perception of productivity, which the authors argue is resulting from reflection on the performance against the initial expectations. In the empirical framework making use of SEM, that looked at dedicated survey data collected amongst train passengers in the UK, demonstrating feasibility of collecting a comprehensive data on travel time behaviour using an online survey. They also showed effects of somewhat less explored factors, such as Covid-19 measures (facemasks) privacy perception and delays on the experienced productivity and well-being.

Jara-Diaz (2024) extended the time allocation model theory to allow for presence of travel time activities in the microeconomic formulation of utility-maximising consumer (traveller). His incorporation of travel time activities in an additional time constraint follows that from, for example, Pawlak et al. (2020) and accounts for the productivity differences not by relying on the notion of relative productivity term but by introducing a travel-time specific wage rate, i.e. essentially a mode-specific monetary reward for a unit of work. In terms of the relation to value of travel time, he

concluded that the impact from travel time activities is such that *‘improving the conditions to perform either leisure or work reduces the VTT’* (Jara-Diaz, 2024 p.4). Interestingly, the proposed derivation indicated that VTAT (or VTT) could be obtained by investigating the differences in marginal utilities of leisure or work while travelling relative to leisure or work in non-travel contexts.

Santos de Sa et al. (2024) proposed a conceptual model that formalises the ideas of Mokhtarian (2018), looking into the idea of activity transferability and factors driving feasibility and preferences for it. They distinguished between undertaking activities while travelling as a means of reducing the boredom (‘killing the time’) and a means of replacing (or ‘transferring’) activity from time elsewhere during the day. Their modelling focused on understanding what drives whether a specific activity undertaken during travel is a manifestation of the former or the latter behaviour. Their empirical results have provided evidence for the link between preference for faster journeys and lower transferability, which is consistent with the postulated theoretical results in other studies.

Overall, the contributions highlight maturation of the field of modelling travel time use and its role within the broader activity-travel behaviour fabric. The recent decade has seen a progression from anecdotal evidence, implicit recognition of the possibility of such phenomenon (e.g. deflated cost coefficients, unexplainable preferences for modes of transports allowing better use of travel time), towards comprehensive conceptual frameworks, extensions of microeconomic paradigms and explicit links towards investment appraisal, including value of time. Whilst the following section outlines some of the empirical attempts seeking to elicit impacts of travel time use on VTT, the account of the studies certainly emphasises gradual requirement to arrive at consensus: nomenclatural (‘travel time use’, ‘travel-based multitasking’, etc.), methodological (e.g. activity categorisation, agreed theoretical foundations) and empirical, e.g. magnitude of the phenomenon or relative importance of trip-, traveller- and technology-related covariates. Arguably, the studies have not yet explicitly embraced how travel time use and its value may be affected by multi-modal nature of the trip, what is the role of travel time use during disrupted trips and uncertain travel conditions more broadly (and how travel time use may evolve under such circumstances) or whether and how it is related to the wider lifestyle choices, such as flexible work arrangements.

### 5.9 Estimating the effects of travel time use on the VTT

In addition to theoretical developments or empirical studies looking at models of behaviour while travelling, several studies attempted quantifying the effects and magnitude of reflecting travel time use in the context of the VTT. Broadly speaking, the studies can be divided into ones relying on revealed preference (RP) and stated preference (SP) data.

The first group of RP approaches rely on estimating parameters of the Hensher’s Equation (or a subset of them), which could be used subsequently to apply appropriate discounting to the standard VTT, or more precisely, VTT under absence of travel time activities. Several such studies have been listed by Wardman and Lyons (2016), as per **Figure 17**. Pawlak et al. (2017) proposed a model in which they modelled duration of activities using a hazard-based approach (allowing modelling of multiple spells of episode), whilst productivity was modelled as log-normally distributed random variable. They have also demonstrated, how such a framework can be used to simulate allocations to work and non-work activities as well as the associated productivity, which essentially provide a means of computing ‘p’ and ‘q’ parameters in the Hensher Equation.

In another set of contributions, Calastri et al. (2022) and Palma et al. (2023) attempted to compute allocations to various activities conducted while travelling using multiple discrete-continuous models (Multiple Discrete-Continuous Extreme Value – MDCEV, and extended Multiple Discrete-Continuous – eMDC). Their results have largely proven the feasibility of employing such methods, confirming the

impacts of the recognised covariates, such as mode of travel, trip purpose or duration, but also paving the way towards employing such methods to elicit impacts of various policy levers as well as simulate allocations of travel time across different activity types (and hence enabling simulation of the Hensher Equation's 'p').

Varghese and Jana (2018) was the first study that explored the issue of travel time use impacts in an emerging economy (India), using revealed preference data. In particular, they collected travel diary data for 1,123 individuals across the city of Mumbai and reconstructed alternatives for the reported, though it was not clear how multitasking behaviours and opportunities were reconstructed for the unchosen modes. They observed that the value of travel time reduced by 26% for individuals who performed multitasking. As for the specific activities, reading on a mobile device, usage of social media, messaging or talking to someone on phone, and gaming, led to reduction of the value of travel time by 25%, 37%, and 16% respectively.

Malokin et al. (2021) used RP data covering a 2011-12 sample of 2,216 North California commuters to investigate how VTT and willingness to pay differed in 'millennial' (people born in 1980s and 1990s) and 'non-millennial' cohorts when taking into account travel time activities across a number of commuting modes (bicycle, commuter rail, transit (including local bus, express bus, light rail, and metro rail), shared ride, and drive alone). The underpinning hypothesis has concerned how the millennial cohort, associated with being more 'digitally native', i.e. accustomed to using digital technologies, may experience travel time when equipped with ICT that can enable conduct of activities whilst travelling. Compared to non-millennials, the mode choice of millennials is found to be less affected by socio-economic characteristics and more strongly influenced by the activities performed while traveling. Young adults are found to have lower VTT than older adults for both in-vehicle (15.0% less) and out-of-vehicle travel time (15.7% less), and display higher willingness to pay (in time or money) to use a laptop (a proxy for more worthwhile use of travel time), even after controlling for demographic traits, personal attitudes, and the propensity to multitask.

Hartwig et al. (2024) looked at RP data collected (more than 13,178 trips made by 851 individuals) in a week-long mobility survey between September 2019 and August 2020 in Austria. They relied on a standard travel survey data collection protocol, though allowing for reporting of the secondary activities while travelling to be reported in an open text field (with confirmation required if none reported) and subsequently recoded according to Harmonised European Time Use Survey methodology (Eurostat, 2019). Whilst the study provided a cross-modal insights, it relied heavily on imputation of travel time use for non-chosen modes, itself making use of rather high-level imputation techniques. As for the effects of travel time use on the value of travel time, the authors observed that whilst the value of travel time was lower for public transport than for car drivers even without multitasking (5.50 vs. 14.32 €/h respectively), it would further drop more substantially with different levels of multitasking (3.95 and 3.35 €/h for level 1 and 2) compared to car drivers (only insignificantly to 14.05 €/h).

Complementing the RP approaches, another set of studies relied on SP approaches, in which scenarios describing conditions or feasibility of travel time activities are proposed against varying journey cost or duration. This permits direct estimation of the effect of travel time activities on cost or duration sensitivity, and thus the value of travel time.

In a report commissioned by the UK Department of Transport, SDG (2016) made use of SP on the sample of ca. 2,000 UK to estimate willingness to pay for improved voice connectivity, different levels of Internet service provision (effectively speed or bandwidth) and its reliability. The focus was on the rail context, which on the one hand has been seen as one of the modes with high potential for

productive experience but on the other hand suffers from poor connectivity. The study revealed that respondents were willing to pay up to a 29% higher fare to move from zero to perfect voice connectivity, and up to 26% more for reliable, high bandwidth data connectivity. The authors also observed heterogeneity in terms of WTP for improved voice and data connectivity depending on trip purpose (commute, business, leisure) and distance (short, medium, long), though without a consistent clear pattern as to which segment values such improvement most (especially when comparing business travellers and commuters). The report also demonstrated that if provided with better connectivity (higher bandwidth and reliability), two-thirds of respondents would increase their participation in digital activities, either by consuming more data or by partaking in additional activities. Consequently, it is possible to claim that provision of improved connectivity and therefore enablement of activities, if not charged for, would act in a similar manner to a subsidy. This would allow the users to accept higher monetary cost for a given journey duration or, equivalently under 'trade-off', longer journey duration for a given travel cost. This provides indirect evidence for the role of travel time activities on value of time, though naturally holding for users whose intended time use requires access to connectivity.

Kouwenhoven and de Jong (2018) relied on SP data components contained in the Dutch Value of Travel Time survey data from 2011. They observed that travellers who indicated that a shorter trip duration was useful, or a longer trip duration was very inconvenient, had a higher value of travel time. Respondents who indicated that a shorter travel time was useful to them had on average a  $22.1\% \pm 3.7\%$  higher VTT compared to others. Moreover, looking across modes to explore the hypothesis that value of travel time decreases when the travel time itself can be spent in a useful manner, they observed respondents on local public transport (bus, tram, metro) who were able to spend their travel time usefully having on average respectively a  $20.0\% \pm 5.5\%$  and a  $20.9\% \pm 5.6\%$  lower VTT than other respondents. At the same time, for car users, no significant effects could be observed concerning ability to spend their travel time in a useful way on the value of travel time. In retrospect, the authors indicated that the survey should have included questions on what respondents would have been doing without a specific device they had at hand.

Bounie et al. (2019) presented an approach to valuing connectivity to mobile phone and Internet networks in the context of public transport, based on SP of 501 inhabitants of the Paris region. The respondents were provided with hypothetical scenarios of trade-offs between time reductions and improved connectivity (compared to previous studies which looked at changing fare level). The analysis looked at combinations for which utilities would equalise, establishing sensitivities to changes in the travel duration and connectivity quality. This has yielded an average 12% reduction in the value of travel time due to improved connectivity. They also noted that the values ascribed to better connectivity (and thus reduction in value of time caused by an improved connectivity) are higher when users perform various tasks with smartphones or tablets during their trips and when they experience many connectivity problems in the reference situation. In other words, the effects of travel time activities on value of time are most pronounced, arguably, if the current travel conditions differ much from ideal conditions for travel time use. Interestingly, the authors made use of time multipliers reflecting the marginal rate of substitution between "as currently" and "optimally" connected travel times (i.e. the ratio between marginal disutilities of time), which the authors estimated empirically for a variety of socioeconomic attributes, time use profiles and travel conditions.

Molin et al. (2020) proposed to estimate the impact of travel time activities on value of travel by constructing a within-person choice experiment, which involves that the same respondents making choices in a context that enables conducting activities, as well as in a context that does not enable conducting activities. Their study used data from 820 train travellers in the Netherlands and referred

to the context where the respondent intends to conduct an activity requiring a specific piece of equipment, e.g. reading required a book. The study compared sensitivity to travel and cost parameters for the same respondent travelling with and without the necessary piece of equipment (and hence being and not being able to engage in the intended activity). Their results demonstrated that VTT in the activity context is significantly lower than the VTT in the non-activity context. Reduction in VTT due to conducting onboard activities was around 30% for commuters, while for leisure travellers this could amount to about 50% (**Table 21**). Importantly, they observed that the difference depended strongly on the type of activity which was referred to in the experiment. For example, enabling work would lead to about 34% VTT reduction compared to 31% for reading, whilst for leisure travellers the respective changes would be 15% and 47% respectively. What follows from this result is that reduction in the VTT appears to be inherently driven by a combination of activity and journey purpose, which confirms the earlier findings that preference for specific activities is affected by journey purpose.

*Table 21: Value of time reductions (in Euros) per purpose-activity group (in Euros per hour), derived from the context with activity (AC) and no activity (NAC). Source: Molin et al. (2020)*

|                                                                | Commuters |        |        | Leisure travellers |        |        |
|----------------------------------------------------------------|-----------|--------|--------|--------------------|--------|--------|
|                                                                | Work      | Read   | Music  | Work               | Read   | Music  |
| <b>VTT<sub>AC</sub></b>                                        | 12.42     | 11.22  | 10.26  | 6.54               | 3.77   | 5.74   |
| <b>ΔVTT</b>                                                    | 6.36      | 4.98   | 3.63   | 1.16               | 3.39   | 0.69   |
| <b>VTT<sub>NAC</sub></b>                                       | 18.78     | 16.20  | 13.89  | 7.70               | 7.16   | 6.43   |
| <b>% VTT reduction<sup>1</sup></b>                             | -33.9%    | -30.7% | -26.1% | -15.1%             | -47.1% | -10.7% |
| <sup>1</sup> % VTT reduction = ΔVTT/VTT <sub>NAC</sub> * 100%. |           |        |        |                    |        |        |

Wardman et al. (2020) relied on analysis of stated preference data collected in 2014. They proposed 3 types of SP, which they called: ‘standard’, ‘strategic’ and ‘tactical’. The ‘standard’ SP emulated conventional SPs used in eliciting values of travel time. The ‘strategic’ SP introduced possible restrictions on the use of technology, ranging from complete ability to do whatever the respondent intended, through inability to use mobile technology, to inability to undertake any activities. The ‘tactical’ SP reflected choice of more elaborate time use profiles, where some of the activities could be restricted in certain alternatives and time reallocated to other activities. Overall, the authors observed about 15.5% (11.7% when rescaled due to sample composition and distribution of trip distances and purposes) of reduction in the value of travel time when comparing ‘no digital devices’ to ‘as is’, though noting the 2014 empirical context and the impact of composition and importance of digital activities on the overall average. The authors observed that when assuming the digital revolution to have started in 2000, the obtained results figures imply 1.20%, 0.88% and 0.73% annual reductions in VTT respectively. Taking 2005 as a starting point yields corresponding figures of 1.87%, 1.37% and 1.14%. Consequently, such annual reductions would be consistent with downrating value of travel time increases attributed to income growth, which is consistent with the observed studies that demonstrate much less than the expected growth in VTT.

#### 5.10 VTT impacts of travel activities in autonomous vehicles

There have been several attempts to elicit impact of travel time activities and worthwhile travel time use on the value of time in the context of Autonomous Vehicles (AVs). Indeed, the ‘Technology’ scenario within DfT’s Common Analytical Scenario’s collates some relevant evidence to arrive at a recommendation for a 21% downrating of the VTT (Department for Transport, 2023e) (DfT, 2023f).

Kato et al. (2024) discussed value of travel time savings for leisure trips made using autonomous vehicles in the context of Tokyo, based on data from a stated choice survey. VTT in AVs was found to be lower than that in non-AVs by approximately 7.3 % on average, though this result was not statistically significant. Preference for in-vehicle multitasking, age and gender were found to not affect

the reduction of VTT in AVs though the VTT varies significantly across these subgroups. The finding that VTT reduction in AVs is not statistically significant agrees with other findings in the literature for leisure trips (see Kolarova et al., 2019). Most notably, Kato et al. (2024) suggest that there may be differences in VTT between a non-AV driver, a non-AV passenger, and drive-free AV occupants; in particular, guidelines in Japan already acknowledge that non-AV passengers have a lower VTT compared to non-AV drivers by about 13.6%.

Choi et al. (2023) estimated MNL logit mode choice models stratified by trip length, of the choice between AV with a manual driving option (AVMD), AV with a self-driving option (AVSD), shared AV (SAV), and public transit (PT), using an SP survey administered to Korean commuters in 2019. The results show that the value of in-vehicle travel time savings for AVSD (on average, \$7.61/hr) is lower than for the other two AV modes (AVMD = \$10.26/hr, SAV = \$13.67/hr). Moreover, as travel distance increases, VTT for shared modes (SAV and PT) tends to decrease.

Kolarova et al. (2019) similarly used stated preference survey data from Germany and found no changes in VTT for leisure or shopping trips using AVs, though they found a VTT reduction of 41% for commuting trips for autonomous driving as compared to driving a conventional car. This was based on the analysis of data from an online SP survey conducted in Germany in 2017. The same data was further analysed to reveal that a reduction of VTT by 22% is found for riding autonomously compared to driving manually for long-distance trips (Kolarova and Steck, 2020).

Malokin et al. (2019) studied commuters in Northern California using revealed preference data collected along the Sacramento-San Francisco Bay Area transportation corridor in 2011-2012. For each combination of mode and travel activity, the authors formulated a binary logit model to understand propensity to multitask during travel. In addition, a multinomial logit commute mode choice model was estimated which included a propensity for productive travel multitasking. The MNL is then implemented in AV diffusion (full adoption) scenarios, for example, by setting the 'propensity to use a laptop' for shared ride and driving alone to be equal to the corresponding propensity for commuter rail (if greater). The results show an increase in the drive alone market share at the expense of all the other modes.

Research by Fu et al. (2019) incorporates the possibility to multitask (owing to automation of the taxi and flying taxi). A stated preference questionnaire was designed and distributed in Munich, and multinomial/mixed logit models were estimated of the choice between private car, public transport, autonomous taxi (AT) and autonomous flying taxi (AFT). Multitasking possibility was set as a constant attribute in the stated choice experiments, with a value of 'yes' for AT and AFT, 'no' for private car, and 'in part' for public transport. In this manner, the VTT estimates derived from the analysis take into account the potential to multitask with the autonomous modes -- 27.55 €/hour for car, 27.47 €/hour for PT, 32.57 €/hour for AT and 44.68 €/hour for AFT). Nonetheless, these results need to be treated with caution (especially direct empirical values) given the possibility of hypothetical bias due to the experimental nature of AT and AFT modes. It is notable that the models estimated found no significant impact of walking and waiting time for AT and AFT, suggesting that the respondents may have been influenced by the novelty of the new modes. In addition, the VTT estimated for the car and PT modes are also higher than average values of time from German national studies. It is also difficult to directly infer the effect of travel time use from these results (due to the design of the SP experiment, with multi-tasking being a categorical attribute).

Pagoni et al. (2022) present a meta-analysis of the value of time for autonomous driving based on 13 studies published after 2018. It is found that the study characteristics (e.g., country, area type where the autonomous vehicle is travelling, trip purpose), the AV characteristics (e.g., level of autonomy,

whether the AV is private or not), and model specification characteristics affect the value of time estimations.

Empirically, there is a stream of work that investigates the value of not sharing a trip with anyone (potentially indicating the importance of privacy/intimacy) as part of a shared AV service, using SP surveys. **Table 22** below summarises some of the findings in terms of Willingness to Share (WTS) metrics. The metrics are broadly similar across the study areas (US, Dallas in the US, Netherlands, and Southern Ontario in Canada), with higher values on WTS emerging when there are additional stops required to pick up a passenger, which would of course confound the willingness to share with a willingness to increase the length of the trip.

*Table 22: Comparison of VTT studies for shared AV services*

| Country           | USA                                                   |                                                                                | Netherlands                                                                                                        | Canada                                                           |
|-------------------|-------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------|
| Region            |                                                       | Dallas                                                                         |                                                                                                                    | Southern Ontario                                                 |
| Definition of WTS | avoid an additional stop to pick up another passenger | money value attributed to traveling alone, not to have an additional passenger | monetise how much money individuals are willing to pay to (not) share their trip with other individuals=disutility | the valuation of sharing an on-demand car with another passenger |
| Value             | <b>£1.05/trip - £5.44/trip</b>                        | Work trip= <b>£0.39/trip</b><br>Leisure trip= <b>£0.71/trip</b>                | <b>0.44 £/trip</b> (1 or 2 additional passengers)<br><b>2.43 £/h</b> (shared with four additional passengers)      | <b>£0.59/trip - £2.36/trip</b>                                   |
| Source            | Sheldon and Dua (2024)                                | Lavieri and Bhat (2019)                                                        | Alonso-González et al. (2021)                                                                                      | Sweet (2021)                                                     |

Nordström and Engholm (2021) carry out a morphological analysis to illustrate the complexity of determining the value of travel time for a diverse range of mobility concepts. They suggest that future transport research should move away from specifying VTT for each existing and potential travel mode. VTT models should instead employ a range of travel and service characteristics to reflect the heterogeneity of VTT in the context of rapid technological advancements. The paper considers four categories of parameters that directly or indirectly impact the value of travel time: (i) vehicle characteristics, (ii) operating principles, (iii) journey characteristics and (iv) traveller characteristics. The parameters and respective attributes result in a morphological matrix that spans all possible solutions. Five plausible mobility concepts are discussed in this context, as an example, and **Table 23** summarises the conclusions drawn from the morphological analysis for each of the mobility concepts.



Table 23: AV concepts proposed by Nordström and Engholm (2021)

|                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|----------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Feeder shuttle</b>                                    | <i>From the perspective of reliability and variability, the VTT should be lower than for a regular public transport bus. However, the qualitative conditions of a feeder shuttle will generally be lower than on a regular bus making it impossible to spend the travel time productively by e.g. performing work. Based on the analysis, we [the authors] do not believe it is warranted to assume the VTT of a feeder shuttle to be lower than for a regular bus nor it warranted to assume it to be higher without further studies</i> |
| <b>Public transport shuttle on-demand in rural areas</b> | <i>The vehicle is assumed to be shared but given the rural context, it is probably rarely crowded. While the characteristics of shared travel will impact the sense of privacy and potentially comfort, the extent of the impact should be evaluated further.</i>                                                                                                                                                                                                                                                                         |
| <b>Private premium self-driving car</b>                  | <i>Compared to a non-automated vehicle, it has been assumed that the in-vehicle time in a private premium self-driving car can be utilized better by the traveller and thus the VTT is lower compared to the current mode 'car'.</i>                                                                                                                                                                                                                                                                                                      |
| <b>Shared self-driving taxi</b>                          | <i>In summary, while the VTT can be lower compared to traditional car travelled, the lack of privacy and safety concerns as well as higher levels of variability make it difficult to assume that the VTT can be compared to other, non-shared modes without further investigations.</i>                                                                                                                                                                                                                                                  |
| <b>No-frills self-driving service</b>                    | <i>A high rejection rate heavily impacts reliability of the service in general as well as the VTT directly since the availability of the service cannot be trusted. Moreover, the concept is assumed to be 'departure' based and does not allow for travel to be planed according to estimated arrival time. Additionally, the in-vehicle level of comfort is presumed to be low. All in all, even though the service is self-driving, the VTT is comparable to low-cost bus travel, at best.</i>                                         |

A study by Correia et al. (2019) extended the microeconomic model by Jara-Díaz (2007) to conclude that the VTT for work travel should be reduced in the context of autonomous vehicles, whereas the VTT for leisure travel should remain unchanged. The authors then go on to conduct a stated choice experiment for the choice between using a conventional car, an AV with 'office interior' and an AV with 'leisure interior' for work and leisure travel; the interiors are well-described with suitable images. A second set of stated choice experiments use chauffeur-driven rather than autonomous vehicles as alternatives with all else being the same. Analysis of the data generated from a representative sample of the Dutch population shows that mean VTT for AV-office travellers is in general lower compared to conventional car users and AV-leisure users. On the other hand, the mean travel time valuation of AV-leisure users is slightly higher than the VTT of conventional car users, however, this is not statistically significant. Unlike the more qualitative analysis by Nordström and Engholm (2021), Correia et al. (2019) acknowledge that their quantitative analysis does not include several factors that should be considered, such as travel time reliability and the details of the types of activities that are possible to undertake in the AV.

Overall, there has been a decent amount of empirical work concerning autonomous modes of transport, with either implicit or explicit treatment of the idea of travel time use when travelling by such modes. There is reasonably consistent evidence that the on-board environment allowing usable activities is associated with reduced VTT for work-related trips though not so much for leisure trips.

At the same time, what constitutes an autonomous vehicle is subject to potential (mis)interpretation by the respondents and warrants careful definition as well as sensitivity analyses, to date largely absent from the research. Moreover, there is potential for conflating VTT reductions due to productive travel time use and the apparently higher VTT in emerging modes such as flying taxis and autonomous flying taxis.

Furthermore, due to the nature of autonomous vehicles, there is very limited knowledge and confidence concerning how travel time use on-board such modes may be affected by factors such as

traffic conditions (and therefore vehicle kinematics), road conditions or competition from other activities (such as rest or leisure).

### 5.11 Novel aspects concerning measurement of travel time activities

Yin and Cherchi (2024) used virtual reality (VR) experiments to test the ability of such technology to enable elicitation of preferences for stated preference surveys in the context of autonomous taxis (in comparison to conventional taxis with drivers). They compared results from a survey administered in a conventional way (via an online partner) and through an immersive VR technology. Whilst the study did not directly look at travel time use, they provided insights of relevance in the context of considering use of such approach for travel time use studies. They observed that such technology can have an impact on the preferences elicited with stated choice experiments, as compared to traditional surveys, in three ways. Firstly, it can impact perception of attributes that the respondents are able to experience (also partially) within the VR environment, e.g. observing people waiting for taxis influencing sensitivity to waiting time. Secondly, the impact was seen with respect to those attributes that are typically more difficult to measure, either because of their visual nature (such as urban environment) or because their value results from evaluation of impact exerted by action of others (such as social conformity attributes). Thirdly, they observed that VR experiments were more capable of eliciting systematic heterogeneity in the individual preferences for the alternatives evaluated and their characteristics. From the point of view of travel time activities, the authors noted that they had not been captured in the design of the experiment. Nonetheless, considering the above findings, part of which are inevitably linked to the motivations for travel time activities, one could expect VR-based SC to lead to different conclusions concerning their impact on VTT.

Abeille et al. (2022) looked at the issue of measuring productivity during travel time activities, by performing semi-structured interviews with knowledge workers from an American firm to investigate in-depth the motivation and practices of mobile work, use of ICT and perception of productivity with use of ICT. Their analysis led to suggestions that individuals may find it easier to identify factors hampering productivity, in what they term the 'inverse' approach, i.e. understanding '*un*productivity'. The authors observed that that individuals experienced difficulties speaking about productivity and productive tasks while finding it easier to discuss what made them *un*productive. In addition, they reported on the requirement to account for people's confounding of experience of process productivity (i.e. experiencing environment which does not prevent effective activity undertaking) and output productivity (i.e. being able to complete a specific task). In particular, the act of completing a pre-planned task could lead to perception of being productive even if the person had not been efficient. This naturally introduces the aspects of planning and expectations, which appear to have very strong impact on people's subjective perception of how productive the travel context may be.

Abeille et al. (2024) reflected on the shortcomings concerning a singular self-reported metric of productivity. They proposed to treat productivity as a latent (unobserved directly) construct, which can only be measured through a set of indicators. Such an approach allows flexible incorporation of various observable metrics, self-reported and exogenously measured. In their operational structural equation model (SEM), using 2023 UK rail passenger data, they used indicators concerning 'positive indicators on work quality': satisfaction with quality of work and feeling productive; 'positive indicators on work output': satisfaction with number of tasks completed and 'negative indicators' of work activity: impacts of disturbances of the surrounding environment. This approach reflected findings from their earlier, in-depth interview study. Arguably, this approach can provide more robustness against self-reporting biases, confounding of process and outcome productivity or reporting on 'positively impacting' vs. 'negatively impacting' factors. Lastly, the study pointed out the issue of productivity

possibly varying at different stages of travel (beginning, middle, end), due to variation in travel conditions, e.g. crowding, as well as travellers' cognitive abilities, e.g. fatigue.

### 5.12 Key takeaways on worthwhile use of travel time

The phenomenon of travel time use has moved, over the past two decades, from a niche research topic to becoming one of the key, challenging considerations concerning value of travel time considerations. This has been reflected in emergence of several complementary theoretical and conceptual frameworks, accompanied by dozens of empirical studies across various geographical, modal and purpose contexts. The growth in studies has, however, led to some confusion in terminology, e.g. travel time activities, activities on-the-go, activities during travel or travel-based multitasking. Furthermore, no clear consensus has emerged as to how travel time activities should be presented in travel behaviour surveys.

Theoretical framing of the problem has matured substantially with several frameworks linking different considerations behind travel time use, its experience as well as value of travel time. Importantly, the frameworks are not contradictory to each other, nor to the Hensher Equation approach. In particular, the proposed frameworks attempt to explain better where different terms of the Hensher Equation may arise from (behaviourally) or result from relaxing implicit assumptions in the existing conceptual and economic frameworks underpinning travel behaviour models and the concept of value of travel time.

The dominant perspective tends to be that of the traveller (or employee in the context of business travel), with employer perspectives virtually absent from the mainstream transport research literature. A study looking into how the valuation of travel time use value may differ depending on the employee vs. employer perspectives is currently missing, especially for complex circumstances such as when the purpose of travel may be mixed, costs shared and benefits from activities during travel accrued to both employer and the employee.

The empirical work has relied on either RP or SP data, with more recent contributions leaning towards the latter given the direct control over multitasking conditions and scenarios, otherwise difficult to know for the unchosen modes. However, there is lack of consensus concerning how to represent travel time use in SP studies, e.g. specific activities, their durations, experienced quality (including productivity).

As for the existing RP studies, they either look at models of travel time use (which can provide inputs to frameworks such as Hensher Equation) or discrete choice modelling to elicit values of travel times directly. In the latter case, the issue of multitasking for unchosen alternatives has been addressed via imputation approaches, though these are relatively ad-hoc and do not typically provide a means of identifying additional imputation errors.

Although several RP studies have effectively measured the impacts of travel time use on the VTT, no combined SP-RP approach has been proposed to the best of the authors' knowledge. Nor has there been much of a discussion what the use of one versus the other approach implies in terms of findings, strengths, or weaknesses.

One of the specific methodological shortcomings concerned looking at how to effectively elicit time allocation to various activities whilst travelling. While studies customarily ask types of activities, only a handful of studies looked at durations and scheduling. A notable example of the SPURT survey (Fickling et al., 2009) used a simple method of asking the respondents to mark each of 10 boxes (representing the entire journey split into 10 equal episode) spent on work activities (**Figure 18**). A purely methodological study looking at whether such method could be expanded to incorporate more

activity type and more granular representation (e.g. 20 boxes) would be a valuable indication as to whether such approach could be a useful and simple way to elicit activity durations and scheduling in SP studies. Importantly, as Pawlak et al. (2017) noted, the use of such method implies heteroscedasticity in activity duration, which needs to be accounted for in the subsequent modelling.

Q31. Can you please indicate when during **this** rail journey you undertook work-related activities.

**PLEASE DRAW A LINE OR LINES IN THE GRID BELOW**

**THE EXAMPLE BELOW INDICATES WORK DONE IN FIRST 30% OF TRIP, AND AGAIN BETWEEN 60% AND 70% OF TRIP**

The figure displays two horizontal grids used for activity allocation reporting. The top grid is an example with lines drawn in the first three boxes and the sixth and seventh boxes. The bottom grid is empty for completion. Both grids are labeled 'Start', 'Middle', and 'End' at the bottom.

Figure 18: Activity allocation reporting proposed in the SPURT study by Fickling et al. (2009)

Amongst the emerging approaches, virtual reality methods have been explored for an improved presentation of the choice alternatives. When compared against standard SP presentation, systematic differences have been observed. However, such approaches remain experimental and their use in the context of travel time use modelling has not been validated. Interestingly, there has been a rather surprising absence of studies looking at the potential to use activity-based models (ABMs) to explore implications of travel time use on policy considerations.

There has been the broader issue concerning lack of Monte Carlo experiments concerning sensitivity of the modelling outputs to misrepresentation of multitasking. This effectively means that there is no 'right' approach in terms of proven reliability or robustness to misspecification. Palma et al. (2023) attempted to test various time allocation specifications and multitasking components, though in a specific empirical context only.

There is very clear evidence for the role of journey purpose in how travel time activities impact upon the value of travel time. Specifically, the purpose determines the nature of activities that would have been ideally undertaken, which in turn drives possible impacts of travel time use on value of travel time if such activities are restricted wholly or partially (quality, productivity) by travel conditions, and the absence of facilities or equipment. Nonetheless, studies have not looked at multi-purpose journeys, nor whether preferred segmentation requires only purpose as one of the covariates or whether the entire models or decision rules must be segment specific.

In light of the impacts of travel time use varying by activity types, travellers whose activities are easier to undertake whilst travelling are likely to experience lower values of travel time. This naturally leads to equity considerations, such as the extent to which investments in travel improvement vs. reduction in travel duration will accrue across the population. Whilst almost universal proliferation of mobile devices may be seen as mitigating such equity issues, this has been neither confirmed nor refuted in research to date.

Factors which have been identified in empirical studies as key considerations in shaping the extent of travel time use and impacts on VTT include:

- Mode
- Duration (of travel)
- Purpose (business vs. leisure)
- Whether activities undertaken during travel are at least partially substitutes for activities that would have been undertaken, had there been no travel need.

- Pre-planning of travel time use, including expected vs. encountered conditions (as preparation matters).
- Feasibility of undertaking the intended activities as opposed to the actual activities, including availability of facilities (esp. connectivity), travel conditions (esp. crowding) and equipment.
- Measurement of the quality of activities, including productivity, but also enjoyment, fitness aspects, especially for non-business related trips.
- Socioeconomic attributes of the traveller as well as companions.

Table 24: Summary of studies concerning measurement of impacts of travel time use on VTT

| Study                          | Context                                   | Approach – primary models       | Key findings in relation to VTT                                                                                                                                             |
|--------------------------------|-------------------------------------------|---------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Pawlak et al. (2017)           | 2008, rail, UK                            | RP – hazard, log-normal regress | Joint model for activity duration and productivity. Inputs to HE                                                                                                            |
| Calastri et al. (2022)         | 2016, rail, UK                            | RP – MDCEV model                | Allocation of travel time to activities. Inputs to HE.                                                                                                                      |
| Palma et al. (2023)            | 2014-15, multiple modes, UK               | RP – MDCEV, eMDC                | Allocation of travel time to activities. Inputs to HE.                                                                                                                      |
| Varghese and Jana (2018)       | 2016, multiple modes, India               | RP – DCM                        | Reduction in VTT by 15-25% due to travel time use. Variation by mode and activity.                                                                                          |
| Malokin et al. (2021)          | 2011-12, multiple modes, USA (California) | RP – DCM                        | Young adults ('millennials') have higher willingness to pay for using laptop and lower in- and out-of-vehicle time by ca. 15%.                                              |
| Hartwig et al. (2024)          | 2019-20, multiple modes, Austria          | RP – DCM                        | Reduction in VTT by ca. 30-40% for public transport users who can use travel time, compared to no reduction for car users (no reduction even with assumed travel time use). |
| SDG (2016)                     | 2015-16, rail, UK                         | SP – DCM                        | Reduction in VTT by ca. 20-30% due to improved voice and data connectivity enabling activities while travelling.                                                            |
| Kouwenhoven and de Jong (2018) | 2011, multiple modes, the Netherlands     | SP – DCM                        | Reduction in VTT by ca. 20% due to being able to spend travel productively. Variations by mode.                                                                             |
| Bounie et al. (2019)           | 2015, public transport, France (Paris)    | SP – DCM, contingent valuation  | Reduction in VTT by ca. 12% due to allowing better connectivity and the associated activities.                                                                              |
| Molin et al. (2020)            | year not reported, rail, the Netherlands  | SP – DCM                        | Reduction in VTT by ca. 30% for commuters and up to 50% by leisure travellers. Variations by purpose and activity type.                                                     |
| Wardman et al. (2020)          | 2014, rail, UK                            | SP - DCM                        | Reduction in VTT by ca. 10-15% due to allowing travel time use.                                                                                                             |

The role of travel time use and its possible adaptation in the context of uncertain travel conditions, especially during disruptions, has been virtually absent from the research to date. Whilst intuition could point towards travel time use being a coping mechanism for disrupted travel, via the ability to participate in some activities remotely, this has not been researched in studies to date. Naturally, such study would be a valuable endeavour if pointing towards what conditions would be required for passengers to make use of such mitigation strategies to reduce their dissatisfaction and consequently, cost of disruption.

Empirically, the studies (RP and SP alike) report between 10% and 40% reduction in VTT, when comparing conditions allowing undertaking the intended activities versus inability to do so (see **Table 24**). More extreme values have been observed, such as close to zero (for example in the case of driving private car) or above 30% (enabling leisure activities, such as book reading amongst leisure). In practice, impacts differing between purposes of travel can largely be attributed to whether the preferred profile of activities can be undertaken or not. For example, enabling work activities during leisure travel will carry lesser impact on the value of travel time than enabling leisure activities and vice versa.

The field has not yet seen a consensus concerning the measurement of productivity (or more broadly quality of activities). Self-reported measurement relative to office conditions appears to be a convenient way, though the extent to which it accurately describes the actual conditions remains underexplored. Nor has there been an agreement whether, and if so how, to describe quality of such activities in SP and RP surveys. Emerging approaches look at composite indices or use of new concepts, such as 'un-'productivity (looking at barriers to optimal productivity) or introducing the measures of enjoyment and fitness, to more accurately compare activities undertaken while travelling to what they would be in non-travel contexts. A promising avenue appears to be one concerning treatment of such metrics as latent (not directly observable) and acknowledging their possible variation during stages of travel, due to variation in journey conditions and traveller's performance. In short, the measurement of in-travel productivity is still an evolving research interest – the closest we have to an off-the-shelf approach that is reasonably established would be a Hensher style analysis. Additional work may be required in this area, both for business and non-business journeys to arrive at robust measures of the concepts of interest.

The AV context has been explored widely in the VTT literature, given the hypothesis that such environments present the seemingly ideal environment for undertaking activities while travelling. Existing empirical studies (based on SP data) have confirmed, in several contexts, reduction in the value of travel time implied by enablement of travel time activities. At the same time, there has been no consensus about what is meant when the concept of AV is presented, in terms of whether it concerns a private or public mode, the extent of sharing, presence of conditions leading to motion sickness, condition of the interior and so on. In fact, the productive use of travel time is typically represented as a stated preference for travel-based time use or in the form of an alternative where multi-tasking is explicitly described as being possible (e.g. autonomous vehicle). Similar conclusion extends to more futuristic options, such as unmanned aerial vehicles.

Apart from the AV context, the existing empirical studies focus on the rather traditional approaches to mobility, by segmenting by modes of transport, often focusing on the main one. The multi-modality or emerging mobility concepts, such as MaaS, and behaviours, e.g. remote/hybrid work, are not discussed in the current literature concerning travel time use.

## 6. Congestion, reliability, and crowding

Whilst differences in comfort factors across modes of transport have been discussed in detail in **Section 4.4**, another strand of segmentation of the VTT relates to variations of the comfort of travel within mode whilst travelling, and the implications of changes in travel conditions on scheduling decision. Altogether research looking into congestion, reliability and crowding effects and their connection with the VTT show that these concepts are not independent from each other since increases in congestion typically cause changes in travel time reliability due to impacts on delay. Similarly, crowding in public transport services not only affects the comfort of travel inside busses and carriages, but also introduces challenges in additional boarding and alighting times potentially resulting in bunching of public transport services and thereby increasingly unreliable journey times.

A fundamental problem in this area is whether you are seeking to value;

- delay (i.e. unit value of a minute of lateness against some reference – typically the timetable for PT, and free-flow conditions for car);
- reliability (i.e. unit value of variability in travel time);
- congestion (i.e. unit value of a minute spent in congested conditions, possibly discriminating between different levels of congestion);
- crowding (i.e. unit value of a minute spent in crowded conditions, possibly discriminating between sitting and standing conditions).

Valuation studies have looked at all four concepts in some shape or form. Transport models have tended to focus upon delay (in the case of road) and crowding (in the case of rail).

ARUP et al. (2015a) considered three distinct SP games:

- SP1 considered time vs. cost
- SP2 considered time vs. cost vs. reliability
- SP3 considered time vs. cost vs. congestion, and time vs. cost vs. crowding

Intuition and evidence (e.g. National Highways' Myriad model) suggest there exists a relationship between the concepts of congestion (crowding) and reliability – in the sense that variability in traffic levels can lead to variability in arrival time. Indeed, the preamble to SP2 in the 2014/15 study stated that the reasons for (unreliability) could be '*breakdowns, unplanned roadworks, or general traffic*'. On this basis and recalling that the reliability ratio was estimated to be 0.4 across all journey purposes, there is an argument that this captures aspects of both reliability and congestion.

ITS (2022) remarked that: '*...this position is less than ideal, since there should be a clearer distinction between congestion and reliability – with intuition suggesting that the former should relate to traffic conditions in the mass of the travel time distribution, and the latter to more extreme events in the tail*'.<sup>22</sup> In this context, the 2022 report highlighted the SP game developed by David Hensher (**Figure 19**), which includes congestion and reliability as separate attributes within the same game, thereby addressing to some extent the problem of duplication/confounding. On the other hand, ITS (2022) criticised the simplicity of Hensher's treatment of the travel time distribution and were sceptical of how much insight could be gleaned on relative valuations within the mass vs. the tail.

They suggested a possible 'third' way, of developing a new version of SP3 which explicitly distinguishes between FF and delayed time (in the sense of 'number of minutes late') but does not explicitly refer

---

<sup>22</sup> Note that the latter is different from the current TAG definition (see para 6.1.1 TAG unit A1.3).



to congestion. Intuition suggests that congestion will be a principal cause of delay, but the attraction of this approach is that delay is an objective metric universally understood by both SP respondents and network modellers.

ITS (2022) concluded: *'All things considered, there is a good argument for enhanced SP research that explicitly separates out i) recurrent congestion leading to modest delays and ii) relatively rare events leading to substantial delays. In terms of the Department's appraisal practice, this might mean incorporating the value of avoiding delay into Level 1 benefits via the congestion/delay multipliers while leaving the value of reducing delay due to extreme events in Level 2 and analysed through a different type of model'.*

**Transport Study**

Games 1

|                                          | Details of Your Recent Trip | Alternative Road A | Alternative Road B | Alternative Road C |
|------------------------------------------|-----------------------------|--------------------|--------------------|--------------------|
| Time in free-flow (mins)                 | 15                          | 14                 | 16                 | 16                 |
| Time slowed down by other traffic (mins) | 10                          | 12                 | 8                  | 12                 |
| Time in Stop/Start conditions (mins)     | 5                           | 4                  | 6                  | 4                  |
| Uncertainty in travel time (mins)        | +/- 10                      | +/- 12             | +/- 8              | +/- 8              |
| Running costs                            | \$ 2.20                     | \$ 2.40            | \$ 2.40            | \$ 2.10            |
| Toll costs                               | \$ 2.00                     | \$ 2.10            | \$ 2.10            | \$ 1.90            |

If you take the same trip again, which road would you choose?

☒ Current Road
 ☐ Road A
 ☐ Road B
 ☐ Road C

If you could only choose between the new roads, which would you choose?

☐ Road A
 ☐ Road B
 ☐ Road C

[Go to Game 2 of 6](#)

Figure 19: Example of a choice card used by Hensher (2004)

## 6.1 Congested values of travel time (CVTT) and congestion multipliers

### 6.1.1 UK position – 2014/15 study

SP1 of the 2014/15 study considered trade-offs between travel time and cost assuming 'Average' traffic congestion but certainty of travel time. SP3 further decomposed travel time according to three traffic levels – namely 'Free-Flow' (FF), 'Light Congestion' (LC) and 'Heavy Congestion' (HC). Then adopting SP1 as the headline VTT, SP3 was used to derive Congested Values of Travel Time (CVTT) as multipliers of the SP1 values.

On this basis, **Table 25** presents the 2014/15 estimates of CVTT by journey purpose. Note that they encompass a wide range from 0.42 (in the case of Free-Flow for Employer Business) to 1.89 (in the case of Heavy Congestion for Other Non-Work). It will be shown in due course that the HC value is high by accepted standards. Partly for this reason, and partly due to the technical challenges of implementing CVTT in transport models, the multipliers have not to date been adopted in TAG.



Table 25: CVTT multipliers from SP3 – expressed relative to Average traffic conditions from SP1 (ARUP et al., 2015a)

| <b>Traffic conditions</b> | <b>Commute</b> | <b>EB</b> | <b>OtherNW</b> |
|---------------------------|----------------|-----------|----------------|
| Free-Flow                 | 0.51           | 0.42      | 0.47           |
| Light Congestion          | 0.72           | 0.68      | 0.83           |
| Heavy Congestion          | 1.37           | 1.26      | 1.89           |

It is more natural in the modelling context to express CVTT as a multiple of the VTT for FF (as opposed to Average) traffic conditions. **Table 26** converts the multipliers from **Table 25** accordingly. Since the terminology has a degree of precedent in the modelling literature, multiples of the VTT for FF are referred to below as ‘*M*’ multipliers.

Table 26: CVTT multipliers from SP3 expressed relative to FF traffic conditions (‘*M*’ multipliers) from SP3 (ARUP et al., 2015a)

| <b>Traffic conditions</b> | <b>Commute</b> | <b>EB</b> | <b>OtherNW</b> |
|---------------------------|----------------|-----------|----------------|
| Free-flow                 | 1.00           | 1.00      | 1.00           |
| Light congestion          | 1.40           | 1.61      | 1.76           |
| Heavy congestion          | 2.66           | 2.99      | 3.98           |

The distinction between **Table 25** and **Table 26** highlights two broad approaches that could potentially be followed for deriving a headline VTT and associated congestion multipliers, namely:

- **Approach 1:** derive the headline VTT from SP1, and multipliers from SP3.
- **Approach 2:** derive both the headline VTT and multipliers from SP3, where the headline VTT would in this case be a weighted average of time spent in FF, LC and HC from SP3.

#### 6.1.2 UK position – CVTT workstream

Over the last 5 years or so, the Department has undertaken a root and branch review of the case for incorporating congestion multipliers in modelling and appraisal, and in the course of this workstream has commissioned five reports:

1. The 2018 report from WSP, RAND Europe and Mott Macdonald on ‘Congestion Dependent Values of Time in Transport Modelling’, dated March 2018 and in the public domain (WSP et al., 2018). The remit of this study was to explore the implementation of congested VTT multipliers in practical modelling and appraisal.
2. The 2019 follow-on report by WSP and Mott MacDonald (2019), which extended their work to a more substantive implementation in the PRISM multi-modal model of the West Midlands.
3. The 2020 Forward Look report commissioned from ITS Leeds, which reviewed the 2018 and 2019 reports and formulated a position on the applicability of the 2014/15 congestion multipliers to modelling and appraisal (ITS, 2020).
4. The 2022 report by ITS Leeds, which reanalysed the congestion multipliers from the 2014/15 study (ITS, 2022).

5. The 2022 report by WSP and Mott MacDonald, which undertook further PRISM testing (WSP and Mott MacDonald, 2022).

The scope of WSP et al. (2018) was to assess the feasibility of incorporating congestion multipliers into highways and other transport models. The report considers various insights on this topic, namely:

- a technical review of the CVTT estimates emanating from the 2014/15 study by Arup, ITS Leeds & Accent;
- a literature review of CVTT estimates from the UK and elsewhere in Europe;
- empirical testing of the route choice implications of CVTT using Trafficmaster data for a small sample of OD pairs.

WSP et al. (2018) concluded that: *'...there is a strong body of evidence that travel time is valued more highly in congested conditions. However, that evidence is currently insufficient to allow us to formulate VTT as a function of congestion in a way that would allow it to be included in modelling and appraisal. Further work is therefore required to get to the position where congestion-dependent VTT could become a TAG requirement'*.

The scope of WSP and Mott MacDonald (2019) was to:

- provide a fully defined 'proof of concept' for applying CVTT in a transport model with an appropriate level of geographical coverage, that consists of both assignment and demand modelling;
- improve the Department's understanding of the potential impact of including CVTT in the modelling and appraisal of highway schemes;
- highlight any barriers to robust implementation of CVTT in extant models, and identify what additional work could be undertaken to overcome such barriers.

Adopting the PRISM model of the West Midlands as a testbed, WSP and Mott MacDonald (2019) concluded that: *'...there is the potential for congestion multipliers to have a significant impact on the design and appraisal of transport schemes. However, it would not be acceptable to include the multipliers in the appraisal without also using them in the supporting transport modelling...Further work would be required before this can be recommended as the standard approach in TAG'*.

The scope of ITS (2020) was to consider three questions:

- Has the WSP and Mott MacDonald (2019) CVTT study met its objectives?
- How would the Department implement CVTT in appraisal and what are the barriers to this?
- What are the future research needs for CVTT?

ITS (2020) concluded that WSP and Mott MacDonald (2019) had *'demonstrated through the case study that it is technically feasible to represent CVTT in modelling and appraisal. There is no technical showstopper. The issues are (a) whether there is enough behavioural evidence to support a value of M different from unity; (b) whether M multipliers should be deployed in appraisal only or throughout modelling and appraisal; and (c) what testing regime would be required to enable the Department to*

*move forward from case study to implementation in guidance with confidence*. *M* here refers to multiplier relative to free flow time.

ITS (2020) recommended several additional desk studies which could help to move the CVTT workstream forward incrementally, in advance of more substantive testing. These studies included the following:

- a) A small piece of re-analysis of 2014/15 SP1 and SP3 to examine the relationships between CVTT values, journey purpose and journey length.
- b) Further work to compare the relative merits of Approaches 1 and 2 referred to above.
- c) A small piece of re-analysis of WSP and Mott MacDonald (2019) to examine the implied cost elasticities of demand (in terms of trips and distance) with respect to travel time arising from the future year forecasts.
- d) A review of work elsewhere on continuous values of CVTT together with consideration of how such values could be used to map on to model outputs such as FF time and congested time.

The scope of ITS (2022) comprised points a), b) and d), whilst WSP and Mott MacDonald (2022) responds to point c) whilst also examining the use of CVTT during matrix building and its implications for modelling and appraisal.

Key conclusions from ITS (2022) include:

- a) In the near term, the policy focus should be on Approach 1 rather than Approach 2. That is to say, the headline VTT should be derived from SP1 and congestion multipliers from SP3.
- b) Whilst the headline VTTs from 2014/15 would seem robust, there are significant barriers to the implementation of the 2014/15 congestion multipliers in modelling and appraisal. On this basis, there is a good case for commissioning new SP research focussed on re-surveying the congestion multipliers.
- c) Re-surveying the congestion multipliers is not without its challenges. Therefore, the prudent approach would be to first commission a scoping study exploring several technical issues:
  - Context specificity and whether CVTT varies systematically urban/interurban, short/long distance etc.
  - Understanding the overlap between congestion and reliability and arriving at recommendations for both considered together.
  - Deciding whether to formulate further SP work in terms of 'Delay' rather than 'Congestion' – thereby bridging the gap to modelling (e.g. the Hensher approach).
  - Deciding what SP design would be required to get a better handle on the scenarios presented to SP respondents vis-a-vis the continuous function.
  - Whether there is a role for RP data to help validate SP.
- d) Beyond the near term, a broader set of issues concerning guidance on CVTT should be considered, encompassing the following considerations:
  - Review the policy and analysis requirements of headline VTTs vs. CVTT multipliers.

- Against those policy/analysis requirements, assess the relative merits of Approaches 1-3 (where the latter would formulate the headline VTT in terms of FF), as well as the implementation costs of each approach.
- Consider the policy/analysis benefits and risks of offering more than one of Approaches 1-3 in guidance.
- The above considerations should be cognisant of the need for consistency and complementarity across the complete VTT portfolio, encompassing all modes and all types of travel time.

Key conclusions from WSP and Mott MacDonald (2022) were that:

- a) The use of congestion-dependent values of travel time (CVTT):
  - Has a significant impact on forecast flows, with implications for options appraisal and scheme design.
  - Has a significant impact on the economic appraisal of road schemes, with implications for investment decisions.
  - The only technical significant technical difficulty is the lack of the required functionality in some commonly used modelling software.
- b) They state the important qualification that the true values of CVTT and VTT (for FF) are unknown, such that the scale of the above impacts cannot be determined. However, they point to 'robust evidence for appropriate values from the 2015 UK VTT study'.
- c) Defining congestion as a categorical variable (e.g. free flow, medium or heavy congestion) is problematic for both modelling and appraisal. Instead, congestion should be a continuous variable, and ideally one that can be readily calculated from standard transport model outputs (such as travel time, free-flow time, delay, distance etc.).
- d) Notably, traffic assignment models are "fussiest" about the exact form of the function for congestion. Any definition of congestion and congested VTT that works in highway assignment models should work fine in demand modelling and appraisal.
- e) Further work is proposed in the following areas:
  - Review the implementation of CVTT in modelling and appraisal, with the minimum requirement being to quantify M and the value of FF.
  - Review any evidence that scheme flows are consistently underestimated due to the omission of CVTT.
  - Analyse observed routing patterns, to see if they are consistent with a higher VTT in congested areas.
  - Engage with software developers to ensure that CVTT can be represented in the most used modelling platforms.

- Following completion of the above, further testing on a range of schemes to get a broader understanding of the possible implications of using CVTT.

### 6.1.3 Empirical evidence on congestion multipliers

Section 5 of WSP and Mott MacDonald (2019) provides a useful and recent review of evidence, which we assimilate here with the review undertaken by Wardman and Nicolás Ibáñez (2012), whilst also adding more recent work. The evidence base is summarised in **Table 27**. Key observations include the following:

- With around 100 multipliers recorded from the UK and (mostly) from countries not dissimilar to the UK, **Table 27** gives a reasonably informative evidence base.
- Whilst exhibiting variation across the type of congestion and journey purpose, the congestion multipliers in general terms show reasonable consistency in the range 1-2.
- It is possible that there are discrepancies across the studies in terms of the denominator of the congestion multiplier – especially in terms of whether the denominator refers to the value of FF time or average travel time (which could involve a degree of congestion). It stands to reason that multipliers defined in terms of FF will be higher than those defined in terms of average conditions, all else equal.
- In terms of journey purpose, it is generally found that the CVTT for business is highest, the CVTT for leisure lowest, with commuting intermediate. This replicates the ordering one would expect for VTT.
- In terms of traffic conditions, it is generally found that the multiplier for gridlock is highest, the multiplier for FF lowest (and below one when the base is average travel time), with some level of congestion intermediate. This accords with intuition.
- Most multipliers have been estimated using SP, but there is some RP evidence corroborating the SP evidence base.
- Most of the SP-based studies have employed experiments based on either route choice or unlabelled alternatives – but there is no obvious pattern in terms of the multipliers emanating from each.
- Whilst most studies consider a single generic definition of ‘congestion’, some studies successfully estimate monotonically increasing multipliers for up to six levels of congestion.
- All studies that estimated congestion multipliers seem to have used categorical definitions instead of continuous, although they do differ in the number of categories used.
- Some of the studies included reliability as a distinct variable alongside congestion, whilst most did not. Therefore, it is possible that, in many of the studies, the CVTT estimates are confounded to a greater or lesser degree with reliability.

Surveying the range of multipliers presented in the table serves to emphasise that the multipliers for HC from the 2014/15 UK study are at the higher end of the range – but the 2014/15 study is not unique in estimating multipliers markedly in excess of 2.

Table 27: Evidence on congestion multipliers

| Study                                | Context               | Time        | Multiplier | Data                                        | Purpose      |
|--------------------------------------|-----------------------|-------------|------------|---------------------------------------------|--------------|
| Wardman (1986)                       | Not known             | Congested   | 1.39       | SP route choice                             | Commute      |
|                                      |                       | Congested   | 1.46       | SP route choice                             | Business     |
|                                      |                       | Congested   | 1.28       | SP route choice                             | Other        |
|                                      |                       |             |            |                                             |              |
| Hensher et al. (1990)                | Sydney                | Congested   | 1.70       | SP unlabelled alternatives                  | Business     |
|                                      |                       |             |            |                                             |              |
| Kazakov et al. (1993)                | Canada                | Congested   | 1.37       | SP unlabelled alternatives                  | Commute      |
|                                      |                       |             |            |                                             |              |
| Douglas Economics (1996)             | Sydney                | Congested   | 1.31       | SP mode choice                              | All          |
|                                      |                       |             |            |                                             |              |
| MVA (1997)                           | Nottingham            | Congested   | 1.18       | SP mode choice car users                    | Non-business |
|                                      |                       |             | 1.57       | SP mode choice P&R users                    | Commute      |
|                                      |                       |             | 1.75       | SP mode choice P&R users                    | Other        |
|                                      |                       |             |            |                                             |              |
| HCG & Accent (1999)                  | UK national study     | Congested   | 1.70       | SP route choice                             | Commute      |
|                                      |                       | Congested   | 2.04       | SP route choice                             | Business     |
|                                      |                       | Congested   | 1.90       | SP route choice                             | Other        |
|                                      |                       |             |            |                                             |              |
| Calfee et al. (2001)                 | US Metropolitan areas | Congested   | ≈3.00      | SP ranking                                  | Commute      |
|                                      |                       |             |            |                                             |              |
| Hensher (2001a)                      | NZ Inter-urban        | Slowed down | >2.70      | SP unlabelled alternatives plus current     | All          |
|                                      |                       | Stop-start  | >5.70      | SP unlabelled alternatives plus current     | All          |
|                                      |                       |             |            |                                             |              |
| Hensher (2001b)                      | NZ Suburban           | Slowed down | 1.33       | SP unlabelled alternatives plus current     | All          |
|                                      |                       | Stop-start  | 2.82       | SP unlabelled alternatives plus current     | All          |
|                                      |                       |             |            |                                             |              |
| Beca Carter Hollings & Ferner (2002) | NZ Urban              | Slowed down | 1.07       | SP unlabelled alternatives                  | All          |
|                                      |                       | Stop-start  | 1.49       | SP unlabelled alternatives                  | All          |
|                                      | NZ Rural              | Slowed down | 1.15       | SP unlabelled alternatives                  | All          |
|                                      |                       | Stop-start  | 1.67       | SP unlabelled alternatives                  | All          |
|                                      |                       |             |            |                                             |              |
| Nielsen et al (2002)                 | Copenhagen            | Congested   | 1.31       | SP route choice                             | Commute      |
|                                      |                       |             | 1.64       | SP route choice                             | Business     |
|                                      |                       |             | 2.16       | SP route choice                             | Other        |
|                                      |                       |             |            |                                             |              |
| Steer (2004)                         | 2004                  | FF          | 1.00       | SP mode choice & route choice               | All          |
|                                      |                       | Busy        | 1.20       | SP mode choice                              | All          |
|                                      |                       | LC          | 1.11       | SP route choice                             | All          |
|                                      |                       | HC          | 1.46/1.47  | SP mode choice & route choice               | All          |
|                                      |                       | Stop start  | 1.67/1.61  | SP mode choice & route choice               | All          |
|                                      |                       | Gridlock    | 1.96       | SP mode choice                              | All          |
|                                      |                       |             |            |                                             |              |
| Eliasson (2004)                      | Stockholm             | Slow queue  | 1.42       | SP unlabelled alternatives                  | Morning      |
|                                      |                       |             | 1.53       | SP unlabelled alternatives                  | Afternoon    |
|                                      |                       |             | 1.44       | SP unlabelled alternatives                  | Business     |
|                                      |                       |             |            |                                             |              |
| Nielsen (2004)                       | Copenhagen            | Congested   | 1.61       | SP route choice and unlabelled alternatives | All          |
|                                      |                       |             |            |                                             |              |
| Faber Maunsell (2005)                | Brisbane              | Stop-start  | 1.65       | SP route choice                             | All          |
|                                      |                       |             | 1.38       | SP unlabelled alternatives                  | All          |
|                                      |                       |             | 1.89       | RP route choice                             | All          |
|                                      |                       |             | 1.28       | RP/SP route choice                          | Business     |
|                                      |                       |             | 1.34       | RP/SP route choice                          | Commute      |
|                                      |                       |             | 1.52       | RP/SP route choice                          | Other        |

|                           |                                    |                       |      |                                         |              |
|---------------------------|------------------------------------|-----------------------|------|-----------------------------------------|--------------|
|                           |                                    |                       |      |                                         |              |
| Deloitte (2006)           | Auckland                           | Congested             | 1.12 | SP unlabelled alternatives              | All          |
|                           |                                    |                       |      |                                         |              |
| Fosgerau (2006)           | Denmark                            | Congested             | 1.52 | SP unlabelled alternatives              | Non-business |
|                           |                                    |                       |      |                                         |              |
| Confidential (2007)       | US major city                      | Congested             | 1.60 | SP unlabelled alternatives              | All          |
|                           |                                    |                       |      |                                         |              |
| Hensher et al. (2007)     | Sydney                             | Slowed down           | 1.25 | SP unlabelled alternatives plus current | Non-commute  |
|                           |                                    |                       |      |                                         |              |
| Rich & Nielsen (2007)     | Copenhagen                         | Congested             | 1.15 | RP route choice                         | Commute      |
|                           |                                    |                       | 1.65 | RP route choice                         | Business     |
|                           |                                    |                       | 2.00 | RP route choice                         | Leisure      |
|                           |                                    |                       |      |                                         |              |
| Fosgerau et al. (2007)    | Denmark                            | Congested             | 0.88 | SP unlabelled alternatives              |              |
|                           |                                    |                       |      |                                         |              |
| Rose et al. (2008)        | Sydney                             | Slowed down           | 1.35 | SP unlabelled alternatives plus current | Commute      |
|                           |                                    | Stop-start            | 2.79 | SP unlabelled alternatives plus current | Commute      |
|                           | Santiago <200 min                  | Slowed down           | 1.25 | SP unlabelled alternatives plus current | All          |
|                           |                                    | Stop-start            | 1.43 | SP unlabelled alternatives plus current | All          |
|                           | Taiwan                             | Slowed down           | 1.31 | SP unlabelled alternatives plus current | All          |
|                           | Suburban                           | Stop-start            | 2.08 | SP unlabelled alternatives plus current | All          |
|                           |                                    |                       |      |                                         |              |
| Murphy (2009)             | Rostock 2004                       | Congested             | 1.05 | SP unlabelled alternatives              | All          |
|                           |                                    |                       | 1.14 | RP tolled vs. untolled routes           | All          |
|                           | Montreal 2005                      | Stop-start            | 1.65 | SP unlabelled alternatives              | Commute      |
|                           |                                    |                       | 1.23 | SP unlabelled alternatives              | Business     |
|                           |                                    |                       | 2.16 | SP unlabelled alternatives              | Other        |
|                           | Riga 2005                          | Stop-start            | 1.00 | SP route choice                         | Commute      |
|                           |                                    |                       | 1.07 | SP route choice                         | Business     |
|                           |                                    |                       | 1.07 | SP route choice                         | Leisure      |
|                           | Serbia 2007                        | Stop-start            | 1.73 | SP route choice                         | Commute      |
|                           |                                    |                       | 1.14 | SP route choice                         | Business     |
|                           |                                    |                       | 1.05 | SP route choice                         | Leisure      |
|                           | Brisbane 2007                      | Stop-start            | 1.19 | SP route choice                         | Commute      |
|                           |                                    |                       | 1.24 | SP route choice                         | Business     |
|                           |                                    |                       | 1.22 | SP route choice                         | Other        |
|                           |                                    |                       | 1.04 | RP route choice                         | All          |
|                           |                                    |                       |      |                                         |              |
| Maunsell Australia (2009) | Singapore                          | LC & HC               | 1.16 | SP route choice                         | Commute      |
|                           |                                    | Stop-start & gridlock | 1.71 | SP route choice                         | Commute      |
|                           |                                    | LC & HC               | 1.12 | SP route choice                         | Business     |
|                           |                                    | Stop-start & gridlock | 1.68 | SP route choice                         | Business     |
|                           |                                    | LC & HC               | 1.00 | SP route choice                         | Other        |
|                           |                                    | Stop-start & gridlock | 1.36 | SP route choice                         | Other        |
|                           |                                    |                       |      |                                         |              |
| Hensher & Greene (2010)   | Sydney                             | Slowed down           | 1.68 | SP unlabelled alternatives plus current | Commute      |
|                           |                                    |                       |      |                                         |              |
| Abrantes & Wardman (2011) | UK urban and inter-urban 1986-2008 | Congested             | 1.54 | Meta                                    | All          |
|                           |                                    |                       | 1.43 | Meta                                    | Business     |
|                           |                                    |                       | 1.59 | Meta                                    | Commute      |
|                           |                                    |                       | 1.54 | Meta                                    | Other        |
|                           |                                    |                       | 1.51 | Meta                                    | RP           |
|                           |                                    |                       | 1.55 | Meta                                    | SP           |

|                         |                           |                               |           |                            |          |
|-------------------------|---------------------------|-------------------------------|-----------|----------------------------|----------|
| Wardman & Ibanez (2012) | UK M6 Toll Road 2006      | FF                            | 1.00      | SP route choice            |          |
|                         |                           | Busy                          | 1.14      | SP route choice            |          |
|                         |                           | LC                            | 1.23      | SP route choice            |          |
|                         |                           | HC                            | 1.31      | SP route choice            |          |
|                         |                           | Stop-start                    | 1.45      | SP route choice            |          |
|                         |                           | Gridlock                      | 1.78      | SP route choice            |          |
|                         |                           |                               |           |                            |          |
| Prato et al. (2014)     | Copenhagen                | Congested                     | 1.46-1.50 | RP route choice            | Peak     |
|                         |                           | Congested                     | 1.25-1.26 | RP route choice            | Off-peak |
|                         |                           |                               |           |                            |          |
| Arup et al. (2015)      | UK national study 2014/15 | FF                            | 1.00      | SP unlabelled alternatives | Commute  |
|                         |                           | FF                            | 1.00      | SP unlabelled alternatives | Business |
|                         |                           | FF                            | 1.00      | SP unlabelled alternatives | Leisure  |
|                         |                           | LC                            | 1.40      | SP unlabelled alternatives | Commute  |
|                         |                           | LC                            | 1.61      | SP unlabelled alternatives | Business |
|                         |                           | LC                            | 1.76      | SP unlabelled alternatives | Leisure  |
|                         |                           | HC                            | 2.66      | SP unlabelled alternatives | Commute  |
|                         |                           | HC                            | 2.99      | SP unlabelled alternatives | Business |
|                         |                           | HC                            | 3.98      | SP unlabelled alternatives | Leisure  |
|                         |                           |                               |           |                            |          |
| Flügel et al. (2020)    | Norwegian national study  | FF, car driver                | 0.8       | SP unlabelled alternatives | Commute  |
|                         |                           | FF, car driver                | 0.9       | SP unlabelled alternatives | Business |
|                         |                           | FF, car driver                | 0.9       | SP unlabelled alternatives | Leisure  |
|                         |                           | FF, car driver                | 0.9       | SP unlabelled alternatives | All      |
|                         |                           | FF, car pax                   | 0.9       | SP unlabelled alternatives | Commute  |
|                         |                           | FF, car pax                   | 1.0       | SP unlabelled alternatives | Business |
|                         |                           | FF, car pax                   | 0.9       | SP unlabelled alternatives | Leisure  |
|                         |                           | FF, car pax                   | 0.9       | SP unlabelled alternatives | All      |
|                         |                           | Mod congestion, car driver    | 1.2       | SP unlabelled alternatives | Commute  |
|                         |                           | Mod congestion, car driver    | 1.1       | SP unlabelled alternatives | Business |
|                         |                           | Mod congestion, car driver    | 1.3       | SP unlabelled alternatives | Leisure  |
|                         |                           | Mod congestion, car driver    | 1.2       | SP unlabelled alternatives | All      |
|                         |                           | Mod congestion, car pax       | 1.2       | SP unlabelled alternatives | Commute  |
|                         |                           | Mod congestion, car pax       | 1.1       | SP unlabelled alternatives | Business |
|                         |                           | Mod congestion, car pax       | 1.3       | SP unlabelled alternatives | Leisure  |
|                         |                           | Mod congestion, car pax       | 1.2       | SP unlabelled alternatives | All      |
|                         |                           | Severe congestion, car driver | 2.3       | SP unlabelled alternatives | Commute  |
|                         |                           | Severe congestion, car driver | 1.4       | SP unlabelled alternatives | Business |
|                         |                           | Severe congestion, car driver | 2.4       | SP unlabelled alternatives | Leisure  |
|                         |                           | Severe congestion, car driver | 2.3       | SP unlabelled alternatives | All      |
|                         |                           | Severe congestion, car pax    | 2.0       | SP unlabelled alternatives | Commute  |
|                         |                           | Severe congestion, car pax    | 1.3       | SP unlabelled alternatives | Business |
|                         |                           | Severe congestion, car pax    | 2.0       | SP unlabelled alternatives | Leisure  |
|                         |                           | Severe congestion, car pax    | 1.9       | SP unlabelled alternatives | All      |



## 6.2 The Value of travel time reliability (VTTR) and the reliability ratio

de Jong and Kouwenhoven (2020a) describe the two most prominent conceptual approaches associated with the value of travel time reliability (VTTR), respectively the mean-dispersion model and scheduling models. The former defines the VTTR as the traveller's WTP for reducing the standard deviation of travel time. The latter associates travel time unreliability with penalties for arriving early or late. Bates et al. (2001) and Börjesson et al. (2012a) highlight that the VTTR can be derived from the results from a departure time (i.e. scheduling) model and that hence the two approaches are theoretically equivalent (under certain conditions). Having obtained estimates the VTT and the VTTR, it is not uncommon to calculate the relative importance of saving travel time versus reducing travel time uncertainty. The ratio  $\frac{VTTR}{VTT}$  is commonly denoted as the 'reliability ratio'.

### 6.2.1 Empirical evidence on the VTTR and the reliability ratio

de Jong and Bliemer (2015) provide an overview of empirical findings on the valuation of travel time reliability through the reliability ratio across countries until 2013 as reported below in **Figure 20**.

| Mode             | Purpose  | This study | Norway 2018 | Norway 2010                | UK 2014 |
|------------------|----------|------------|-------------|----------------------------|---------|
| Car              | Commute  | 0.27       | 0.4         | 0.25 (long) - 0.42 (short) | 0.33    |
|                  | Business | 0.21       | 0.4         |                            | 0.42*   |
|                  | Other    | 0.35       | 0.4         |                            | 0.35    |
| Train            | Commute  | 0.32       | 0.4         | 0.54 (long)                | -       |
|                  | Business | 0.11       | 0.4         |                            | -       |
|                  | Other    | 0.27       | 0.4         |                            | -       |
| Bus, tram, metro | Commute  | 0.65       | 0.4         | 0.69 (short)               | -       |
|                  | Business | 0.61       | 0.4         |                            | -       |
|                  | Other    | 0.56       | 0.4         |                            | -       |
| Air              | Business | 0.30       | -           | 0.20 (long)                | -       |
|                  | Other    | 0.28       | -           |                            | -       |
| Recr. Navigation | Other    | 0          | -           | -                          | -       |

\*only applies to employee part of the VTT

**Table 50** Comparison of reliability ratios between the 2022 study and national studies in Norway and the UK

Figure 21 contrasts the results from the most recent Dutch national VTT study against those of the UK and Norway, where

|      |                                         | Value by Purpose |      |          |         |                 |        |
|------|-----------------------------------------|------------------|------|----------|---------|-----------------|--------|
| Mode | Attribute                               | Education        | Work | Shopping | Leisure | Business Travel | All    |
| Car  | Standard deviation                      | 3.21             | 3.45 | 3.51     | 3.09    | 6.54            | 3.61   |
| Car  | VOT (€/h)                               | 3.90             | 4.87 | 4.29     | 4.03    | 8.38            | 4.66   |
| Car  | Standard deviation/VOT                  | 0.7              | 0.7  | 0.7      | 0.7     | 0.7             | 0.7    |
| PT   | Mean expected unscheduled delay         | 4.66             | 5.10 | 4.28     | 4.82    | 15.97           | 5.48   |
| PT   | Mean expected unscheduled early arrival | 1.81             | 1.98 | 1.67     | 1.88    | 6.22            | 2.13   |
| PT   | VOT (€/h)                               | 4.39             | 4.47 | 5.11     | 4.35    | 7.01            | 4.83   |
| PT   | RR: VOR_late/VOT                        | 0.9              | 1.0  | 0.7      | 0.9     | 1.7             | 0.9    |
| PT   | RR: VOR_early/VOT                       | 0.3              | 0.4  | 0.3      | 0.3     | 0.7             | 0.4    |
| Air  | Mean expected unscheduled delay         | na               | na   | na       | 38.44   | 51.27           | 46.60  |
| Air  | Mean expected unscheduled early arrival | na               | na   | na       | 90.16   | 120.25          | 109.30 |
| Air  | VOT (€/h)                               | na               | na   | na       | 25.45   | 38.76           | 33.67  |
| Air  | VOR_late/VOT                            | na               | na   | na       | 1.4     | 1.4             | 1.4    |
| Air  | VOR_early/VOT                           | na               | na   | na       | 3.3     | 3.2             | 3.2    |

Note: RR=reliability ratio (population weighted); PT=public transport; na=not applicable

**Table 2** VOR, VOT, and Reliability Ratio

Figure 22 presents additional evidence from the German national VTT study. Most of the evidence-base supports the conclusion that the VTTR<VTT and hence a reliability ratio smaller than 1 (although there is possibly a transposition error in the table, as the actual calculations of the reliability ratio for Germany appear incorrect). There is, however, a significant degree of heterogeneity in the reliability

ratio across modes of transport, journey purposes and countries. Unfortunately, the evidence-base is not large enough to disentangle the impact across these different dimensions. Comparing the different national VTT studies against each other, the Dutch national VTT study makes a step difference by segmenting the reliability ratio by mode and purpose combinations, whereas the most recent Norwegian study assumed a uniform distribution across modes and journey purposes and the UK only estimated the reliability ratio for car based trips but segmented by journey purpose (see

| Mode             | Purpose  | This study | Norway 2018 | Norway 2010                | UK 2014 |
|------------------|----------|------------|-------------|----------------------------|---------|
| Car              | Commute  | 0.27       | 0.4         | 0.25 (long) - 0.42 (short) | 0.33    |
|                  | Business | 0.21       | 0.4         |                            | 0.42*   |
|                  | Other    | 0.35       | 0.4         |                            | 0.35    |
| Train            | Commute  | 0.32       | 0.4         | 0.54 (long)                | -       |
|                  | Business | 0.11       | 0.4         |                            | -       |
|                  | Other    | 0.27       | 0.4         |                            | -       |
| Bus, tram, metro | Commute  | 0.65       | 0.4         | 0.69 (short)               | -       |
|                  | Business | 0.61       | 0.4         |                            | -       |
|                  | Other    | 0.56       | 0.4         |                            | -       |
| Air              | Business | 0.30       | -           | 0.20 (long)                | -       |
|                  | Other    | 0.28       | -           |                            | -       |
| Recr. Navigation | Other    | 0          | -           | -                          | -       |

\*only applies to employee part of the VTT

**Table 50** Comparison of reliability ratios between the 2022 study and national studies in Norway and the UK

Figure 21). Confidence intervals were not reported in the source studies, but it would not be entirely surprising if the reliability ratios are not all largely overlapping. Across the three studies a significant degree of consistency can be observed, and most reliability ratios are not too distinct from the Norwegian estimate of 0.4.

Comparing over time, the Dutch reliability ratios have dropped significantly relative to their 2009/2011 study. This is partly attributed to the difference in presentation of the SP survey format, and the change in the modelling approach. The most recent study makes consistent use of the average travel time in both presentation and analysis.

The German reliability ratios for car travel (constant across journey purposes) are coming out higher relative to the other national studies. The study is, however, rather distinct in the presentation of the reliability and showed percentage of early, on-time, and late arrival, as opposed to the equal proportion of a limited number of possible travel (or arrival times). Moreover, the results for public transport and air travel are not converted to the traditional standard deviation approach to match with existing modelling approaches in Germany and hence these values are not directly comparable (Ehreke et al., 2015). In contrast, the Dutch, Norwegian and UK national VTT surveys make use of largely comparable presentational formats of the SP experiment related to travel time reliability.

| Study                          | Model and data                | Country                | RR                                               |
|--------------------------------|-------------------------------|------------------------|--------------------------------------------------|
| <i>Car</i>                     |                               |                        |                                                  |
| MVA (1996)                     | Logit on SP data              | UK                     | 0.36-0.78                                        |
| Copley et al. (2002)           | Logit on SP data              | UK                     | Pilot survey: 1.3                                |
| Hensher (2007)                 | Logit on SP data              | Australia              | 0.3-0.4                                          |
| Eliasson (2004)                | Logit on SP data              | Sweden                 | 0.30-0.95                                        |
| Mahmassani (2011)              | Logit on SP data              | USA                    | NCHRP 431: 0.80-1.10<br>SHRP 2 CO4: 0.40-0.90    |
| <i>Expert workshop of 2004</i> | Expert opinion                | <i>The Netherlands</i> | 0.8                                              |
| Significance et al. (2013)     | Latent class model on SP data | The Netherlands        | Commuting: 0.4<br>Business: 1.1<br>Other: 0.6    |
| <i>Train</i>                   |                               |                        |                                                  |
| ATOC (2002)                    | Logit on SP data              | UK                     | is 0.6-1.5                                       |
| Ramjerdi et al. (2010)         | Logit on SP data              | Norway                 | Short trips: 0.69<br>Long trips: 0.54            |
| <i>Expert workshop of 2004</i> | Expert opinion                | <i>The Netherlands</i> | 1.4                                              |
| Significance et al. (2013)     | Latent class model on SP data | The Netherlands        | Commuting: 0.4<br>Business: 1.1<br>Other: 0.6    |
| <i>Bus/tram/metro</i>          |                               |                        |                                                  |
| MVA (2000)                     | Logit on SP data              | France                 | 0.24                                             |
| Ramjerdi et al. (2010)         | Logit on SP data              | Norway                 | Short trips: 0.69<br>Long trips: 0.42            |
| <i>Expert workshop of 2004</i> | Expert opinion                | <i>The Netherlands</i> | 1.4                                              |
| Significance et al. (2013)     | Latent class model on SP data | The Netherlands        | Commuting: 0.4<br>Business: 1.1<br>Other: 0.6    |
| <i>Air</i>                     |                               |                        |                                                  |
| Ramjerdi et al. (2010)         | Logit on SP data              | Norway                 | 0.20                                             |
| Significance et al. (2013)     | Latent class model on SP data | The Netherlands        | Business: 0.7<br>Other: 0.7                      |
| <i>Road freight</i>            |                               |                        |                                                  |
| Fowkes (2007)                  | Logit on SP data              | UK                     | Shippers: 0.38<br>Own-account: 0.19              |
| Halse et al. (2010)            | Logit on SP data              | Norway                 | Shippers: 1.2<br>Carriers: 0<br>Overall: 0.11    |
| Significance et al. (2013)     |                               | The Netherlands        | Shippers: 0.9<br>Carriers: 0.28<br>Overall: 0.37 |

**Table 1** Summary of the empirical findings on the reliability ratio (the value of the standard deviation of travel time versus travel time).  
*Figure 20: Empirical findings of the reliability ratio. Source: de Jong and Bliemer (2015)*

| Mode             | Purpose  | This study | Norway 2018 | Norway 2010                | UK 2014 |
|------------------|----------|------------|-------------|----------------------------|---------|
| Car              | Commute  | 0.27       | 0.4         | 0.25 (long) - 0.42 (short) | 0.33    |
|                  | Business | 0.21       | 0.4         |                            | 0.42*   |
|                  | Other    | 0.35       | 0.4         |                            | 0.35    |
| Train            | Commute  | 0.32       | 0.4         | 0.54 (long)                | -       |
|                  | Business | 0.11       | 0.4         |                            | -       |
|                  | Other    | 0.27       | 0.4         |                            | -       |
| Bus, tram, metro | Commute  | 0.65       | 0.4         | 0.69 (short)               | -       |
|                  | Business | 0.61       | 0.4         |                            | -       |
|                  | Other    | 0.56       | 0.4         |                            | -       |
| Air              | Business | 0.30       | -           | 0.20 (long)                | -       |
|                  | Other    | 0.28       | -           |                            | -       |
| Recr. Navigation | Other    | 0          | -           | -                          | -       |

\*only applies to employee part of the VTT

**Table 50** Comparison of reliability ratios between the 2022 study and national studies in Norway and the UK  
*Figure 21: Reliability ratio across different national VTT studies. Source: Significance (2023)*

| Mode | Attribute | Value by Purpose |      |          |         |                 | All |
|------|-----------|------------------|------|----------|---------|-----------------|-----|
|      |           | Education        | Work | Shopping | Leisure | Business Travel |     |

|     |                                         |      |      |      |       |        |        |
|-----|-----------------------------------------|------|------|------|-------|--------|--------|
| Car | Standard deviation                      | 3.21 | 3.45 | 3.51 | 3.09  | 6.54   | 3.61   |
| Car | VOT (€/h)                               | 3.90 | 4.87 | 4.29 | 4.03  | 8.38   | 4.66   |
| Car | Standard deviation/VOT                  | 0.7  | 0.7  | 0.7  | 0.7   | 0.7    | 0.7    |
| PT  | Mean expected unscheduled delay         | 4.66 | 5.10 | 4.28 | 4.82  | 15.97  | 5.48   |
| PT  | Mean expected unscheduled early arrival | 1.81 | 1.98 | 1.67 | 1.88  | 6.22   | 2.13   |
| PT  | VOT (€/h)                               | 4.39 | 4.47 | 5.11 | 4.35  | 7.01   | 4.83   |
| PT  | RR: VOR_late/VOT                        | 0.9  | 1.0  | 0.7  | 0.9   | 1.7    | 0.9    |
| PT  | RR: VOR_early/VOT                       | 0.3  | 0.4  | 0.3  | 0.3   | 0.7    | 0.4    |
| Air | Mean expected unscheduled delay         | na   | na   | na   | 38.44 | 51.27  | 46.60  |
| Air | Mean expected unscheduled early arrival | na   | na   | na   | 90.16 | 120.25 | 109.30 |
| Air | VOT (€/h)                               | na   | na   | na   | 25.45 | 38.76  | 33.67  |
| Air | VOR_late/VOT                            | na   | na   | na   | 1.4   | 1.4    | 1.4    |
| Air | VOR_early/VOT                           | na   | na   | na   | 3.3   | 3.2    | 3.2    |

Note: RR=reliability ratio (population weighted); PT=public transport; na=not applicable

**Table 2** VOR, VOT, and Reliability Ratio

*Figure 22: Results from the German national VTT study. Source: Ehreke et al. (2015)*

Moylan et al. (2022) provide an interesting analysis showing that using the standard mean-standard deviation presentation of travel time unreliability provides consistent answers with the existing literature for the reliability ratio. However, they also find a strong preference for reducing the probability of extreme events suggesting additional metrics than the standard deviation are required (e.g. to deal with the phenomenon of ‘resilience’). Unfortunately, clearer definitions of metrics capturing such events need defining. Zang et al. (2024) build on the idea of the importance of reducing the probability of extreme events and elicit travellers WTP for reducing the unexpected delay. Although these are interesting areas of academic research, these methods need maturing before they should be considered in the context of national VTT studies.

### 6.2.2 Modelling of travel time reliability in the UK and beyond

de Jong and Bliemer (2015) also discuss the inclusion of reliability in transport forecasting models. Namely, there is a concern that including reliability only in the appraisal (as per current practice in the UK and many other countries) will bias the estimates of benefits (as has been demonstrated for congestion multipliers) unless reliability is also included in modelling exercises.

In terms of the modelling of travel time unreliability, de Jong and Bliemer (2015) propose the use of three distinct methods:

- Method 1: no change to transport models but its outputs in terms of road travel times between origins and destination will be used to calculate the standard deviation of travel time for the reference case (without the infrastructure project) and the policy case (with the project).
- Method 2: reliability will also be included in the mode choice and route choice components of the transport models. This takes more effort and is expected to take three-five years to implement according to the authors.
- Method 3: Monte Carlo simulation will give variations in demand and capacity that lead to differences in travel time, from which one can calculate the average and standard deviation, for use in the cost–benefit analysis. According to de Jong and Bliemer (2015), this is the ideal long-run solution but is more computationally demanding than the first two solutions.

The choice for the use of the standard deviation as the preferred metric of travel time reliability is made by de Jong and Bliemer (2015) based on interviews with international experts on the VTT and VTTR.

In the context of the UK, reliability impacts are often monetised and included in the economic appraisal, following advice in TAG Unit A1.3 (section 6), but we are not aware of reliability being used

in any behavioural models that support ‘real life’ appraisal. For highway schemes the most common approaches are (a) skim time and distance from the assignment and apply the “urban roads” method (section 6.3 of TAG A1.3), or (b) use National Highways MyRIAD spreadsheet (which can only be applied to certain types of schemes). Both methods use the standard deviation of travel time as the reliability metric and the TAG-recommended reliability ratio of 0.4 (i.e. 1 minute of SD has the same value as 0.4 minutes of mean travel time).

WSP and Andrew Gordon Consulting Limited are currently doing research for National Highways to include reliability in traffic assignment models (not currently in the public domain). This has parallels with the earlier referred work on modelling congestion.

One of the motivations for the work was that more people use the M6 Toll in the West Midlands (in preference to the old M6) than would be implied by standard values of time, implying there may be additional benefits of using the toll road, such as better reliability.

Most of the problems for traffic assignment models stem from the metric of reliability being the standard deviation (SD) of travel time. Specifically, the fact that it is not link-additive, i.e. you cannot just sum the SD for each link in a path to get the total path SD – a point which is also highlighted by de Jong and Bliemer (2015). This is a problem for the minimum cost path finding algorithms in commercial traffic assignment software (SATURN, Visum, etc.), which all assume costs are link-additive. Alternative algorithms exist which do not assume link-additivity, but they are extremely slow in comparison and are not practical for commercial applications.

The ideal reliability metric would therefore be link-additive. One strand of the research for National Highways was to look at alternative reliability metrics. Several exist in the literature, for example ‘buffer time’ which is defined as the difference between the 95%ile time and the mean, or ‘mean excess travel time’ which is defined as the average delay (relative to mean) when travel time > mean. However, none of the metrics identified is link-additive. de Jong and Bliemer (2015) on the one hand argue that using the variance may solve the issue of link additivity, but on the other hand reject this metric because in congested networks there is dependence across the links as congestion spreads backwards. In other words, variance is only link-additive if we assume link travel times are independent of each other (zero correlation), which is not the case.

As such, it appears that the ideal link-additive reliability metric may well not exist.

Supply-side modelling (i.e. predicting reliability as a function of flow and/or mean time) seems to have been almost entirely focused on SD as the metric, including the urban road function in TAG<sup>23</sup>. Any move to a different metric would therefore require a new programme of research to develop supply-side functions for the new metric. de Jong and Bliemer (2015)’s conclusions align with this argument and argue that an SD approach is consistent with current practice in Australia and New Zealand.

In summary, there does not appear to be an alternative reliability metric that overcomes the modelling difficulties presented by SD, and most supply-side modelling is based on predicting the SD. There is therefore no compelling reason to use an alternative metric and the working assumption should be that the next VTT should estimate VTTR based on the SD of travel time.

The above focuses on highway modelling. For consistency, we also need to make sure that reliability for other modes is included in modelling and appraisal, even if the metric is different (for PT TAG uses the lateness of arrival time relative to the timetable, and this metric is of particular importance in the

---

<sup>23</sup> In some cases, the functions predict the coefficient of variation, which is just the SD of journey time divided by the mean,

railway industry, notably as an input to the Schedule 4 performance regime for planned disruption<sup>24</sup>). Soza-Parra et al. (2022), for example, make use of smartcard data and GPS information to characterise headway and travel time reliability on public transport in Santiago (Chile), and they model the impact of reliability on mode choice, but acknowledge that the need for considering the impact of service reliability on route choice and passenger assignment models.

### 6.3 Interaction between congestion and reliability

The 2014/15 study, and most national VTT studies present separate SP games for reliability and congestion. However, they are strongly correlated (i.e. the SD of journey times increases with congestion) and there is a concern that using both the congested VTT *and* VTTR would lead to double counting. This is one of the explanations why reliability benefits are commonly adopted in appraisal practices, but not congestion effects (see paragraph 4.4.3 in TAG Unit A1.3, Department for Transport, 2022a). Ideally a new VTT study would isolate reliability and congestion effects and remove the confounding in estimation and double counting in appraisal.

The key challenge is to isolate the components comprised in the CVTT and the VTTR. That is, part of disutility associated with spending time in congestion is associated with discomfort due to increased levels of attention interacting with other traffic, and (or) the discomfort of not having a feeling of progress by not moving at desired speeds or not moving at all. At the same time, the risk of arriving late may also contribute to that level of discomfort and be included in the current congestion multipliers.

Isolating these components is a challenging task because some of the negative impacts from anticipated delay and variability in travel time may already be circumvented by changing, for example, departure times or route choices. In-trip discomfort can therefore perhaps be better associated with unanticipated delay, or unexpected delay, as captured by Zang et al. (2024). Ojeda-Cabral et al. (2021), further add to the discussion, by noting that in the context of public transport adding recovery time to the rail timetable may increase travel time reliability but this comes at a welfare premium due to the frustration caused by standing still outside the station or travelling slower as part of the regular timetable.

ITS (2020) briefly touches upon the confounding issue between congestion and reliability. In Section 3.3 of their report, it is suggested that congestion multipliers can stand as a proxy for ‘regular’ conditions, and that the scope of reliability benefits can be reduced to the assessment of resilience measures affecting the right-hand tail of the journey time distribution. Although this may seem a feasible solution, parts of the impact of regular reliability benefits on scheduling decisions may not be fully captured.

The workstream on CVTT emerged from the ARUP et al. (2015a) results and recommendations, and the studies discussed in **Section 6.1** highlight progress has been made in this dimension. Similarly, the referred work by NH in the context of reliability is making progress on the modelling side. It seems more than natural that the next step in this area is the commissioning of a sequence of studies on the overlap between congestion and reliability both in terms of valuation and modelling.

### 6.4 Crowding effects

Whilst there are significant challenges in separating congestion and reliability effects, the discomfort of travelling in crowded conditions on public transport is easier to separate out. It should be noted that some of the stress related to potentially not being able to make your next train/bus may still partly

---

<sup>24</sup> The Schedule 8 performance regime for unplanned disruption uses a different approach.

be captured by the studies reviewed below (this is distinct from the inconvenience per se of having to interchange, which is conventionally captured through interchange penalties). Across the different national VTT studies a consistent approach of the SP format is applied. That is, alongside regular travel time a distinction is made as to whether passengers can sit or must stand during their journey, and a level of vehicle occupancy is provided. The level of crowding is typically denoted by the number of passengers standing per m<sup>2</sup> or the load factor. **Figure 23** and **Figure 24** respectively provide an example of the presentation of the crowding SP in ARUP et al. (2015a), Flügel et al. (2020) and Significance (2023).

**Figure 2.9: Time vs. cost vs. quality experiment for rail**

|                                 | Option A                                                                                                                                                     | Option B                                                                                                                                                 |
|---------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| One way travel time             | 3 hours 54 minutes                                                                                                                                           | 3 hours 18 minutes                                                                                                                                       |
| One way travel fare             | £18.00                                                                                                                                                       | £24.00                                                                                                                                                   |
| Crowding level when you boarded |  <p>Seated, 100% of seats occupied, eight people stood around each door</p> |  <p>Standing, 100% of seats occupied, one person stood around door</p> |
|                                 | Option A                                                                                                                                                     | Option B                                                                                                                                                 |

Figure 23: Presentation of the crowding levels in public transport in ARUP et al. (2015a).

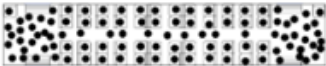
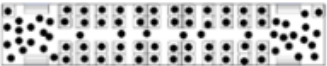
| Treinreis A                                                                                                              | Treinreis B                                                                                                              |
|--------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
| Reistijd in de trein :<br><b>40 min.</b>                                                                                 | Reistijd in de trein :<br><b>33 min.</b>                                                                                 |
| Drukke in de trein :<br><b>100% van de zitplaatsen bezet, er staan overal personen ( 3 personen per vierkante meter)</b> | Drukke in de trein :<br><b>100% van de zitplaatsen bezet, er staan overal personen ( 2 personen per vierkante meter)</b> |
|                                       |                                      |
| Zitten of staan?<br><b>U kunt zitten</b>                                                                                 | Zitten of staan?<br><b>U moet staan</b>                                                                                  |
| Frequentie<br><b>1 trein elke 60 minuten (1 per uur)</b>                                                                 | Frequentie<br><b>1 trein elke 30 minuten (2 per uur)</b>                                                                 |
| Kosten:<br><b>€ 7.20</b>                                                                                                 | Kosten:<br><b>€ 14.20</b>                                                                                                |

Figure 10 - Example of an SP4A choice card (in Dutch, see Appendix A for attribute names in English)

Figure 24: Presentation of the crowding levels in public transport in Significance (2023).

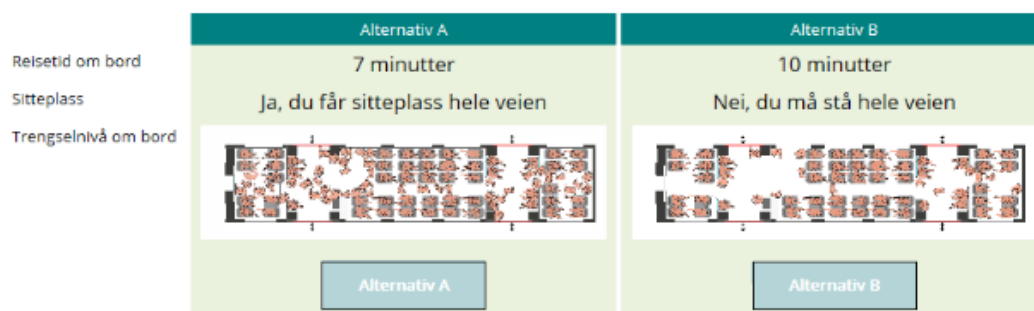


Figure E2. Example of choice task used to estimate values of in-vehicle crowding in public transport.

Figure 25: Presentation of the crowding levels in public transport in Flügel et al. (2020).

Fedujwar and Agarwal (2024) present a systematic overview of 40 academic papers on valuing crowding in public transport. These papers do not include the national VTT studies. **Figure 26** summarises the evidence on standing multipliers, clearly indicating that travellers prefer to sit as opposed to stand. The standing multiplier, like the congestion multipliers, denotes the ratio between the VTT in standing and sitting conditions. A value larger than one indicates that travellers experience more discomfort when standing than when sitting. Most estimates of the standing multiplier are between 1 and 2, but not unexpectedly more variation is observed across studies. A similar overview graph is unfortunately not available for the so-called crowding multipliers. The crowding multipliers denote the extent to which the disutility of travel time goes up as crowding levels, for example the number of passengers standing per  $m^2$ , increases and thereby increase the VTT. Rossetti and Daziano (2024) do provide an overview of crowding multipliers in **Table 28**. The evidence reveals that the size of the crowding multiplier is consistently higher when standing than when being seated as per expectation.

**Table 29** summarises the crowding multipliers reported in the three national VTT studies which included crowding as part of their study design. This was not the case for the German and Singapore studies. Comparing the crowding multipliers for commute and business across the UK and Dutch study we observe a close correspondence for the standing multiplier, particularly towards higher levels of crowding. This indicates that the Dutch have a higher standing multiplier but slightly lower marginal disutility from additional units of travel. Relative to the Norwegian study we observe that the standing multiplier is comparable but that again the UK study observes a higher marginal disutility of standing in increasingly crowded conditions. The business crowding multipliers in Norway appear to be somewhat on the low end, especially when one considers that crowding may impede the ability to work productively in the course of travel. Overall, these numbers are of the same order of magnitude and do not indicate a cause for concern.

The Dutch study signals that crowding multipliers in local public transport are lower than in rail travel, which is most likely due to the shorter distances people travel in such conditions, and the potential for using their time productively on these kind of journeys. The values can, however, not be compared against the results from the UK study because the crowding multipliers estimated for non-rail journeys were using categorical statements as opposed to  $pax/m^2$ . However, the crowding multipliers for 'densely packed' in other PT are comparable to the Dutch outcomes at 4  $pax/m^2$  (see table 7.13 in ARUP et al., 2015a). For consistency purposes this could be improved in future design of the SP experiments.



Flügel and Hulleberg (2023) show that after an increase in the crowding multipliers during the Covid-19 pandemic, the crowding multipliers in Norway have returned in November 2022 back within the confidence bounds of the pre-pandemic values of Nov 2018.

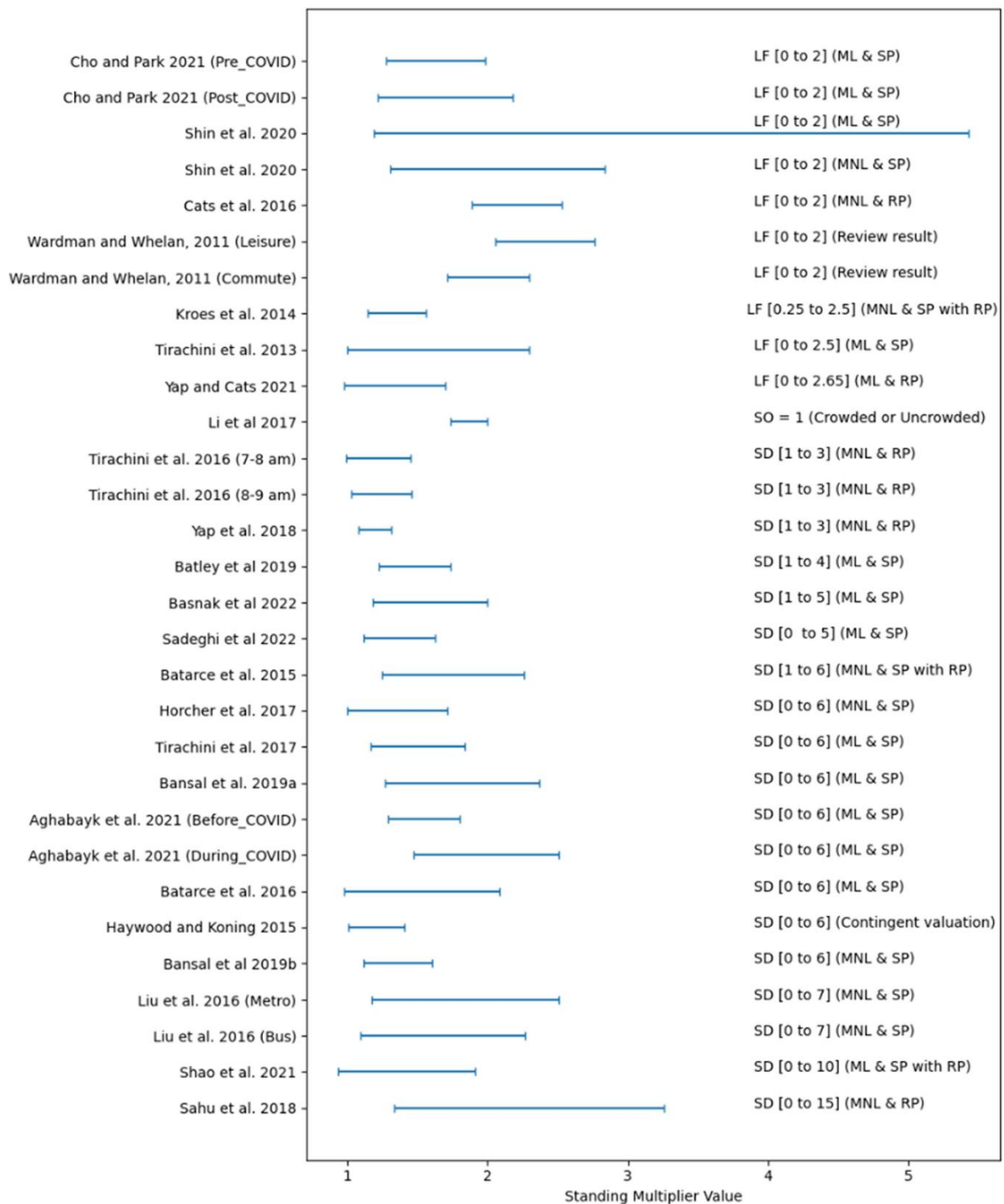


Figure 26: Overview of standing multipliers as reported by Fedujwar and Agarwal (2024) (Fig 3).

Table 28: Summary table from Rosetti and Daziano (2024)

| Reference                 | City or country    | Data type | Crowding                       | CM, sitting | CM, standing |
|---------------------------|--------------------|-----------|--------------------------------|-------------|--------------|
| Hörcher et al. (2017)     | Honk Kong          | RP        | 6 pax/m <sup>2</sup>           | 1.72        | 1.98         |
| Yap et al. (2020)         | The Hague, NL      | RP        | 3 pax/m <sup>2</sup>           |             | 1.34         |
| Bansal et al. (2022)      | Unnamed Asian city | RP        | More than 2 pax/m <sup>2</sup> |             | 1.47         |
| Kroes et al. (2014)       | Île-de-France, FR  | SP        | 8 pax/m <sup>2</sup>           | 1.41        | 1.6          |
| Björklund et al. (2015)   | Sweden             | SP        | 8 pax/m <sup>2</sup>           | 1.5         | 2.13         |
| Batarce et al. (2016)     | Santiago, CL       | SP        | 6 pax/m <sup>2</sup>           |             | 2.51         |
| Tirachini et al. (2017)   | Santiago, CL       | SP        | 6 pax/m <sup>2</sup>           | 1.60–1.67a  | 1.98–2.00a   |
| Sahu et al. (2018)        | Mumbai, IN         | SP        | 7–9 pax/m <sup>2</sup>         |             | 1.94–2.26b   |
| Sahu et al. (2018)        | Mumbai, IN         | SP        | 13–15 pax/m <sup>2</sup>       |             | 3.82b        |
| Bansal et al. (2019)      | New York City, US  | SP        | 6 pax/m <sup>2</sup>           | 2.13–3.16a  | 2.65–5.81a   |
| Shin et al. (2021)        | Seoul, KR          | SP        | 6 pax/m <sup>2</sup>           | 2.15        | 3.22         |
| Sadeghi et al. (2022)     | Ithaca, NY, US     | SP        | 5 pax/m <sup>2</sup>           |             | 1.75         |
| Wardman and Whelan (2011) | United Kingdom     | Meta      | 150% load factor               | 1.4         | 2.5          |

a Depending on type of model used.

b Multipliers for standing times under 10 min and between 10 and 20 min respectively.

Table 29: Summary of crowding multipliers across the UK, Dutch and Norwegian national VTT studies

|                        | UK (2014) |       |          | NL (2022) |          | Norway (2018)*     |          |
|------------------------|-----------|-------|----------|-----------|----------|--------------------|----------|
|                        | Rail      |       |          | Rail      | Local PT | All PT             |          |
|                        | Commute   | EB    | Other NW | All       | All      | Commute + other NW | Business |
| <b>Seated</b>          |           |       |          |           |          |                    |          |
| 50% load               | 1         | 1     | 1        | 1         | 1        | 1                  | 1        |
| 75% load               | 1.084     | 1.013 | 1        | 1         | 1        | 1                  | 1        |
| 100% load              | 1.378     | 1.326 | 1.396    | 1         | 1        | 1.22               | 1.05     |
| 1 pax/m <sup>2</sup>   | 1.499     | 1.503 | 1.587    | 1.18      | 1.04     | 1.3                | 1.08     |
| 3 pax/m <sup>2</sup>   | 1.806     | 1.797 | 1.947    | 1.36      | 1.12     | 1.45               | 1.1      |
| <b>Standing</b>        |           |       |          |           |          |                    |          |
| 0.5 pax/m <sup>2</sup> | 1.595     | 1.705 | 1.693    |           |          |                    |          |
| 1 pax/m <sup>2</sup>   | 1.646     | 1.835 | 1.767    | 1.93      | 1.27     | 1.65               | 1.65     |
| 2 pax/m <sup>2</sup>   | 1.813     | 2.073 | 2.189    | 2.11      | 1.43     | 1.78               | 1.68     |
| 3 pax/m <sup>2</sup>   | 2.168     | 2.133 | 2.494    | 2.29      | 1.58     | 1.92               | 1.72     |
| 4 pax/m <sup>2</sup>   | 2.563     | 2.689 | 3.026    | 2.47      | 1.73     | 2.05               | 1.75     |

\* values approximated based on graph E3 in Flügel et al. (2020)

With respect to the use of crowding multipliers in appraisal, TAG Unit M3.2 discusses that its inclusion in transport assignment models has significant practical implications. Namely, during assignment generalised costs will change due to increasing or reducing crowding levels. This requires to iteratively evaluate the assignment model until convergence which directly influences run times. For these reasons “crowding should only be modelled where it is likely to have a significant effect on traveller behaviour or where one of the objectives of the scheme is associated with crowding.” (Department for Transport, 2024k)

To adjust for crowding levels in generalised costs so-called in-vehicle weights are used. Different weights are recommended across regions and for regional and inter-urban journeys. Although these recommended values make use of the same metric, i.e. pax/m<sup>2</sup>, the evidence is, however, not the same as results reported in ARUP et al. (2015a). The evidence is based on a separate study commissioned by the DfT and the Passenger Demand Forecasting Council (PDFC) (Department for Transport, 2024h). For example, for London and the South East, the reported in-vehicle weights are somewhat lower than those identified by ARUP et al. (2015a). We do not believe that PDFC has subsequently re-surveyed these weights, and it will therefore be advisable for a new VTT study to include crowding within its remit – especially as crowding levels and attitudes towards crowding may well have changed following the Pandemic. This will also ensure that up-to-date weights are estimated in a consistent fashion across all relevant public transport modes.

### 6.5 Conclusions in relation to congestion, reliability, and crowding

Whilst ARUP et al. (2015a) covered all three aspects of congestion, reliability and crowding, only the reliability results were eventually implemented in TAG Unit A.1.3 (Department for Transport, 2022a). In relation to congestion and crowding, TAG Unit A.1.3 is very explicit in stating that “*more research is needed before reliable values can be derived on a robust basis*” for congestion and crowding. For rail appraisal this means that the recommendations set out in PDFH for crowding should be used, but only when the scheme is aimed at reducing crowding or where the crowding impacts are likely to be large, whereas for other modes congestion and crowding should not be implemented.

Starting with the treatment of reliability, since this is presently represented in TAG, but only in an adjusted BCR, the SP presentational format adopted in ARUP et al. (2015a) is largely comparable with those presented in the most recent national VTT studies in Norway and the Netherlands. Moreover, the estimated reliability ratios across these three studies are in the same order of magnitude. The primary difference which emerges across these three studies is, however, in the level of segmentation. Where for the UK study reliability ratios are only estimated for car journeys but segment by journey purpose, Norway adopts a uniform reliability ratio across modes and journey purposes, and the Netherlands segments the reliability ratio by mode and journey purpose. In a future study, the desired level of segmentation can be revisited easily by contrasting different model specifications with additional parameters.

Both in the context of valuation and modelling research has been conducted into the use of alternative metrics moving beyond the currently adopted mean-standard deviation approach. Especially, in the context of transport modelling the use of the standard deviation of travel time is challenging because the metric is not link-additive. Alternative metrics are, however, not available and may well not exist. As such, there is no compelling argument to change the present approach. Naturally, communicating the notion of the standard deviation in travel time in the SP exercise requires ongoing attention.

In sharp contrast are the conclusions with respect to congestion. The follow-up work on CVTT following the 2014/15 study clearly highlights that the adopted categorical approach to congestion in SP3 together with the unspecified conditions of travel in SP1 made their implementation problematic both in terms of valuation and modelling. The subsequent research conducted on modelling recommends the alternative use of ‘free flow’ and ‘delay’ metrics. A change to these metrics, however, highlights that congestion and reliability are not independent. Altogether, it is recommended that this valuation and modelling research agenda needs to be continued before robust evidence on congestion can change the present status quo on the matter in TAG Unit A.1.3. Recent empirical evidence from the wider literature and other national VTT studies do not make progress on the continuous measurement of congestion nor on its interaction with reliability.

For crowding a more positive story can be told. The approach adopted in ARUP et al. (2015a) for rail is consistent with other international VTT studies and the wider literature. The SP design captures the relevant aspects of crowding in term of a continuous density (pax/m<sup>2</sup>) and accounts for passengers travelling in sitting and standing conditions. Moreover, the estimated crowding multipliers are in the same order of magnitude relative to the international literature. With respect to other forms of public transport improvements in the design can be made by moving away from the categorical levels of crowding used in the previous study, alike the case for congestion. We are not aware of any subsequent updates made to the evidence base for PDFH, and together with empirical evidence from Norway that crowding multipliers have returned to pre-pandemic levels, see a good case for obtaining robust evidence on crowding in a next VTT study with minimal changes in approach.

A similar interaction between crowding and reliability may exist as between congestion and reliability. We consider this interaction the be smaller in magnitude and of less concern in public transport than for car journeys. Nevertheless, if additional research is commissioned into congestion and reliability, the interaction between crowding and reliability should be considered within scope.

## 7. Active and emerging modes of travel

Estimating the VTT for active modes of travel (e.g. walking and cycling) is conceptually the same to estimating the VTT for any other mode of transport. That is, the relevant metrics provide information on the shadow prices of time and costs. As with any other transport mode the (dis)utility of travel time is associated with its opportunity costs and the level of comfort experienced when spending time walking or cycling. Within the population of interest there will be a degree of heterogeneity associated with the (dis)utility of travel time. Namely, some people may particularly enjoy the physical challenge and health benefits of walking and cycling and opt for longer journeys where the comfort factor may outweigh the opportunity costs and hence result in a negative VTT. In the same vein, contextual factors like the quality and safety of the provided transport infrastructure, and other environmental factors such as the weather and hilliness of the terrain may affect the level of comfort of the journey. Alike with other modes of transport, the purpose of the journey (commute or leisure) and the role of the active travel in the overall journey (e.g. main mode or access/egress mode) can all be influential on the importance of saving time.

Despite its similarities with other modes of transport, the primary challenge experienced in valuing travel time in relation to active modes of travel is the introduction of relevant cost (or income) trade-offs allowing to translate the above referred marginal utility of travel time by active modes in a monetary valuation. The root cause for this challenge is that travelling on foot or by bicycle (assuming ownership) is generally considered to be free. Together with the public good nature of urban transport infrastructure it is typically considered to be unrealistic for users of active modes of travel to pay more for routes saving travel time or improving safety. Accordingly, adopting the traditional within-mode SP1-based time vs money route choice trade-off has proven to be challenging and alternative survey formats have accordingly been adopted in the wider literature.

### 7.1 Survey design in ARUP et al. (2015b)

In the 2014/15 VTT study walk and cycling as active modes of travel were commissioned as 'optional modes', and its results were published in a separate report (ARUP et al., 2015a). The dearth of evidence on the VTT for active modes of transport in general gave the corresponding study efforts the label of 'exploratory research'. Moreover, the research on the topic proved challenging.

ARUP et al. (2015b) provide a full report on the selection and testing of different design options, with Section 8.3 describing the final design formats that were adopted for walking and cycling. SP1, as displayed in **Figure 27** for cycling, describes the introduction of a cycle way for which users can pay on a trip-by-trip basis, which is then contrasted against the conditions of the existing route. To arrive at this design, two earlier design versions were discarded after extensive testing. The original specification involved a hypothetical toll bridge reducing travel time. This was discarded based on in-depth interviews and revised into mode choice and cycle hire setting which focus groups revealed as not being ideal.

|                     | Existing Route                                                                                                                                               | Cycleway                                                                        |
|---------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|
| <b>Travel time</b>  | #x# minutes in conditions where there is little motor traffic<br>#x# minutes in conditions where there is a lot of motor traffic<br>#x# minutes at junctions | #x# minutes to Cycleway<br>#y# minutes on Cycleway<br>#z# minutes from Cycleway |
| <b>One way cost</b> |                                                                                                                                                              | £#Y#                                                                            |

Figure 27: Final design of SP1 for cycling in ARUP et al. (2015b)

For walking SP1 centred around egress times to car parking options for car users, and egress times for bus from the bus stop to the destination. Additionally, in-vehicle time on the main mode of transport were presented alongside respectively parking costs or bus fares as the monetary attribute (see **Figure 28** and **Figure 29**).

|                                                          | Car Park A  | Car Park B  |
|----------------------------------------------------------|-------------|-------------|
| <b>One way walking time from car park to destination</b> | #X# minutes | #Y# minutes |
| <b>Time in Car</b>                                       | #A# minutes | #B# minutes |
| <b>Parking cost</b>                                      | £#X#        | £#Y#        |

Figure 28: Final design of SP1 walk – parking choice ARUP et al. (2015b)

|                                                  | Bus journey A | Bus journey B |
|--------------------------------------------------|---------------|---------------|
| <b>Walking time from bus stop to destination</b> | #X# minutes   | #Y# minutes   |
| <b>Time on Bus</b>                               | #A# minutes   | #B# minutes   |
| <b>Bus fare</b>                                  | £#X#          | £#Y#          |

Figure 29: Final design of SP1 walk – bus stop choice in ARUP et al. (2015b)

The SP2 design did not change from its original format and presents for both walking and cycling the time spent in different travel conditions (amount of traffic and type of walking/cycle infrastructure). **Figure 30** and **Figure 31** presented the respective design formats adopted. Note that SP2 did not include a monetary component and only allows identification of the (dis)comfort of travelling in different travel conditions.

|                                     | Route A                                                                                                                                            | Route B                                                                                                                                  |
|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| Additional time spent at junctions  | 8 short waits (15 seconds) at junctions<br>4 long waits (45 seconds) at junctions                                                                  | 2 short waits (15 seconds) at junctions<br>6 long waits (45 seconds) at junctions                                                        |
| Time in different cycling condition |  30 mph road, lot of traffic, good quality cycle lane: 11 minutes |  20 mph road, little traffic, no facilities: 9 minutes |
|                                     |  Lot of traffic but segregated by barrier: 7 minutes              |  Lot of traffic but segregated by barrier: 3 minutes   |
|                                     |  Lot of traffic on shared use footpath: 5 minutes                 |  Cycleway away from roads: 11 minutes                  |
|                                     | Route A                                                                                                                                            | Route B                                                                                                                                  |

Figure 30: Final presentation of the SP2 design for cycling in ARUP et al. (2015b)

With sample sizes of around 200 respondents each for the walking and cycling experiment, covering both commuting and other non-work purposes, sample sizes are significantly smaller than those collected for the main motorised modes of transport as discussed in **Section 2.2**. The modelling results are therefore presented with significantly lower degrees of confidence and no national VTT measures for walking and cycling are reported.

Instead, ARUP et al. (2015b) provide five recommendations. Firstly, because the SP1 format for walking worked well, they recommend rolling this out for a full field survey. Second, for SP1 in the context of cycling the format needs to be continuously reviewed and revised to avoid strategic bias against paying for the use of a dedicated cycleway. Third, the format for SP2 needs to be simplified by reducing the number of conditions that any individual is presented with. Fourth, the SP2 exercise needs to be constructed around an actual (reference) trip to improve realism. Fifth and finally, they recommend additional work translating behavioural VTTs for walking and cycling into appraisal values, particularly because there may be issues with the reporting of short walk and cycling trips in the NTS. To address the underreporting of short walk trips, the NTS has changed its approach since 2017 and has started recording these on day 1 of the diary as opposed to day 7 (Department for Transport, 2024e).



|                                               | Route A                                                                                                                                 | Route B                                                                                                                                  |
|-----------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Additional time waiting to cross roads</b> | 2 short waits (15 seconds) to cross the road<br>8 long waits (45 seconds) to cross the road                                             | 10 short waits (15 seconds) to cross the road<br>10 long waits (45 seconds) to cross the road                                            |
| <b>Time in different walking condition</b>    |  Lot of road traffic but lot of pedestrians: 3 minutes |  Pedestrian route but poor quality surface: 11 minutes |
|                                               |  Lot of road traffic and few pedestrians: 11 minutes   |  Little road traffic but lot of pedestrians: 3 minutes |
|                                               |  Little road traffic but lot of pedestrians: 9 minutes |  Footway shared with cyclists: 7 minutes               |
|                                               | Route A                                                                                                                                 | Route B                                                                                                                                  |

Figure 31: Final presentation of the SP2 design for walking in ARUP et al. (2015b)

## 7.2 Active travel in other national VTT studies

**Table 30** presents the recommended VTT values emerging from the Norwegian and Dutch national VTT studies (Flügel et al., 2020, Significance, 2023). The consistent picture which emerges is that – besides the use of much larger sample sizes than in the UK – the VTT for walking is higher than for cycling, and that the VTT for commuting is larger than for other non-work journey purposes. Where in the Dutch study the VTT for cycling trips are of the same order of magnitude for commuting and other non-work trips, in Norway these values are higher for commute journeys, but lower for other non-work journeys by car. The VTT for walking can be considered high relative to the VTT values presented earlier in Table 16. Relative to the existing international literature on active travel related VTT, which is rather small, the obtained Dutch values are much smaller than those obtained in Sweden (Börjesson and Eliasson, 2012, Björklund and Mortazavi, 2013) and only a bit smaller than a previous study in the Netherlands (Ginkel, 2014). The general lack of empirical evidence on the VTT for cycling in the literature is echoed by Tordai et al. (2023). Whilst the German national VTT study, which was also reviewed in **Section 2.4**, did include walking and cycling in its study design, it did not report recommended values to be used in policy appraisal.

Table 30: Recommended national VTT measures for walking and cycling in 2010 GBP per hr

|          | Netherlands |         | Norway   |          |           |           |  |  |
|----------|-------------|---------|----------|----------|-----------|-----------|--|--|
|          | Walking     | Cycling | Walking* | Cycling* | Walking** | Cycling** |  |  |
| Commute  | £ 10.19     | £ 6.52  | £ 30.29  | £ 11.46  | £ 23.47   | £ 10.46   |  |  |
| Business | £ 9.44      | £ 7.19  | £ -      | £ -      | £ -       | £ -       |  |  |
| Other    | £ 7.54      | £ 6.69  | £ 19.83  | £ 6.09   | £ 14.28   | £ 5.82    |  |  |
| All      | £ 7.60      | £ 6.67  | £ 20.74  | £ 10.55  | £ 15.28   | £ 10.28   |  |  |

\* Not controlling for accident risk (deaths/severely injured)

\*\* Controlling for accident risk (deaths/severely injured)



Where the SP1 setting for motorised modes of transport allowed for an effective comparison of recommended VTT values, the heterogeneity in stated preferences methods applied to arrive at the above values for walking and cycling makes this comparison less clean. The UK's approach to opt for a within mode route choice design is distinct from the mode choice approach adopted in the other three national VTT studies.

We briefly summarise the main SP design features used in the three other recent European national VTT studies considered in this report:

1. Germany adopts a mode choice approach where either walking or cycling is contrasted against travelling by public transport and by car. The only attribute used to describe walking and cycling is travel time, implying that active travel is free of any costs. Any estimated cost-coefficient thus comes from costs associated with the alternative modes of transport, respectively public transport and car.
2. Norway adopts two separate SP studies, where the first one is comparable to the German setup. In this mode choice experiment, walking or cycling are contrasted against travelling by car or public transport. Again, only the alternative mode is associated with travel costs. The largest difference emerges by the walking and cycling journeys being characterised by travel time and the main road type (e.g. off-road walk and cycle path), the number of signalled crossings, and the number of other crossings. It is this first SP which is used to derive the VTT estimates. The second SP exercise is a within-mode route choice contrasting varying in travel time, negative effects of physical effort, and positive effects of physical effort. For the cycling setting there is also an attribute describing the use of a regular or electric bicycle.
3. In the Netherlands also two distinct SP designs are adopted. The main SP design a mode choice design is used, but instead of presenting multiple alternative options the use of a regular bicycle (or walking) is contrasted against one other mode of transport respectively being an electric bicycle or a car. The only attributes used here are travel time and travel cost, where the travel cost of using the regular bike (or walking) is as per the other studies free. In the second SP setting a route choice setting is adopted describing each route by means of five attributes (travel time, cycle/footpath configuration including speed of cars, type of paving of cycle path / width of footpath, number of bypassing cars, and the beauty of the route).

The biggest contrast with the UK design is that the VTT in the other national studies is always derived from contrasting the active mode of travel with other modes of transport, as opposed to the use of a within mode route choice for cyclists with an artificially constructed payment vehicle for the use of cycling infrastructure. A similar observation can be made for walking, where in the UK design it is treated as an access or egress mode, whereas in the other national studies it is used as a main mode of transport.

For completeness, **Figure 32 - Figure 37** present examples of the SP formats used in three referred national VTT studies outside of the UK. Differences in the configuration of the SP designs, especially the difference in the number of attributes considered between the Dutch (2) and the Norwegian (5) study muddles the comparison of the VTT estimates. This issue is closely related to the SP1 vs SP2/SP3 discussion presented in **Section 3.2.1** and made prominent by Hess et al. (2020).

| Zu Fuß                                                                           | Öffentlicher Verkehr       | Auto                       |
|----------------------------------------------------------------------------------|----------------------------|----------------------------|
| Gehzeit 1:15 h                                                                   | Gesamtzeit 0:23 h          | Gesamtzeit 0:18 h          |
|                                                                                  | davon Fahrtzeit 0:14 h     | davon fahrend 0:12 h       |
|                                                                                  | davon Wartezeit 0:06 h     | davon im Stau 0:02 h       |
|                                                                                  | davon Fußweg 0:03 h        | davon Fußweg 0:04 h        |
|                                                                                  | Umsteigen 1 Mal            |                            |
|                                                                                  | Kosten 1,80 €              | Kosten 1,70 €              |
|                                                                                  | (29 €/Monat bei 8 Fahrten) | (27 €/Monat bei 8 Fahrten) |
|                                                                                  | Fährt alle 30 min          |                            |
|                                                                                  | Anteil verspätet 10 %      | Anteil verspätet 10 %      |
| Wahl: <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |                            |                            |

Figure 32: German mode choice design contrasting walking against public transport and car (Axhausen et al., 2014)

| Fahrrad                                                                          | Öffentlicher Verkehr      | Auto                      |
|----------------------------------------------------------------------------------|---------------------------|---------------------------|
| Fahrtzeit 0:11 h                                                                 | Gesamtzeit 0:09 h         | Gesamtzeit 0:06 h         |
|                                                                                  | davon Fahrtzeit 0:05 h    | davon fahrend 0:02 h      |
|                                                                                  | davon Wartezeit 0:02 h    | davon im Stau 0:02 h      |
|                                                                                  | davon Fußweg 0:02 h       | davon Fußweg 0:02 h       |
|                                                                                  | Umsteigen 0 Mal           |                           |
|                                                                                  | Kosten ,20 €              | Kosten ,30 €              |
|                                                                                  | (2 €/Monat bei 4 Fahrten) | (2 €/Monat bei 4 Fahrten) |
|                                                                                  | Fährt alle 60 min         |                           |
|                                                                                  | Anteil verspätet 10 %     | Anteil verspätet 10 %     |
| Wahl: <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |                           |                           |

Figure 33: German mode choice design contrasting cycling against public transport and car (Axhausen et al., 2014)

Please choose...

|                                       | Alternative A                | Alternative B      | Alternative C | Alternative D |
|---------------------------------------|------------------------------|--------------------|---------------|---------------|
| Travel mode                           | Bicycle                      | Bicycle            | Car           | Car           |
| Travel time                           | 35 min.                      | 30 min.            | 25 min.       | 32 min.       |
| Travel cost                           |                              |                    | 60 NOK        | 47 NOK        |
| Main road type (cycling)              | Off-road walk and cycle path | On-road cycle lane |               |               |
| Signalised crossings                  | 5                            | 4                  |               |               |
| Other crossings                       | 5                            | 6                  |               |               |
| 1. I prefer the least...              | Alternative A                | Alternative B      | Alternative C | Alternative D |
| 2. I prefer the most...               | Alternative A                | Alternative B      | Alternative C | Alternative D |
| 3. Of the remaining ones, I prefer... | Alternative A                | Alternative B      | Alternative C | Alternative D |

Figure 34: Example of the first SP game for cyclist in the Norwegian VTT study, approach to walking is the same (Flügel et al., 2021)

Please choose...

|                                      | Alternative A                                                                                                                   | Alternative B                                                                                                                  |
|--------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|
| Type of bicycle                      | Regular bicycle                                                                                                                 | Regular bicycle                                                                                                                |
| Travel time on bicycle               | 32 minutes                                                                                                                      | 40 minutes                                                                                                                     |
| Negative effects of physical efforts | Some panting, some sweat, somewhat sore muscles                                                                                 | Minor panting, minor sweat, no sore muscles                                                                                    |
| Positive effects of physical efforts | High number of calories burned, large contribution to future health, large contribution to desirable level of physical activity | Low number of calories burned, minor contribution to future health, minor contribution to desirable level of physical activity |
|                                      | Alternative A                                                                                                                   | Alternative B                                                                                                                  |

Figure 35: Example of the second SP game for cyclists in the Norwegian VTT study (Flügel et al., 2021)

| Elektrische (huur)fiets      | Fiets                       |
|------------------------------|-----------------------------|
| Reistijd:<br><b>18 min.</b>  | Reistijd:<br><b>25 min.</b> |
| Huurkosten:<br><b>€ 0.40</b> | Kosten:<br><b>€ 0.00</b>    |

Figure 36: Example of the SP design for cycling in the Dutch VTT study (Significance, 2023)

| Rit A                                                                                | Rit B                                                                              |
|--------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|
| Route:<br><b>Zeer mooi</b>                                                           | Route:<br><b>Mooi</b>                                                              |
| Verharding:<br><b>Stoeptegels</b>                                                    | Verharding:<br><b>Stoeptegels</b>                                                  |
| Aantal voorbijrijdende auto's:<br><b>Zeer veel</b>                                   | Aantal voorbijrijdende auto's:<br><b>Zeer weinig</b>                               |
| Reistijd:<br><b>28 min.</b>                                                          | Reistijd:<br><b>38 min.</b>                                                        |
| Fietspad:<br><b>Fietsstraat (auto's toegestaan, voorrang voor fietsers, 30 km/u)</b> | Fietspad:<br><b>Vrijliggend fietspad langs een weg (50 km/u)</b>                   |
|     |  |

Figure 37: Example of the Dutch VTT study on quality of cycling infrastructure (Significance, 2023).

In the context of active travel, the description of the travel conditions is potentially more important than for other modes of travel because of perceived levels of safety, and potential health benefits being reflected as an implicit part of elicited VTTs. Significance (2023), for example, reveals that multipliers on the VTT for active travel (relative to the VTT for the generic setting in **Figure 36**) increase when directly interacting with cars on the same road, particularly at higher speeds, and with the quality of the provided infrastructure in terms of quality of the surface, width of the infrastructure and degree of separation of other traffic. This picture is confirmed by the Norwegian VTT study in **Table 31**, where again travelling in unfacilitated conditions is associated with the highest multiplier on travel time.<sup>25</sup>

Interestingly, the Norwegian study proceeds by downscaling the recommended VTT values when controlling for accident risk (deaths/severe injuries). That is they assume that the elicited VTT partly comprises a WTP for reducing exposure to mortality risk, where such WTP information may already be embodied in other benefit metrics in CBA, such as the Value of Statistical Life (or Value of a Prevented Fatality in the vocabulary of HMT – see Section 6.6 of HMT (2024)). To avoid double counting they recommend using the downscaled VTT estimates when accident costs are included separately in the cost-benefit analysis. Although this is a valid argument, the question emerges regarding the extent to which this is also required in the context of other modes of transport.

Flügel et al. (2021) go even a step further and argue that the VTT for active travel should be scaled down to control for the internalisation of potential health benefits of physical effort (i.e. this is distinct from the safety benefits discussed in the preceding paragraph). Again, this can be positioned in the context of avoiding double counting, as benefits from physical activity are also (partly) accounted for in the Department for Transport (2022b)'s guidance on active mode appraisal through the WHO's Health Economic Assessment Tool (HEAT) (WHO, 2024). Whilst Flügel et al. (2021) find reductions in the VTT for cycling between 20-65% after accounting for, largely short term, health benefits, it should

<sup>25</sup> For the Norwegian case the multipliers are calculated from Tables E4 and E5 in Flügel et al. (2020)

be recognised that it is very hard (or perhaps nearly impossible) to determine the extent to which perceived health benefits at the trip level can be separated from long term health benefits and mode choices in general. As such, whilst we recognise that there could be additional wellbeing effects of active mode use, it seems there is little ground yet to isolate such health effects from the overall VTT for active modes and clearly additional conceptual and empirical research is needed in this area.

| Cycling path configuration |                                                           |                                                                                     |            | Walking path configuration                                                  |                                                                                       |            |  |
|----------------------------|-----------------------------------------------------------|-------------------------------------------------------------------------------------|------------|-----------------------------------------------------------------------------|---------------------------------------------------------------------------------------|------------|--|
|                            | Description                                               |    | Multiplier | Description                                                                 |    | Multiplier |  |
| 1                          | Car road – bikes allowed (30 km/h)                        |    | 1.085      | Walking on road with cars and bikes (30 km/h)                               |    | 1.397      |  |
| 2                          | Car road – bikes allowed (50 km/h)                        |    | 1.233      | Walking on road with cars and bikes (50 km/h)                               |    | 1.598      |  |
| 3                          | Bike street – cars allowed, bikes have priority (30 km/h) |    | 0.956      | Sidewalk directly next to road with cars and bikes (30 km/h)                |    | 1.100      |  |
| 4                          | Bike lane in the road (30 km/h)                           |    | 0.993      | Sidewalk directly next to road with cars and bikes (50 km/h)                |    | 1.192      |  |
| 5                          | Bike lane in the road (50 km/h)                           |  | 1.034      | Sidewalk at 2 metres from road with cars and bikes (30 km/h)                |  | 1.005      |  |
| 6                          | Bike lane next to the road (50 km/h)                      |  | 0.963      | Sidewalk at 2 metres from road with cars and bikes (50 km/h)                |  | 1.034      |  |
| 7                          | Bike lane next to the road (80 km/h)                      |  | 1.122      | Shared bike/pedestrian path (no car traffic)                                |  | 0.818      |  |
| 8                          | Bike path (no other road around)                          |  | 0.862      | Pedestrian path (no car traffic)                                            |  | 0.832      |  |
| Type of pavement           |                                                           |                                                                                     |            | Walking path width                                                          |                                                                                       |            |  |
|                            | Description                                               | Multiplier                                                                          |            | Description                                                                 | Multiplier                                                                            |            |  |
| 1                          | Paving stones                                             | 1.183                                                                               |            | Narrow path (less than 1 m) with obstacles (parked bikes, flower beds etc.) | 1.072                                                                                 |            |  |
| 2                          | Sidewalk tiles                                            | 1.169                                                                               |            | Narrow path (less than 1m) without obstacles                                | 1.008                                                                                 |            |  |
| 3                          | Concrete slabs                                            | 1.085                                                                               |            | Normal path (1 – 2m wide)                                                   | 0.997                                                                                 |            |  |
| 4                          | Asphalt                                                   | 0.971                                                                               |            | Wide path (more than 2m wide, boulevard-like )                              | 0.952                                                                                 |            |  |

Figure 38: Multipliers on the VTT for different travel conditions for cycling and walking (Source: Significance (2023))

Table 31: Multipliers on the VTT for different travel conditions for cycling and walking (Source: Flügel et al. (2020))

|         |                    | Not facilitated<br>(road with cars<br>or on the<br>pavement) | Walk and<br>cycle path | Cycle lane on<br>the road | Separate cycle<br>path | All |
|---------|--------------------|--------------------------------------------------------------|------------------------|---------------------------|------------------------|-----|
| Cycling | Commuting          | 1.30                                                         | 0.97                   | 1.06                      | 0.80                   | 1   |
|         | Leisure            | 1.28                                                         | 0.96                   | 1.22                      | 0.90                   | 1   |
|         | Business           | 1.26                                                         | 0.97                   | 1.06                      | 0.83                   | 1   |
|         |                    | Not facilitated                                              | Walk and<br>cycle path | Pavement                  | Separate walk<br>path  | All |
| Walking | Commuting          | 1.05                                                         | 0.55                   | 0.52                      | 0.56                   | 1   |
|         | Leisure            | 1.33                                                         | 0.46                   | 0.47                      | 0.43                   | 1   |
|         | All<br>observation | 1.28                                                         | 0.46                   | 0.46                      | 0.42                   | 1   |

### 7.3 Moving forward with obtaining VTT estimates for active travel

As argued in **Section 7.2**, the VTT related evidence base for active travel is rather small and this is especially the case when walking and cycling are considered as the main modes of transport. To move this forward, several challenges need addressing. These challenges go beyond the design of the SP survey in an upcoming national VTT study, although significant progress can still be made in this domain too.

In the UK, TAG Unit A5.1 deals with the appraisal of active travel (Department for Transport, 2022b). Different key impact activities are identified covering health benefits (physical activity and absenteeism), journey quality impacts, safety benefits, environmental benefits, decongestion benefits and indirect tax impacts, and travel time savings. Only brief reference is made to travel time savings, where the analyst is referred to applying the rule-of-a-half, i.e. estimating the benefits of travel time savings based on changes in demand, in correspondence with the values included in the TAG Databook. In the context of the arguments presented by Flügel et al. (2021), this suggests that there is a risk of double counting present in this application because physical activity impacts are covered separately.

The different national VTT studies provide consistent estimates that journey quality improvements - provided in the form of appropriate walking and cycling infrastructure - reduce the travel time disutility. The challenge, however, is that it is difficult to disentangle these quality improvements into pure comfort factors, health benefits (e.g. being further away from cars and hence air pollution) and safety concerns. One potential way of dealing with this is by being as explicit as possible regarding the journey characteristics. This can be achieved by including attributes, not only describing the quality of cycling infrastructure and number of intersections, but also the level of risk exposure on a given route (e.g. the number of annual mortalities and severe injuries), and health benefits. The challenge, however, is whether credible experimental designs can be developed isolating these dimensions of quality, safety and health because segregated active travel infrastructure is most likely always associated with lower levels of risk and less exposure to certain health hazards. Moreover, recent research for the Department for Transport has shown that these well-being benefits, apart from improvement in life expectancy, are presently not captured by TAG but that it is challenging to capture such long-term benefits through standard SP research (Department for Transport, 2024e).

Even if these benefit components can be isolated, the question emerges whether these are already accounted for elsewhere in the appraisal of active travel transport investment, and hence are at risk of double counting. If such risks of double counting exist, one can either decompose the VTT as

described above using appropriate SP designs or develop a downscaling mechanism as adopted in the Norwegian VTT study. We therefore recommend additional conceptual research into how travel time savings fit in the overall appraisal of active travel infrastructure. The report by Nellthorp (2023) for Active Travel England suggests that progress is being made in improving appraisal of active travel investments but that additional research effort is needed. Zapata-Diomedí and Kroen (2023) highlight that these challenges are not only faced in the UK but are common in the international guidance on transport appraisal and apply to both the modelling as to the valuation of active travel investment. These discussions fit in a wider literature on making CBA more suitable for evaluating active travel policies where travel time savings only form part of a wide range of benefits of active travel (van Wee and Börjesson, 2015, Standen et al., 2019, Rich et al., 2021). Our argument here is that proper scoping of which benefit components to value, and research into how to best capture these value components in isolation is essential to progress this field of research.

An additional challenge related to eliciting the VTT for active travel is the presence of selection bias in users selecting walking or cycling as their main mode of transport. Similar discussions on selection bias in the context of differences in the VTT across modes for commute and other non-work journeys were presented in **Section 4.4**, making it difficult to separate mode from user effects. The previous UK VTT study indicates this selection bias is particularly relevant for active travel, as *“walkers and cyclists might want a time saving but they do not necessarily want a mode switch”*. This was the main motivation for ending up with the adopted final designs using within mode trade-offs for cycling. A clear solution for this is, however, not available but needs consideration in the design of new SP formats.

Mode choice SP surveys, as implemented in the recent national VTT studies outside of the UK, nevertheless offer an important advantage of being able to contrast free active modes of travel against other modes of travel associated with a monetary cost. We have, however, identified that there is a clear need to improve the detail on the travel conditions especially in relation to the Dutch and German SP designs, especially when one has an interest in decomposing the time saving benefits by quality, health and safety. This discussion resembles the argument by Hess et al. (2020) presented in **Section 3.2.1**, where it was argued that it is unclear what particular travel conditions conventional SP1 exercises relate to. We believe that the emergence of new micromobility options such as e-bikes and e-scooters allows the construction of increasingly credible and relevant choice scenarios. These can make use of the same infrastructure but achieve travel time savings at the expense of costs (either hire or charging and higher purchase costs for private ownership). We have identified initial SP mode choice designs emerging capturing the VTT associated with these new modes of travel (Baek et al., 2021, Krauss et al., 2022, Esztergár-Kiss et al., 2022) and encourage further developing these SP formats. The mode choice format also offers good potential for connecting with the potential use of RP.

ARUP et al. (2015b, Section 5.2) discuss a range of alternative design formats which can be considered. One alternative option put forward by ARUP et al. (2015b) is the use of cycling hire, which resembles the SP1 approach adopted in the Dutch national VTT study. After consultation with focus groups this option was not taken forward due to concerns of realism. As reviewed in **Section 2.5**, cycling and e-scooter sharing schemes have rapidly developed in the past ten years and such a setup may work in this context today. The real challenge that, however, remains is that of selection bias and the type of trip sampled. That is, commuters who are regularly travelling by bike or walk are unlikely to opt in to a hire scheme and make use of their personal bike or scooter. For them a long-term scenario contrasting the choice between a faster but more expensive e-bike and a regular bike may be more relevant. Shared mobility schemes are much more compatible with last mile journeys in dense urban areas, or as part of a multi-leg journey. Care therefore needs to be taken with designing SP formats for a relevant audience and type of trips to make the presented choice scenario relevant.

One aspect which is often referred to in the context of the appraisal of active travel time is the issue of small travel time savings (Mackie et al., 2001). We consider the treatment of small travel time savings a policy discussion of which the implications span beyond active travel. Dubernet (2019, Section 2.3) provides an overview of the treatment of small travel time savings in the literature and international transport appraisal guidance and reaches the conclusion that the issue has not yet been addressed satisfactorily.

To conclude, our review of the VTT for active travel modes has revealed that across the national VTT studies applied over the last ten years there is a large degree of consistency in the way the sensitivity to travel time has been decomposed in relation to the quality of the provided infrastructure (e.g. type of cycle lane and number of intersections). However, in relation to the specification of the main choice task involving trade-offs between travel time and travel costs large variations in approaches exist. The use of within mode payments for use of cycling infrastructure are not overly realistic but avoid the challenges of presenting relevant alternative options in a mode choice setting where strong user effects may be present. The increasing presence of micro-mobility options may overcome the realism issue of a mode choice context with increased presence of e-bikes and e-scooters in everyday life, but the experimental design will need to be adjusted for the type of trip made because not for all trips hiring shared mobility options are relevant. As such, determining the right level of segmentation is essential and different segments should be presented with different framings of a valuation exercise.

We have, however, highlighted broader challenges related to separating out mode and user effects, but also additional health and safety benefits which may be implicit in mode and route choices for active travel. Together this calls for broader research into the appraisal of active travel to avoid potential double counting and the treatment of small travel time savings.



## 8. Aviation

This chapter summarises recent practice and evidence on the VTT for aviation. It is fair to say that the available information is rather patchy. The main reason for this is that VTT studies tend to focus on surface modes and there is not a strong tradition of undertaking similar studies for air. A second, but perhaps related, issue is that official guidance manuals on VTT exhibit a range of approaches in terms of whether and how they deal with aviation.

In what follows, the note summarises the official position in several countries – this is intended to be indicative rather than exhaustive. This is followed by a summary of a small number of studies from the academic and national VTT literature. The focus has been restricted to recent evidence since the last UK study, i.e. 2015 onwards.

### 8.1 Official national guidance

#### 8.1.1 UK position

TAG Unit A5.2 (Department for Transport, 2023f) notes that aviation interventions may generate time savings to passengers, in respect of:

- a) their journey through the airport;
- b) their time waiting at the airport;
- c) their in-vehicle time on the plane.

To the above, one might also add:

- d) airport access and egress time, which can be non-trivial;
- e) transfer time between airports, which is especially relevant to the London airports;
- f) flight delays.

The guidance advises: *‘Where possible and practical to do so, the time savings associated with an aviation policy intervention should be assessed, considering the change in the generalised journey time (GJT) and associated value(s) of time across the end-to-end journey of the passenger, and ensuring travel time benefits are not double counted across Generalised Journey Time components’.*

It should however be highlighted that the TAG Databook (Department for Transport, 2024f) does not issue specific recommendations on the VTT for aviation – and indeed aviation was explicitly scoped out of the most recent 2014/15 UK national study (ARUP et al., 2015a). More specifically, the guidance states: *‘In the absence of aviation specific values of time, the surface mode values of time can be used for leisure and commuting journeys. Since business values of time are available for various surface modes and distances it is advised that you contact the department for guidance on this matter’.*

In its Aviation Modelling Suite the DfT makes reference to the VTT for aviation (Department for Transport, 2024b) and present a VTT for business and leisure travellers in 2024 of respectively £26.63, and £7.41 (expressed in 2015 prices). These values apply to the VTT for surface access to airport and have been derived using the results from ARUP et al. (2015a) after applying corrections for the – on average – higher income of the population travelling by air.

For specific infrastructures bespoke analysis has been undertaken where users of the infrastructure are split into airport passengers accessing the airport zone and non-airport passengers accessing the remaining zones. These different types of passengers then have different values of time applied to them. An example where this has been the case is on the Western Rail Link to Heathrow, and in theory this approach can be used on any piece of infrastructure providing benefits to both airport and non-airport users.

### 8.1.2 US position

By contrast, the US DoT does issue guidance on VTT for aviation alongside surface modes. These values are based on evidence from the literature – including evidence from other countries. The *values recommended by US DoT (2016)* are presented below in **Table 32**.

*Table 32: Values per person hour, as a percentage of total earnings (plus range); hourly earnings in 2015 \$ pph; VTT in 2015 \$ pph (plus range) (US DoT, 2016)*

|                                       | Business         |        |                |       |             | Personal         |       |                |       |             | All   |             |
|---------------------------------------|------------------|--------|----------------|-------|-------------|------------------|-------|----------------|-------|-------------|-------|-------------|
|                                       | % total earnings | range  | total earnings | VTT   | range       | % total earnings | range | total earnings | VTTS  | range       | VTT   | range       |
| Surface models except HSR – local     | 100              | 80-120 | 25.40          | 25.40 | 20.30-30.50 | 50               | 35-60 | 27.20          | 13.60 | 9.50-16.30  | 14.10 | 10.00-17.00 |
| Surface models except HSR – intercity | 100              | 80-120 | 25.40          | 25.40 | 20.30-30.50 | 70               | 60-90 | 27.20          | 19.00 | 16.30-24.50 | 20.40 | 17.20-25.80 |
| Air & high speed rail                 | 100              | 80-120 | 63.20          | 63.20 | 50.60-75.80 | 70               | 60-90 | 36.10          | 36.10 | 31.00-46.50 | 47.10 | 38.90-58.30 |

Focussing specifically on aviation, a comprehensive guidebook was also produced in 2015 by National Academies of Sciences and Medicine (2015a). This guidebook was prepared as part of Airport Cooperative Research Program (ACRP) 03-19: Passenger Value of Time, Benefit-Cost Analysis and Airport Capital Investment Decisions.

The abstract to the guidebook states:

*‘Travel time is often a significant consideration in benefit-cost analysis (BCA) for transportation projects and policies, and in recent years, there has been greater emphasis on use of BCA for airport capital investment. The Federal Aviation Administration (FAA) requires benefit-cost analysis for airport capacity enhancement projects funded through the Airport Improvement Program (AIP). Similarly, US Department of Transportation (USDOT) considers the result of cost-benefit analysis when selecting projects for discretionary funding programs, such as the recent TIGER (Transportation Investment Generating Economic Recovery) Program grant rounds. The Office of Management and Budget (OMB) has also long required the USDOT to estimate the incremental costs and benefits during Regulatory Impact Analysis (RIA) for rulemaking, such as occurred for the 2009 Enhanced Airline Passenger Protections (or “Passenger Bill of Rights”) rule.’*

*‘The same need to understand the benefits and costs of investments applies to airport managers. Airports invest billions of dollars in infrastructure and service investments, yet airport owners and operators have relatively limited information on how customers value the impact of these investments. In addressing the question of how to best allocate limited resources, one can ask “would air travellers prefer improvements to the airport access roads, improved security processing times, or a people-mover connection between terminals?” This guidebook provides a method for airport owners and operators to determine how their customers value the travel time impacts of efficiency improvements.’*

The recommended values are a combination of previous evidence from the literature as well as a survey of air travellers conducted in 2013. The survey was based on 1,260 travellers who made flights between 172 distinct origin airports and 148 distinct destination airports throughout the country –

thus the scope was restricted to domestic flights. The survey employed SP experiments, inviting travellers to select among various combinations of alternative scenarios that involved (a) different times to complete various segments of trips, as well as (b) different trip costs. A detailed description of the analysis is provided elsewhere (National Academies of Sciences and Medicine, 2015b).

The recommended values as presented in **Table 33** distinguish between four stages of the journey, namely:

- Ground access time (departure) and ground egress time (arrival)
- Terminal groundside (departure) – Terminal access time (parking, shuttle bus), check-in and security screening time, time to reach the gate area, gate time
- Airside (flight) – Flight time, including boarding time and time to connect between flights at an intermediate airport, unexpected flight delay.
- Terminal groundside (arrival) – Time to reach baggage claim or exit the terminal (if no checked bags), baggage claim wait time and time to exit the terminal with claimed bags.

*Table 33: National Academies of Sciences and Medicine (2015b) recommended values in US \$ per hr (2013 prices)*

| Time category                    | Business | Leisure | Airport composite |
|----------------------------------|----------|---------|-------------------|
| Ground access time               | 18.60    | 16.95   | 17.60             |
| Ground egress time               | 18.60    | 16.95   | 17.60             |
| Terminal access time             | 33.85    | 26.00   | 29.15             |
| Check-in and security time       | 37.20    | 28.45   | 32.00             |
| Time to reach gate area          | 32.25    | 22.85   | 26.65             |
| Gate time                        | 20.50    | 17.60   | 18.75             |
| Flight time                      | 51.00    | 34.90   | 41.40             |
| Unexpected flight delay          | 286.30   | 123.30  | 189.15            |
| Time to reach bag claim/exit     | 32.25    | 22.85   | 26.65             |
| Baggage claim wait time and exit | 37.20    | 28.45   | 32.00             |

### 8.1.3 Dutch position

There is a tradition of including aviation within official VTT studies in the Netherlands. The most recent Dutch national VTT study (Significance, 2023) estimated aviation values for in-vehicle time as well as for airport access/egress by park & fly, kiss & fly, by public transport, by taxi and by any mode. SP was again used, with the experiment offering a choice between two unlabelled journeys embodying a trade-off between travel time, cost, and travel time reliability. Separate games were devoted to the main stage and the access/egress journey, but the two games were modelled using a joint estimation approach. Both domestic and international flights were admitted.

Some key covariates from the Dutch study indicate that:

- There is a positive relationship between VTT for aviation and income.
- The VTT for business trips is typically 19% higher than for non-business.

- The VTT for airport access is between 78% and 89% lower than the VTT for the in-flight component of the air trip, depending on the mode used for the access trip.
- The VTT for airport egress is between 54% and 76% lower than the VTT for the in-flight component of the air trip, depending on the mode used for the egress trip.
- For airport egress, the VTT is higher than for airport access.

In terms of the recommended national average values for aviation, as presented in Table 34 this implies that – due to correlations across explanatory variables – the in-vehicle VTT for business is nearly double than that for leisure.

Table 34: Recommended values for aviation in the Dutch VTT study (values in € per hr, 2022 prices) (Significance, 2023)

|                                   | Business        | Leisure        | All           |
|-----------------------------------|-----------------|----------------|---------------|
| <b>In-vehicle Time</b>            | 110.22 +/-1.45  | 53.80 +/-3.18  | 61.79 +/-0.38 |
|                                   |                 |                |               |
| <b>Airport access/egress time</b> | <b>Business</b> | <b>Leisure</b> | <b>All</b>    |
| Park & fly                        | 31.49           | 13.90          | 15.57         |
| Kiss & fly                        | 18.59           | 11.12          | 11.86         |
| Taxi                              | 35.62           | 13.72          | 21.51         |
| Train or train & bus              | 15.02           | 8.26           | 9.11          |

#### 8.1.4 Norwegian position

Another country which includes aviation within official guidance is Norway. The most recent national study was undertaken in 2020 (Flügel et al., 2020). Values of in-vehicle time for domestic aviation were based on a simple time vs. cost SP experiment. Further experiments employed time-based trade-offs to derive transfer and cancellation penalties (quantified in IVT units per transfer/cancellation), as well as valuations of access time by a range of options (as multipliers of in-vehicle time).

Table 35: Recommended values for aviation in the Norwegian VTT study (values in NOK per hr, 2018 prices unless otherwise stated) (Flügel et al., 2020)

|                                             | Business | Commute | Leisure | All  |
|---------------------------------------------|----------|---------|---------|------|
| <b>In-vehicle time</b>                      | 792      | 450     | 267     | 495  |
| Transfer penalty (mins of IVT)              | 13       | 53      | 53      | -    |
| Cancellation penalty (hrs of IVT)           | -        | -       | -       | 11.8 |
| Access time by car (multiplier of IVT)      | -        | -       | -       | 0.8  |
| Access time by Airtog (multiplier of IVT)   | -        | -       | -       | 1.0  |
| Access time by Train (multiplier of IVT)    | -        | -       | -       | 0.8  |
| Access time by Airbuss (multiplier of IVT)  | -        | -       | -       | 1.0  |
| Access time by Rutebuss (multiplier of IVT) | -        | -       | -       | 0.9  |
| Access time by Taxi (multiplier of IVT)     | -        | -       | -       | 0.9  |

#### 8.1.5 German position

Axhausen et al. (2014) consider aviation as a relevant mode choice for long distance journeys ( $\geq 500km$ ) in their SP design and contrast the option against car and public transport as suitable alternatives. The air journey was characterised by its in-vehicle travel time, access, and egress time, waiting time, travel costs, number of transfers, flight frequency, and the chance of delay. As can be observed in

Table 5, aviation was considered for the mode choice component of the short term SP surveys but not for the other types of SP choices (Dubernet and Axhausen, 2020). This is, however, in contrast to Axhausen et al. (2014) who also provide examples of choice cards for within-mode SP choices contrasting alternative flights in terms of their reliability amongst other attributes (see Figures 15 and 16 in Axhausen et al., 2014). **Table 36** provides the recommended values emerging from the German study, which are much lower than those identified in the Dutch VTT study, but the exchange penalty is much higher than those found in the Norwegian study.

*Table 36: Recommended values for aviation in the German VTT study (values in € per hr, 2012 prices unless otherwise stated) (Axhausen et al., 2014)*

|                                                        | Leisure | Business | All   |
|--------------------------------------------------------|---------|----------|-------|
| <b>In-vehicle time</b>                                 | 24.88   | 36.82    | 32.33 |
| Access time (multiplier of IV)                         | 1.1     | 1.1      | 1.1   |
| Waiting time (whilst transferring) (multiplier of IVT) | 2.1     | 2        | 2     |
| Exchange penalty (mins of IVT)                         | 62      | 61.1     | 61.2  |
| Delayed arrival (multiplier of IVT)                    | 1.6     | 1.6      | 1.6   |

## 8.2 Other studies

A recent study of possible interest is Merkert and Beck (2017)'s study in Australia. By way of introduction, Merkert & Beck comment on p30 that *'While some of the often cited national transportation VTT studies specifically exclude aviation (such as the studies in Sweden ..., Denmark ..., and Switzerland, others such as the Netherlands include it at the inter-urban (>100 km) level, and again others such as the UK... and Norway... include it at both the inter-urban levels of less than 50 km and more than 50 km. Indicatively, Abrantes and Wardman (2011) show with their UK meta-analysis that air transport users have on average a VTT value around 4.25 times larger than car users, and Shires and de Jong (2009) confirm, in line with the national VTT studies, that aviation has on average much higher VTT values than other modes of transport, particularly for business and long distance travel'*.

Merkert and Beck (2017) estimated values via an SP experiment of mode choice for domestic journeys, distinguishing between the main stage of the journey and the onward journey to the final destination. The modes offered were plane & taxi (the latter for the onward journey); plane & integrated shuttle; drive in own car; and coach/tour bus. The attributes considered for the main stage were: departure time; average travel time; travel time reliability; cost. The attributes considered for the onward journey were mode; wait time for service; travel time; cost; payment for service. Data was collected via an online panel. The results are reported in **Table 37**.

*Table 37: Results from Merkert and Beck (2017) SP study on the VTT for aviation in Australia (Values in AU\$ in 2014 prices)*

|                |     |                |     |
|----------------|-----|----------------|-----|
| Business       |     | Leisure        |     |
| Lower (95% CI) | 29  | Lower (95% CI) | 63  |
| Median         | 153 | Median         | 99  |
| Upper (95% CI) | 312 | Upper (95% CI) | 144 |

Wardman et al. (2016) report an extensive meta-analysis of values of time, covering 3109 monetary valuations from 389 European studies conducted between 1963 and 2011. An objective of the paper is to explain how valuations vary across studies, including over time and between countries. As well as in-vehicle time, the paper covers valuations of walk time, wait time, service headway, parking space search time, departure time switching, time in congested traffic, schedule delay early and late, mean

lateness and the standard deviation of travel time. The paper also covers the full range of modes, including aviation. In passing, Wardman et al. notes that official national values for air travel have at some time existed in Norway, Sweden, the Netherlands, Germany, and France.

The dataset includes 27 observations from the long-distance market where air competes with rail, and Wardman et al. (2016) reports that, on average, air values are 62% greater than rail values, all else equal. Having introduced and described this data, they then proceed to estimate a meta-model, which serves to demonstrate the key dimensions by which values show variability. Finally, the model is applied to forecasting, to derive values by country and mode used. This includes values for business travel by air, where value of 51.82 for the UK, in euros per hour in 2010 income and prices, is derived.

### 8.3 Findings and conclusions

Ideally, a national VTT study would encompass aviation values. These are needed for policy and analysis work at DfT and comparative national VTT studies in northern Europe have included aviation within their scope. Aviation was, however, out of scope in the previous national VTT study.

It is feasible to estimate robust values for many time components of aviation journeys. Aviation values can be estimated using similar or the same research methods as used for other modes. Key time components potentially of interest to DfT would be in-vehicle time, delay, access and egress time, and transfer time within and between airports.

It should be acknowledged that there are some idiosyncrasies of air which introduce additional considerations, e.g. the scope should potentially be restricted to domestic flights to maintain consistency with the NTS which is typically used to sample enumerate for national representativeness, although alternative datasets may potentially be available through the CAA. Given the security provisions around airports, this may introduce additional complications for survey work, especially if employing an intercept approach.

## 9. Is this the right time for a new national VTT study?

The evidence base for the present VTT values implemented in TAG are becoming more than ten years old. Our review of societal developments since the 2014/15 study in **Section 2.5** revealed that during this ten-year period significant changes in travel behaviour have taken place, not least due to the Covid-19 pandemic and the cost-of-living crisis. Altogether, these events, the increasing electrification of cars and buses, and the increasing presence of micro-mobility options, are likely to have influenced the VTT. We are therefore of the opinion that obtaining new estimates of the VTT is desirable, even if only to validate the present values used in transport policy appraisal.

Although the recovery from the Covid-19 pandemic is still ongoing, aggregate travel patterns have largely recovered and have sufficiently stabilised. Moreover, and despite changes in individual travel patterns (e.g. frequency, mode choice, destination choice, time-of-day etc.) being substantive, our view is that this should not be a significant barrier to conducting a new national VTT study.

In contrast to the scoping studies conducted by ITS (2010) and ITS (2013), the DfT finds itself in a somewhat different position. At the time of commissioning the ARUP et al. (2015a), TAG guidance existing at that time was based on data which had been collected in 1994 – some 20 years prior. Fast-forwarding to the present day, current TAG guidance is based on data collected in 2014 – some 10 years ago. Our view is that, *in broad terms*, the approach followed in 2014-15 remains sound and can be repeated. A conservative approach to a new national VTT study can therefore aim for a repeat of ARUP et al. (2015a) and deliver a robust update of the VTT in the UK.

Whilst supportive of the broad approach followed in 2014-15, it perhaps comes as no surprise that, given the experience of that study, in combination with subsequent developments in methods, we would not necessarily recommend an exact repeat of the study, but to take the study as a robust point of departure. With the benefit of hindsight, it may be unwise to repeat certain aspects of the previous study and/or, with the passage of time, certain aspects may have become redundant or can be improved upon.

In the remaining sections of this report we will set out a recommended approach to a new national VTT study which largely builds on the design of the ARUP et al. (2015a) study, makes suggestions for improvements, but also presents some more progressive options either relating to the use of revealed preference methods, and the design of stated preference methods for active modes of travel and aviation.

The specifications that follow should be taken as illustrative, rather than a definitive approach. Specifically, the study team eventually commissioned to conduct a new national VTT study should be given some degree of freedom in shaping the study specification and respond to challenges in the piloting stages of the study.



## 10. Evaluating the design of the ARUP et al. (2015a) study

The ARUP et al. (2015a) study had three overall objectives:

1. Provide recommended, up-to-date national average values for in-vehicle travel time savings.
2. Improve understanding of what drives the VTT and the uncertainty around these values.
3. Consistently estimate values for trip characteristics of related factors, e.g. reliability and crowding.

In terms of the objectives of a new study, we do not see reasons to fundamentally change these. The ultimate objectives of the study are clearly reflected, and it is only in the more detailed set of requirements, such as the desired level of segmentation or the related factors deemed most relevant, that minor changes may need to be specified.

The study was conducted in two phases involving a developing and testing phase followed by a phase involving the collection and analysis of survey data. The second phase included the translation of the behavioural VTTs into recommended national average VTT values. With an overall timeline of 10 months the timescales were considered (and proved to be) challenging by the original study team. Specifically, additional time for the translation of the behavioural VTTs into appraisal VTTs is recommended to be factored in. Based on recent national VTT studies in other European countries, we would strongly recommend a programme duration of around 18-24 months to design and conduct the data collection efforts, deliver the analysis work – plus additional time to undertake stakeholder consultation and achieve the necessary signoffs to commit to TAG.

### 10.1 Segmentation by mode and purpose

The ARUP et al. (2015a) study focussed on passenger VTT, and considering the recent completion of a dedicated road freight VTT study (AECOM et al., 2023), there is no reason to expand the scope of a new study beyond passenger VTT.

The previous VTT study considered three ‘required’ (car, bus, and rail) and two ‘optional’ (walk & cycle, and other public transport<sup>26</sup>) modes of transport. The study was able to successfully elicit behavioural VTTs for the required modes of transport and other PT but not for walk and cycle. We recommend that the core of the new study is focused around the same four motorised surface modes of transport, but that two separate tranches of research are developed covering the areas of active travel and aviation respectively.

Our review of methods and evidence indicated that the four motorised surface modes of transport are largely consistent with the level of segmentation implemented in other national VTT studies and hence form a good basis for a new study. For active travel and aviation such a robust study approach is not readily available and needs careful designing of the actual study objectives and suitable approach. This refers to the separation of quality, health and safety factors contributing to the VTT for active travel and the segmentation of the VTT for different components of the journey. In principle, these two tranches of research can be completed independently from the ‘main’ study, especially if and when the focus of the SP designs remains on using within-mode route choices. When mode choices are considered a more appropriate SP design format, there may be a benefit of bringing the three (or some) of the study components together in a final study. Naturally, the timing of the different pieces of research should be aligned to avoid time lags between VTTs for different modes of transport. The

---

<sup>26</sup> Underground, tram and light rail.



Norwegian national VTT study has successfully implemented such independent data collection efforts. That said, we return to active modes of travel and aviation in **Sections 12.3** and **12.4**.

With regards to trip purpose, the three main trip purposes of commute, business, and other-non work are consistent with those considered in other recent national VTT studies. Indeed, the latter trip purpose may be segmented further if desired, but there is no obvious argument to deviate from the present level of segmentation (see also **Section 11**).

For business trips an additional level of segmentation in ARUP et al. (2015a) considered the WTP-approach separately for employees and employers. We recommend that a new study should focus on the employee WTP-approach. The primary reason for this is that the recruitment of employers was extremely challenging – in 2014/15, contacting the relevant staff responsible for the company’s travel policy turned out to be difficult in practice. Furthermore, focussing on employees’ WTP would bring the UK into line with other national VTT studies.

In contrast to ARUP et al. (2015a), where the Hensher equation was considered out of scope for the valuation of travel time for business trips, the review of methods and evidence highlights that the Hensher equation remains a relevant alternative to the valuation of travel time for business travellers. Appendix H of ARUP et al. (2015a) reported on the respective parameters of the Hensher equation and as such its inclusion in a new study only requires a review and updating of the relevant questions at marginal additional costs.

## 10.2 SP-Experiments

The main approach to data collection adopted in ARUP et al. (2015a) was Stated Preference (SP), specifically through three different games – referred to as SP1, SP2 and SP3. Simple route-choice based travel time vs travel cost trade-offs in SP1 were used to estimate national average values in TAG. The present review of methods and evidence finds that this approach remains consistent with recent national VTT studies in Norway and the Netherlands. The German VTT study primarily relies on SP data but makes use of an alternative experimental design focused on mode choice as opposed to route choice. Accordingly, there is therefore no direct concern regarding the feasibility of repeating the route choice approach of SP1.

In addition to SP1, ARUP et al. (2015a) implemented two sets of additional SP-games per traveller, respectively SP2 and SP3, capturing different components of the VTT, such as reliability (SP2) and quality aspects such as congestion and crowding (SP3). Again, this approach is consistent with international practice where the use of additional SP games is common practice. The general approach to SP2, using the mean-standard deviation approach, is also reflected in other national VTT studies and is still considered as the state-of-the-art. The same applies to the rail-based crowding-based studies in SP3 eliciting crowding multipliers (in pax/m<sup>2</sup>) and standing multipliers. Options do exist for improvements on various points of detail, as discussed below.

The series of follow-up studies on congested values of travel time (CVTT) outlined in **Section 6.1** have highlighted that there remain unresolved issues in relation to the implementation of the categorical levels of congestion used in the SP3 car games, and the potential risk of double counting when congestion multipliers are implemented alongside the value of travel time reliability. Whilst from a modelling perspective a redefinition into free flow time and delay is considered a better formulation, the research suggests this only increases the risk of double counting. In absence of a clearly preferred SP design capturing the impacts of congestion on the VTT, it is recommended to continue building an evidence base of congestion multipliers especially because of its potential importance in appraisal and policy making. The implemented design, however, needs improving beyond the basic categorical levels

of low and high congestion to facilitate implementation. We return to the topics of congestion and reliability in **Section 13.1**.

The categorical descriptions of crowding levels on buses and other (non-rail) forms of public transport in SP3 suffers from similar challenges as the categorical description of congestion levels for cars. In this case, however, the rail-based crowding approach making use of occupancy levels (pax/m<sup>2</sup>) and standing multipliers can easily be adopted and hence this is what we recommend changing.

The SP3 games in ARUP et al. (2015a) also identify separate VTT multipliers for in-vehicle time, access/egress time, waiting times etc.. This strategy remains feasible, and the precise specification/definition of these travel time components can be changed as the Department wishes, without significant implications for the approach to be adopted.

ARUP et al. (2015a) presented each respondent with all three SP games comprising five choice tasks each. At the level of the individual respondent, this limits the amount of information per respondent within a given SP game. The alternative approach is to present each respondent with only two SP games but increasing the number of choices within each SP game. This approach is adopted in the most recent Norwegian and Dutch studies, presenting respondents with eight choices per SP game. This change will not have ramifications for overall sample sizes but will shift the evidence base somewhat more towards SP1.

The Dutch VTT study provides a good reason for making this change. Significance (2023) highlights that the chosen experimental design approach in the 2014/15 national study, based on d-efficient designs, significantly influenced the time vs cost trade-offs variation presented. A more consistent degree of time and cost variations in SP1 can be introduced making use of what is known as the ‘Scandinavian’ designs. The Dutch study introduces an interesting hybrid design making use of both methods with reported success on reducing non-trading behaviour in the sample. Irrespective of the design strategy chosen, it remains vital that pilot studies confirm that across a range of reference trips the presented time and cost trade-offs are credible and relevant for the respondents.

The use of a pivot design approach – i.e. presenting values of travel time and travel cost around those of the reference trip – is still considered best practice because it increases the amount of variation, and thereby information, to learn about the VTT in the sample. Complementing these with ‘fixed’ attribute levels, as was done in SP2 and SP3, also remains best practice.

For reasons of comparability with the 2014/15 sample, it is recommended that at least a subset of the new sample will be presented with the original experimental designs such that any differences in the emerging VTT can to a certain degree be attributed to changes in the VTT over time, changes in the data collection method, and (or) changes in model specification. In the Dutch study, 10% of non-business respondents were presented with the design specification from the previous national study.

### 10.3 Covariates

Section 1.4.3 of ARUP et al. (2015a) presents a range of covariates recorded in the survey. These include income, distance, productive time, trip type, group size and composition, driver vs. passenger, trip frequency, and time-of-day. Our review of methods and evidence did not identify any obvious missing covariates in the survey for the previous national VTT study.

In relation to the income variable, a new study may wish to elicit more detailed household information such that the income variable can be translated into approximations of alternative definitions of income, such as equivalised household income. To make such transformations effective the chosen

income categories presented in the survey need refining in relation to relevant benefit thresholds at the personal and household level.

#### 10.4 RP component

A small part of the brief for the 2014/15 study covered Revealed Preferences (RP) with the aim of validating the SP evidence for rail travel. This was conducted as a within-mode trade-off between real world train operators operating on the same route providing trade-offs in terms of travel time and travel cost. The implementation was not considered successful, and our review of methods and evidence highlighted that better RP alternatives are available now (e.g. Flügel et al., 2022, Tsoleridis et al., 2022). We therefore do not recommend including the same RP component. Instead, we recommend a separate research programme developing a suitable RP approach taking a more prominent role a new national VTT study in **Section 12.2**.

#### 10.5 Recruitment methods

Two recruitment methods were implemented for the main SP surveys in ARUP et al. (2015a), respectively intercept and telephone recruitment.

Intercept recruitment is still considered a viable means of sampling respondents. The implemented method of sending a follow-up link to the actual questionnaire to intercept respondents also remains good practice in the literature and there are no indications pointing towards changing this approach.

The optimal choice for selecting the intercept locations needs to be re-evaluated based on the current road network and aims of the survey. One area of attention is the approach to sampling (intercepting) short distance trips. The 2014/15 study was clearly affected by the higher potential of intercepting on long distance trips, during which people take breaks relative to short distance trips on which travellers do not stop.

Tailored methods of recruitment are therefore needed, also to accommodate the fact that not all respondents may have access to the internet to complete the survey. In principle there is no objection against re-using telephone recruitment for this purpose. It should, however, be considered that fewer people have landlines, and response rates to cold calling on mobile phones may have reduced over time.

**Section 3.3** discussed alternative methods of recruitment and the general sentiment is against the use of online survey panels as they tend to be associated with selection bias from respondents with a low VTT. Higher quality internet panels are available making use of random probability address-based sampling improving the representativeness of online samples, and alternatively address-file postal surveys can be used. Relative to intercept sampling these methods may be cheaper but they need consideration as to how each respondent is linked to a specific reference trip, as opposed to linking them to the trip they are intercepted on. These methods may, however, improve the sampling of shorter trips.

To attract respondents to the survey, ARUP et al. (2015a) offered £10 vouchers incentivising them to complete the main questionnaire, which increased in the latter stages to £20 to meet certain quota. The appropriateness of these incentives will need to be evaluated during the piloting stages of the survey, keeping in mind that higher incentives may be needed to convince people with higher values of travel time to partake in the survey, but that at the same time adverse selection affects may be stimulated (i.e. attract even higher rates of respondents with a low VTT).

## 10.6 Sampling targets

Table 2.9 in ARUP et al. (2015a) sets out the SP sampling targets for the different mode purpose combinations. Under the assumption that the levels of segmentation and assurance for the intended VTT measures remain broadly the same as 2014-15, our review of methods and evidence has not indicated an obvious need for change. However, it may be worth considering increasing the target samples for bus and other PT to 1,000 per mode-purpose combination. This would allow for the same level of detailed analysis as for car and rail.

Table 38: Target SP sample sizes in ARUP et al. (2015a)

|          | Employee's business | Commute | Other non-work | Totals |
|----------|---------------------|---------|----------------|--------|
| Car      | 1,000               | 1,000   | 1,000          | 3,000  |
| Bus      | N/A                 | 500     | 500            | 1,000  |
| Rail     | 1,000               | 1,000   | 1,000          | 3,000  |
| Other PT | 500                 | 500     | 500            | 1,500  |
| Totals   | 2,500               | 3,000   | 3,000          | 8,500  |

The above sampling targets are relatively generic and do not give a large degree of steer on the types of travellers and trips sampled. Our review of methods and evidence highlights that there is additional benefit of identifying relevant sampling strata for, for example, distance and income levels within these eleven mode-purpose combinations. We are not arguing for a representative sampling approach, but ensuring that, for the relevant variables of interest, sufficient variation is included in the data to confidently estimate interaction effects between the VTT and covariates of interest. Being more prescriptive, however, comes at a monetary costs of additional sampling efforts needed to fill specific sampling quota.

## 10.7 Developing and testing approach

Alike the process adopted to develop and test the chosen SP approach in ARUP et al. (2015a, Section 2.13), we recommend that cognitive in-depth interviews and pilots are put in place to test the wording of the survey, its flow, comprehension, and the SP exercises. Although there are no clear indications that from the analysts' point of view that significant changes are required, this does not imply that changes in society may warrant, for example, rewording of questions or other changes to the survey. The use of sample sizes of around 5-10% of the target sample size in the pilot is considered sufficient such that around 50 respondents are consulted within each mode-purpose combination.

## 10.8 Data analysis and behavioural VTTs

Having made recommendations regarding the approach to collecting SP data, the remainder of this chapter takes a much less prescriptive stance regarding the analysis of the data. First and foremost, analysts of a new national VTT study need the freedom to respond to the nature of the collected data and any challenges emerging in the process. Secondly, we encourage an open conversation between the DfT and the analysts such that informed decisions can be made on data cleaning, data transformation, final model specifications, and translation of behavioural VTTs into appraisal VTTs.

Where feasible, the specifications of the 2014/15 behavioural VTT models need to be replicated on the newly collected data such that any differences in the VTT from the current values can be attributed to actual changes in the VTT as opposed to a mixture of changes in the VTT and the adopted model specification. Naturally, the final model specifications may – and probably should – vary from those described in ARUP et al. (2015a).

As indicated by **Section 3.4**, the state-of-the-art on analysis has not significantly changed over the past ten years. Therefore, the review gives us no reason to fundamentally question the validity of the estimation methods used in 2014-15 and the resulting values which appear in TAG. On this basis, a direct repeat of the ARUP et al. (2015a) data analysis is not necessarily discouraged.

That said, where the final model specifications in ARUP et al. (2015a), but also in other recent national VTT studies outside of the UK, were largely driven by the empirical fit of the model to the data collected, we recommend that micro-economic theory should play a more prominent role in future studies. Key features accounted for in the behavioural models of ARUP et al. (2015a), like reference dependence, cost damping, and multiplicative error terms, are not consistent with neoclassical micro-economic theory. The latter is concerning given the importance of VTT in policy appraisal, with many of these features providing challenging in the context of aggregating (and averaging) VTT measures. This does, however, not mean that the use of alternative theories originating in behavioural economics are invalid. On the contrary, they provide valid insights that the VTT may vary depending on the application context. Our point is that the implications of such context dependent behavioural phenomena for policy appraisal are all but straightforward. We therefore recommend that the DfT should actively engage with the analysts to discuss, and challenge chosen functional forms, and thereby obtain a better picture as to how modelling choices made affect the behavioural VTT and what the corresponding implications are for deriving representative VTTs suitable for appraisal purposes.

Ultimately a balance between the empirical and the theoretical approach will need to be struck, but at present the balance weights, in our opinion, perhaps too much towards the empirical approach. This does not mean that modern approaches to data analysis, such as machine learning, should be disregarded. Indeed, machine learning may offer further flexibility in functional form, but we recommend that such techniques are employed to assist model specification of choice models such that structural core and desirable features of such models, e.g. theoretical consistency, remain. As described by van Cranenburgh et al. (2022), machine learning can assist in finding the optimal utility function, and capturing systematic and random preference heterogeneity. These machine learning methods can *“determine an optimal utility function from an exponentially large set of possible utility functions in a purely data-driven manner”*. As such this avoids subjective choices as to which variables to include (or exclude) and which potential interactions to test but based on this empirical model the researcher needs to make informed decisions regarding the extent the information contained by the data can be captured by a functional form which aligns with economic theory. Inevitably, this process will be associated with subjective judgements of the modelling team.

A topic of discussion common across recent national VTT studies using separate SP games, including ARUP et al. (2015a), is that of differences in the VTT across SP games. Unfortunately, this is inherent to the chosen data collection approach and is difficult to fully resolve. One option is to solely rely on a single SP game where all the relevant attributes describing the route or mode are presented (e.g. reliability, congestion, crowding). Hess et al. (2020) are a strong proponent of such ‘complete’ SP designs but this specification easily becomes difficult to process for respondents as the number of attributes grows. The use of separate SP games provides a feasible solution to overcome this burden on the respondent but is inherently associated with potential differences in the interpretation of the travel time presented in the simplest SP games, and inherently respondents may make choices which reveal differences in the modelled VTT from the different SP games. By using joint estimation, as applied in ARUP et al. (2015a), these underlying VTT estimates can either be forced to be the same resulting in potential biases, or recognised to be different. In ARUP et al. (2015a) this resulted in the estimation of distinct multipliers across SP games leaving the Department with a decision as to which VTT to use for appraisal purposes. The consistent choice across national VTT studies has been to use

the VTT from simple time-cost trade-offs (i.e. SP1). Again, there is no right or wrong approach here. Relying on SP1 presents a tested approach and ensures consistency with preceding national VTT studies, but a case can be made to consider SP games with larger number of attributes.

Estimated behavioural VTT models should include a range of explanatory variables capturing traveller and trip characteristics, including income, cost, time, distance, etc. The analysts should keep in mind the extent to which these explanatory variables are correlated, and the extent to which identified parameters capture a pure effect, or whether something like distance ultimately approximates the impact of a range of variables (time, cost, income etc.) on the VTT. Although this is not such an issue for behavioural VTTs, this is much more relevant in the translation to appraisal VTTs.

### 10.9 Translation into appraisal VTTs

ARUP et al. (2015a) made use of sample enumeration to translate the behavioural VTTs to appraisal VTTs, i.e. recommended national average values. The approach taken was distinct from other national VTT studies in two ways because it i) relied on three years of NTS data to which the behavioural VTT models were applied, as opposed to the more common approach in other national VTT studies of applying the estimated models to the sampled survey data; and ii) the use of distance weighting.

With regards to the former, the benefit of making use of the NTS is having a representative sample of trips, and relevant expansion weights, to which the model can be applied. It is, however, unlikely that the NTS covers all the explanatory variables included in the behavioural models, and a disconnect between the estimated relationships and the sample enumeration process may therefore arise. In ARUP et al. (2015a) this was solved by taking the average for such missing variables from the SP sample for given mode-purpose combinations. If this approach is repeated, additional levels of segmentation can be implemented (e.g. income, or travel distance) to determine relevant averages for the missing variables. Good practice would be to test the sensitivity of the average VTT to the assumptions made.

Implementing the sample enumeration at the level of the SP sample does not introduce these concerns since all variables are included in the survey data. Another challenge, however, arises requiring the reweighting of the sampled population and trips to represent the national travel population. It is essential to have a good understanding of which dimensions determine the expansion weights in the NTS, as each of these 'cells' need good representation in the sampling of survey respondents. This may, however, conflict with other sampling interest (or pose additional sampling requirements, i.e. quotas) where, for example, we may wish to over sample certain types of trips to understand how the VTT varies by a given variable which is not central to the expansion weights. The literature does not provide a clear recommendation on which approach (NTS or SP sample-based sample enumeration) is most desirable.

Beyond the standard expansion weights, decisions need to be made regarding the appropriateness of using additional weights, such as distance weighting, and the use of alternative income assumptions (see income options 1 and 2 in ARUP et al. (2015a)). In terms of the former, the use of distance weighting inflates the nationally recommended VTT because longer distance journeys are associated with higher VTTs. The case made in the UK, and in the Netherlands, for using distance weights is that by travelling longer these travellers are more likely to experience travel time savings and hence should be given more weight. This practice is not adopted in Norway and Germany. In terms of the latter, we follow the recommendations of ARUP et al. (2015a) that income option 1 (individual specific income) should be used when the VTT is sufficiently segmented by mode and purpose. Income option 2 (average income for all) introduces similar challenges as described above in relation to imputing values for missing variables in the NTS database. Ultimately, decisions on the use of alternative income measures and sample enumeration weights remain policy-driven.

## 10.10 Conclusions regarding the design of ARUP et al. (2015a)

In summary, we are confident that, with minor improvements, ARUP et al. (2015a) approach is still consistent with best practice in the international VTT literature and will produce a set of robust new VTT estimates for the existing level of segmentation in the UK policy context.

This does not dismiss the broader critique around VTT methods. The use of SP as the primary data collection approach continues to attract questions regarding its validity, not least because of their hypothetical nature. SP presentations will always remain abstract and incomplete (or imperfect) representations of potential journeys. SP1 is particularly exposed to this criticism due to the use of only two attributes describing the journey. A straightforward solution is not however available, despite the German national VTT study's attempts at addressing some of these issues by framing the SP choices in the context of mode and long term residential and job location choices.

Testing the acceptance and relevance of the adopted SP designs by survey respondents therefore remains of the upmost importance, alongside validating the emerging VTT values using revealed preference (RP) data.

In the next three sections our focus moves beyond the ARUP et al. (2015a) study. **Section 11** reflects whether the level of segmentation of the appraisal VTT, i.e. the final output of a new national VTT study, should differ from the mode-purpose approach adopted in the UK and in most other national VTT studies. **Section 12** identifies essential areas deserving additional attention as these were either unresolved in ARUP et al. (2015a), respectively the separation of mode- and user-effects, and active travel; new opportunities have emerged to increase the use of revealed preference methods; or the scope of the national VTT study should be expanded to include aviation. **Section 13** identifies two exploratory areas of research focusing on the relationship between congestion and reliability, and long-term VTTs (i.e. long-term SP choice designs). Finally, **Section 14** summarises the recommendations made by this scoping study and sets out a timeline to conduct these different areas of interest.



## 11. Reflection on the level of segmentation

ARUP et al. (2015a) recommend in Section 8.3.1 (R8) that “...the Department should disaggregate VTT by distance or some geography typology (e.g. urban/inter-urban) that reflects differences in distance. This will require further work to identify appropriate distance disaggregations. Such work would involve the use of distance profiles from real scheme appraisals, to explore the full implications of the approximations to the ‘real’ VTT of the distance-weighted VTT under different distance disaggregations.”

Whilst we subscribe to the idea that the VTT increases with distance, we are not of the opinion that the necessary progress has been made to increase the level of segmentation of appraisal VTTs beyond the present segmentation by mode and trip purpose (and distance bands for employee’s business).

Before further segmenting the VTT by distance, and potentially other factors like income, we would recommend additional research focusing on i) the intended effect to be represented and potential confounding with other effects; ii) implications for study design and sampling protocols; and iii) modelling practices.

From an interpretational point of view expressing the VTT as a function of distance may act as a source of confusion, because the expressed variation in the VTT is unlikely to represent a pure distance effect. Significant correlations are present between distance, travel time, travel costs, and income. Any segmentation relying solely on distance will automatically pickup their joint impacts on the VTT. This includes higher VTTs due to increasing discomfort (distance), increasing opportunity costs (time), and the behavioural effects of time and cost damping (time and cost). As a case in point, TAG Unit M2.1 (Department for Transport, 2024i) varies the value of travel time with distance, arguing that travellers’ sensitivity to cost declines more rapidly with distance than their sensitivity to time. In other words, distance is used as a pragmatic tool to represent alternative behavioural effects. Ideally, in estimation additional effort is put in isolating these individual effects on the VTT such that robust and independent distance, time, cost and income elasticities can be estimated. If this is not feasible, it should be articulated, as a minimum, that proposed segmentations by distance are a proxy for a multitude of effects. **Section 2.4** revealed that amongst the new national VTT studies only the Norwegians have segmented by distance for non-business VTT (short vs long distance), whereas the other new national VTT studies did not segment by distance.

Decisions relating to segmentation by distance have implications for the sampling approach (i.e. identification of distance elasticities), study design (i.e. the separation of distance from mode and user effects), sample enumeration (i.e. the choice for (not) using distance weighting), and modelling practices (i.e. transport assignment models prefer a continuous relationship between distance and the VTT over a categorical classification). Specifically, increased interest in the role of distance will influence the demands on sampling strata and inevitably increase recruitment costs.

Two related issues should be considered in this context. First, whilst segmentation by distance may work well *within* modes, when working across modes speed differences also become relevant. **Section 4.1** provides illustrative discussions highlighting that the notion of short and long distance may vary substantively across modes of transport. Second, the issue of small travel time savings cannot be disconnected from distance. Particularly on short distance trips, the potential for time savings decreases and potentially becomes marginal. If, alike the Welsh government (Welsh Government, 2022), the Department at some point decides to set a lower bound on the level of time savings to be considered in appraisal, it should most likely also implement a lower bound on the relationship between the VTT and distance.



In relation to income, similar observations can be made. To isolate the impact of income on the VTT, specific design choices are required. In terms of sampling, this means ensuring a good coverage of income categories within each mode-purpose combination and across distance bands. The latter is especially important to separate the income effect from distance effects because wealthier people are known to travel longer distances. Our review has shown that satisfying such constraints is difficult for two reasons. First, there is potential self-selection taking place such that increased efforts, perhaps through increased incentives, need putting in place to ensure people with high values of time partake in the data collection (see **Section 3.3**). Second, challenges in accurately capturing income are inherent to survey-based research because drop out and non-reporting rates increase when detailed income data is requested (see **Section 4.5**). Small improvements can be made by asking additional questions on household composition and benefits received, such that better pictures on personal and household disposable (or equivalised) income can be captured. However, these provisions may not materially change the patterns of variation in VTT by income estimated by the models.

Segmenting the VTT by income inevitably implies a departure from the uniform (or equity) values of travel time currently applied for specific mode-purpose combinations. With the common understanding that the VTT is increasing with income, due to assumed decreasing marginal utilities of income, this additional level of segmentation of the VTT may unintentionally prioritise transport investments for richer parts of the populations which may go against strategic goals despite being more efficient. Such a move closer towards the use of behavioural values of travel time may be desirable to illustrate the varying degrees at which different parts of the population benefit from specific transport interventions. The distributional implications can be controlled for in a different part of the appraisal process using distributional weights. Our recommendation is to implement any form of distributional weighting at the project level (i.e. after aggregating all benefits), not at the level of individual benefits such as travel time savings or safety, to avoid aggregating a variety of potentially inconsistently weighted and unweighted benefits.

In summary, we consider the present level of segmentation of the VTT by mode-purpose combinations to be appropriate, especially considering the additional requirements and complications that further segmentation imposes on the study design and practical implementation in appraisal and modelling. This does not preclude exploring these relationships, but the Department may wonder if a new national VTT study is the best place to do so.

## 12. Essential areas to consider beyond ARUP et al. (2015a)

In this chapter we reflect on four areas which deserve additional attention and thereby constitute larger deviates from the ARUP et al. (2015a) study design. Respectively, **Section 12.1** concerns research efforts needed to separate mode- and user effects. **Section 12.2** makes a case for increased use of revealed preference methods. **Section 12.3** concerns eliciting the VTT for active modes of travel, and **Section 12.4** reflects on eliciting the VTT for aviation.

### 12.1 Separation of mode- and user-effects

A key shortcoming associated with a repeat of the ARUP et al. (2015a) study design is that variations in the VTT across modes of transport may be associated both with mode (i.e. comfort) and user effects. In adopting the recommendations of the ARUP et al. (2015a) study, the DfT decided to implement uniform VTTs across modes of transport for commute and other-non work trips. In contrast, in other countries – adopting similar study designs – the appraisal VTT was segmented by modes across all trip purposes and the separation of mode and user effects received limited to no attention in the final study reports.

Two alternative study designs are available to separate mode and user effects. The first approach follows the Norwegian approach (e.g. Flügel et al., 2020) where respondents are asked to complete an additional SP1 (route choice) experiment for an alternative mode to their primary mode of transport. The second approach is based on the German approach (e.g. Axhausen et al., 2014) where mode and user effects are separated by the use of mode choice experiments. We recommend the use of the former approach because it is closer to ARUP et al. (2015a)s way of eliciting the VTT, and because mode choice experiments can be associated with rather strong modal preferences limiting the extent to which analysts learn about the trade-offs between time and cost. That is, if large parts of the population consistently select their preferred mode of transport, there is the risk that the mode specific constants dominate the model implying little is learned regarding the trade-off between travel time and travel costs.

Implementation of the additional SP1 experiment requires respondents to identify relevant alternative modes of transport, and ideally reference time and costs levels, for the journey they were intercepted on. In the context of the current levels of segmentation used, the alternative modes considered should primarily be bus, car, rail, and other PT. Respondents may wish to express active modes of travel or taxis for given trips but, for assigning a relevant alternative SP1 design, we recommend randomly allocating one of the main alternative modes of transport. For car journeys, these would respectively be bus, rail and if relevant other PT.

In relation to the SP design for the alternative mode of transport, the experimental designs can be picked treating the alternative mode as the main mode of transport. To avoid ordering effects, it would be desirable to alternate the order of presenting the SP for the main and the alternative mode of transport. Moreover, we recommend conducting this as a separate data collection effort, thereby avoiding adding the need to combine this additional SP game with SP1 and SP2 (or SP3), which would bring us back in the domain of the three SP game format. In terms of sample sizes, we recommend collecting at least 200 (additional) respondents for each respective mode-purpose combination, allowing for robust sampling of different users and journey characteristics. Joint estimation techniques can be implemented with the main sample to improve the robust estimation of all the covariates of interest.

## 12.2 Revealed preferences

As expressed in **Sections 10.4 and 10.10**, questions remain around the validity of SP-based VTT estimates. It is therefore important that the DfT considers the use of RP-based methods where feasible. The primary objective of such an RP exercise should, in our opinion, be to validate the main SP-based VTT by mode and trip purpose. Whilst significant improvements have been made in the context of RP methods in the past ten years, implementing a drastic change in the data collection method needs to be carefully developed and tested before it can be rolled out at a national scale. It is for the Department to decide how ambitious they wish to be in shifting towards the use of revealed preferences relative to stated preference. As such, two extremes can be envisaged where RP (does not) become(s) the main mode of data collection. In this case, we recommend at least a small scale repeat of ARUP et al. (2015a) to understand the potential sources of differences in the VTT relative to the 2014/15 study. When SP remains the predominant data collection method for the purpose of consistency, the role of RP can be seen as validating and the start of a transition to increased use in future studies.

Following the comprehensive review of RP methods by Flügel et al. (2022), we recommend the DfT to focus its efforts on the use of GPS-based travel diaries where respondents are recording their trips using a fit-for-purpose application on their mobile phone (examples of such applications include Fourstep, rMove and TimeUse+). The benefit of these applications is that individual level travel data is available at a high level of precision, and this data be supplemented by standard survey questions and SP questions. This includes verification questions regarding trip purpose etc. By having full control over the presentation of the additional questions this approach to RP enables the closest replication of the types of models estimated in the SP exercises, especially in terms of the included covariates.

An important first step is the selection or development of a relevant mobile phone application satisfying the needs of the RP data collection but also complying with the DfT's requirements regarding GDPR and data storage and security. The Norwegians decided to develop their own application (Fotefar) to be in control of the of all steps of the data value chain. Especially when making the data available a certain degree of anonymisation will need to take place to avoid being able to identify a user's home location amongst other things.

Although the successful use of GPS-based travel diaries for the estimation of VTT has been proven (e.g. Tsoleridis et al., 2022), the main challenge is identifying strategies to scale its application to the national level. When implemented in regional applications (e.g. Tabasi et al., 2023), a relatively homogenous set of trips in terms of geographical context, distance, time and cost is likely to be captured facilitating the estimation of relevant mode choice or route choice models. In part, the same challenge applies when designing and analysing SP experiments for short- and long-distance journeys, but in RP many more context variables may influence the observed choices potentially reducing the explanatory power of the core variables of interest, i.e. travel time and travel cost. In line with **Sections 10.5 and 10.6**, appropriate recruitment methods and sampling methods are therefore needed to deal with such challenges.<sup>27</sup>

Flügel et al. (2022) also see merit in the use of the National Travel Survey (NTS) or equivalent survey in conjunction with the use of the National Transport Model (NTM). Whilst Fox et al. (2022) illustrate the successful implementation of this approach in the context of the UK, and Ireland. We do not recommend the use of such a method due to the loss of a certain degree of precision, and because the estimated mode-destination choice models are generally used to model forecast demand, not to estimate the headline appraisal VTT. Nevertheless, valuable lessons can be learned from such studies

---

<sup>27</sup> This goes beyond ensuring that strategies are in place ensuring that people without mobile phones, or those with limited confidence in using them, can be represented in such a data collection effort.

in terms of sampling strategies, information contained in RP data itself, and using models like the NTM to get data, particularly travel time and travel cost, on non-chosen alternatives.

When working in a defined spatial region, a well-established transport model like the NTM may be available with sufficient geographical detail to accommodate short-distance trips and with sufficient geographical coverage such that long distance trips can be included in the analysis. The alternative is the use of application programming interfaces (APIs) querying Google or other applications for the travel time on the actual and other available modes of transport. If travel costs are not directly available, functional forms will need to be used to derive these. Again, this approach can build on the experience from using the NTS as a basis for analysis. Notably, ARUP et al. (2015a) made use of the NTS in converting behavioural VTTs to appraisal VTTs and had to make similar calculations to approximate travel costs.

In terms of sample size, the NTS and similar surveys typically encompass many thousands of households, but regional applications of GPS tracking devices often make use of less than 1,000 respondents. The number of reported (i.e. tracked) trips per respondent is, however, larger using mobile phone apps, thereby correcting to some extent for smaller sample sizes. We highly recommend additional research, ideally using Monte Carlo simulations, establishing minimum required sample sizes to estimate minimum viable model specifications (e.g. linear time and cost choice models) and slightly more advanced model specifications (e.g. accounting for cost damping, income, and distance effects).

In terms of timelines, such a desk-based scoping study setting out the requirements and developing a cost specification for an RP study aimed at validating the SP-based VTT estimates could take place within a timescale of 4-6 months.

We consider this an important first step in reinvigorating the role of RP in national VTT research in the UK. If the validation strategy turns out to be successful, there is a case to start prioritising the use of RP again. This could be done by using RP data from GPS-based tracking applications, as set out above, to capture the main VTT and exploit additional information included in the GPS data on the scheduling of journeys to explore the relationship between the VTT and reliability and congestion. This would imply a significant change from current practice and would be associated with a more extensive scoping study to determine which dimensions of the VTT can best be captured by RP and which should be supplemented by SP research.

### 12.3 Active travel

Based on ARUP et al. (2015b) and the broader review of methods and evidence in **Section 7**, an important topic in need of consideration is developing a suitable approach to elicit the VTT for active travel (i.e. walking and cycling). Notably, the Active Mode Appraisal Toolkit does not (yet) cover travel time savings (Department for Transport, 2022b), and Nellthorp (2023) identifies this as one of several areas of improvement. In light of the conclusions made by Nellthorp (2023), we recommend the commissioning of a study deriving the VTT for active travel as part of a separate and broader research programme aimed at improving active mode appraisal, ideally involving a consortium of relevant stakeholders including the Department for Transport, Active Travel England and others. If that is not feasible, additional research on successfully eliciting the VTT for active modes of travel is still recommended.

Whilst active travel model can be embodied as a potential mode of transport in revealed preference mode choice studies, the focus of such a study should primarily be focused on developing suitable stated preference survey formats for active travel. **Section 7.2** highlighted that SP remains the primary

data collection mode to elicit the VTT for active travel, but that there is large heterogeneity in the way these stated preference surveys are designed.

We consider that there is room for improvement to the SP designs used across the reviewed national VTT studies. First, we recommend moving away from the within-mode SP designs tested in ARUP et al. (2015b) because they construct an artificial payment vehicle for active modes of travel which are typically free of charge. This problem can potentially be addressed by moving to a mode choice setting where alternative modes of transport are associated with a travel cost. Our review of methods and evidence indicated that increased levels of electrification in the micro-mobility sector (e.g. e-bikes and e-scooters) may provide basis for realistic alternative options for active travel trips, as opposed to taking the car or bus. Notably, differences in user groups (owners vs. non-owners; and main mode vs access mode) should also inform the specification of the best survey format in alternative journey contexts.

Second, the mode choice SP designs implemented in Germany, Norway and the Netherlands can improve in the characterisation of the active travel mode. Namely, Germany and the Netherlands only described the active mode of travel by the overall travel time, whereas in Norway the active mode of travel was additionally characterised by the features of the provided cycling infrastructure and the number of crossings. The Norwegian design is more desirable because it considers the comfort factor of travelling by active modes of travel. Unfortunately, the estimation of multipliers on the VTT for travelling on different types of infrastructure does not allow isolating the contributions of travel time savings, journey quality, and safety to the choices made. This is considered an important next challenge, especially if safety and health concerns are captured in other parts of the appraisal process and risk double counting. Isolating the importance of each of these benefit categories in mode choices is in our opinion preferred over ex post downscaling the VTT for safety benefits (Flügel et al., 2020) and health benefits (Flügel et al., 2021). This may require being explicit regarding the level of risk exposure in the SP design.

Defining the most suitable specification of the SP design should follow similar steps taken in ARUP et al. (2015b) using qualitative research and cognitive in-depth interviews before running pilot studies. The options set out in Section 5.2 of ARUP et al. (2015b) and the learnings from the actual data collection can still be used as a starting point.

A detailed scoping study on the VTT for active travel should last for at least six, but ideally ten to twelve, months allowing for sequential testing and adjusting of alternative SP formats. The design and success of these SP formats should be informed by a qualitative approach and running a pilot study. Naturally, these timescales need expanding when a broader approach to active travel appraisal is considered.

We recommend research on the VTT for active travel to take place as a standalone exercise, especially because the data collection methods are at present less robust as for the other main modes of travel described in **Section 10**. If the timescales of such an exercise are not too distinct from the main data collection efforts, this should have no impact on the interpretation of the elicited values. As a case in point, the Norwegian study has successfully conducted the data collection for different types of SP games and modes at different points in time without significant implications. If the main approach to the SP study design for the main modes of transport remains based on within-mode route choice experiments, there is little risk of such a standalone exercise to be associated with duplicate costs (e.g. recruitment, analysis etc.). Only when a shift is made to a mode choice approach (see **Section 12.1**) can economies of scale be achieved as active travel can then be integrated in the SP design.

## 12.4 Aviation

Aviation, as a mode of transport, was not considered within the scope of previous national VTT studies in the UK. In contrast, **Table 15** of this scoping study highlights that in several national VTT studies outside of the UK (i.e. Germany, Norway, and the Netherlands) aviation fell within the scope. The UK, however, does have a VTT for aviation and it has direct impact on how the Department conducts forecasting for future aviation demand using the DfT Aviation model.

The aviation modelling suite described by Department for Transport (2024b) refers to specific VTT values for access and egress by surface transport. These values are based on ARUP et al. (2015a) and adjusted for the population travelling by air having (on average) a higher income. Additionally, data from the CAA's passenger survey is used for estimating the choice of airport. The estimated choice model accounts for differences in travel time and fare and can in principle be used for estimating a VTT for aviation. Our understanding is that due to the discontinuation of existing surveys conducted by the CAA, there is an interest in including aviation within the scope of an upcoming national VTT study. Segmentation only accounts for the journey purposes of business and leisure, but the Department has expressed an interest in further segmenting international leisure journeys in a 'holiday' and 'visiting friends and relatives (and other)' segment.

Since the Aviation model does not differentiate the total journey time into sub-components, there is, at present, limited interest in identifying variations in the VTT across these sub-components. For example, the Aviation model considers waiting time to include passing through the airport, and access and egress time. Moreover, flight delays are not considered in long-term forecasting, and no specific reference is made to transfer time (and penalties) at and between airports. The reviewed evidence (see **Section 8**) highlights that eliciting the VTT for aviation and specific travel time components is nevertheless feasible.

Accordingly, the challenge is for the Department to set out its objectives for aviation in the context of a national VTT study. That is, the core objective of such an exercise ought to be the elicitation of a robust VTT estimate for the relevant segments of interest. This may then feed into relevant future demand models as a core parameter. If the overall objective is to estimate future demand (e.g. for choice of airport), and the VTT is a byproduct of this objective the need to conduct this as part of a national VTT study can be questioned.

Stated Preference methods are considered a suitable mechanism to elicit the VTTs of interest for aviation. In this context, the standard within mode route choice setup only trading off travel time and costs (i.e. SP1) may not be the ideal situation, because there is little opportunity to gain in-vehicle travel time without changing the choice of airport at the origin and destination, unless there is an opportunity to reduce taxi times at larger airports. We see more merit in a within mode stated choice survey, alike the Norwegian setup, contrasting travel time and travel costs alongside other attributes such as the number of transfers, and access (or egress) time to (from) the airport. In essence, this approach is not distinct from decomposing the travel time associated with the use of public transport and such methods are proven to be effective. This approach is also suitable for international travel. Another option is to conduct long-distance mode choice experiments for which different travel time components can be specified across the presented modes of transport, as per the German national VTT study design. Notably, as the distance increases the number of feasible alternative modes of transport decreases, especially when considering intercontinental journeys.

For comparability with the other aspects of the national VTT study it may be of interest to focus on domestic in the first place, and it may not be unreasonable to assume that strong but distinct relationships between distance (as a proxy for a variety of effects) and the VTT exist in an international

context, also due to the nature of the journey. As such, the current assumption of a constant VTT for all journeys other than trip purpose (leisure vs. business) may need revising.

A further question is the extent to which an aviation component of the VTT study can be treated as a separate research track with its own time scales and requirements, or whether it may benefit from including in the main study and associated scale benefits. The more the scoping of such a study focuses on travel time components occurring at the airport, the more independent the design of such a study becomes relative to other modes of interest. Moreover, when opting for an intercept sampling approach significant efforts are required in obtaining the relevant permissions to sample in and around the airport and airport stations and parking, let alone sample behind customs. As this requires additional survey teams being stationed at carefully selected locations an independent track is relevant. When the intercept sampling approach is abandoned, these considerations are no longer relevant and specific algorithms need defining to allocate 'air' as the main mode of transport in the distribution of the survey.<sup>28</sup>

Relative to the other essential areas, designing and implementing an SP study for aviation is less complicated and associated with lower degrees of risk. Therefore, we only recommend a short pre-study of 3 up to 6 months, largely driven by the DfT itself, setting out the requirements for such a study and pre-testing initial SP designs. The main data collection effort can then take place alongside the main study.

## 13. Exploratory areas of research

In this section we set out two areas of research which are important, but not essential for running a new national VTT study. The first area builds on recent research projects on the VTT for congestion and reliability and specifically explores the relationship between these two concepts. The second area reflects on the use of long-term SP games.

### 13.1. Relationship between congestion and reliability

Whilst the review for methods and evidence established that the mean-standard deviation approach to the value of travel time reliability (VTTR) used in SP2 is considered best practice from both a valuation and transport modelling perspective (see **Section 6.2**), the story is entirely different for the treatment of congestion (see **Section 6.1**). Namely, the categorical description of congestion in SP3 is considered to hamper the implementation of congestion multipliers in valuation (i.e. treatment of travel conditions) and in transport modelling (i.e. preference for a free flow + delay approach).

Importantly, this does not mean that reducing the discomfort associated with congestion should not be considered an important benefit of congestion relief schemes. The emerging challenge is, however, that there is a clear connection between reliability and congestion and additional research efforts are needed to examine the extent to which both metrics can co-exist in the appraisal framework without running the risk of double counting.

We recommend a research programme starting with a desk-based review of alternative theoretical measures of reliability and congestion explicitly covering the overlap between these measures. This review should build on earlier work by ITS (2022) and result in crisp definitions of which aspects of congestion are attributed to reliability and comfort factors respectively. Subsequently, a limited set of potential metrics should be reviewed in the context of transport assignment modelling where the core question is whether the current preference for a mean-standard deviation approach solely capturing

---

<sup>28</sup> The DfT recommends avoiding airside surveys to avoid complications, delays and costs despite these producing higher quality outputs.



reliability should be overturned. Having established suitable metrics for congestion and reliability, the next stage of the research project should then focus on how these metrics can best be represented in SP games.

Given that a series of projects has already been conducted on congested values of travel time (CVTT), we should highlight that this approach is associated with high risks. Although research into reliability and congestion has received significant attention over the past decades, this issue has not been satisfactorily resolved. Moreover, it may very well be that no feasible solutions are available since the mean-standard deviation is already considered a second-best approach from a modelling perspective because the standard deviation is not link-additive. Perhaps a good starting point is theoretically contrasting the mean-standard deviation approach against the free flow time – delay approach to get a sense of overlap.

### 13.2. Long term SP choices

The German national VTT study (Axhausen et al., 2014) presented alongside its short term mode choice SP questions two long term SP questions eliciting the extent to which changes in travel time influence residential and job location choices respectively. An important consideration here is that the provision of travel time savings leads to additional changes in behaviour beyond mode and route choices, and that the welfare benefits can therefore also be experienced by obtaining access to better jobs and houses. This broadening of the scope is of interest to the VTT literature, but it is only in its early stages of development. In its current application, simple SP games are used which not only encounter challenges as to how sensitivities to changes in different cost components (e.g. housing prices, salaries, and travel costs) should be dealt with to arrive at a robust VTT estimate, but this change in perspective also needs alignment with the existing appraisal framework which is predominantly trip based and hence aligned with the short term VTT perspective.

By adopting a long-term perspective, one recognises that benefits from transport investments may extend (or dissipate) to other markets than the transport market, and this needs embodying in the appraisal framework itself, potentially in the form of Spatial Computational General Equilibrium Models (S-CGGE)<sup>29</sup>, or Land Use and Transport Interaction (LUTI) models<sup>30</sup>. It may very well be the case that in the short-term benefits are experienced through the transport market, but that these get replaced by changes in residential and job location. The extent to which long term SP games can provide meaningful information which can be fed into the appraisal framework thereby deserves more attention as part of a broader stream of research aligning short- and long-term benefits from transport investments.

## 14. Summary of recommendations

To summarise, we recommend the Department to start planning for a new national VTT (and VTTR) study. **The study design used by ARUP et al. (2015a) is still considered international best practice and can form the basis for a new national VTT study subject to some adjustments summarised below.** The evidence base for the VTT adopted in TAG is now 10 years old and notable changes in travel behaviour and background factors have taken place during the intervening years. We feel that there are limits to the degree of assurance that can be given, when applying the values from the 2014-15 study to policy and scheme analysis in 2025 and onwards.

---

<sup>29</sup> <https://www.gov.uk/government/publications/spatial-computable-general-equilibrium-s-cge-model>

<sup>30</sup> <https://assets.publishing.service.gov.uk/media/6707bf8730536cb927482f91/dft-review-land-use-transport-interaction-models.pdf>



Before setting out the requirements of a new national VTT study, the Department is recommended to review its desired level of segmentation of the VTT. We consider the present level of segmentation by mode and journey purpose combination to be appropriate and aligned with international best practice. We have expressed that further segmenting the VTT by, for example, distance, geography, and (or) income results in an intermediate solution between 'appraisal' and 'behavioural' VTTs with potential implications for transport modelling, interpretation, and policy outcomes. A deeper exploration of segmentation by, for example, distance is feasible in a new study, even if implementation remains unclear, but requires a more detailed specification of sampling strata which will inherently increase study costs.

In terms of research to be conducted ahead of commissioning a new national VTT study, we recommend three specific research programmes:

1. **Developing a suitable approach to using GPS-based tracking applications to validate the SP-based VTT estimates using revealed preference methods.** Whilst the SP methods recommended for the main study are considered robust and reflect international best practice, there remains a desire to validate the hypothetical and abstract nature of SP with RP methods. GPS-based tracking apps supplemented with survey questions provide a suitable approach to this. However, there is no precedent for implementing these applications in the context of a national VTT study. We therefore recommend the commissioning of a short-term desk-based scoping study of up to 6 months specifying the technical, sampling, and data processing requirements for successfully replacing the within-mode RP validation approach previously adopted in ARUP et al. (2015a). If the validation exercise is successful, follow-up research is recommended exploring the potential for gradually reducing the dominance of SP in national VTT studies (and possibly replacing SP in the long run).
2. **Developing a suitable approach to elicit the VTT for active modes of travel.** With increased strategic interests in promoting active travel, it is essential that business cases can be strengthened using monetary estimates of time savings and journey quality related benefits. Building on the work conducted by ARUP et al. (2015b) and making use of new micromobility options (e-bikes and e-scooters) presently available to transport users, we believe that more viable SP designs can be developed relative to ten years ago. The timelines associated with such a scoping study are recommended to take between 10-12 months – to allow the conduct of in-depth qualitative interviews and pilot studies to improve the study design before rolling this out at a large scale. Since active travel is separate from the mechanised modes of transport, the Department can consider running this concurrently or separately from the main VTT study like its treatment of the recent freight VTT study.
3. **Developing a suitable approach to elicit the VTT for aviation.** There is value for the DfT from including aviation within the scope of the national VTT study as it has a direct impact on how the department conducts forecasting for future aviation demand, including surface access, using the DfT Aviation model. To our understanding the interest to expand the scope of the national VTT study to aviation is driven by the CAA has stopped collecting the necessary survey data previously used to estimate the VTT for aviation. At present, only a segmentation is made between business and leisure trips and the VTT applies to all components of the journey. It is recommended that the Department sets out the desired level of segmentation, including the prominence of international travel vs domestic travel, and that the data collection is designed around these specifications. Alike the VTT for active modes of travel, we recommend this as a standalone stream of

research. Given the feasibility of developing relevant SP designs, this pre-study stage should only last for a maximum of six months.

In the context of the above considerations, **we recommend building on the design of the ARUP et al. (2015a) study subject to several distinct but relatively minor changes and points of attention:**

- *For business VTT:* discontinue the employer's WTP approach and instead include the Hensher Equation within the scope of a new national VTT study.
- *In relation to congestion:* continue the SP3 car games presenting different categorical levels of congestion but consider alternative options to improve the SP design with the ultimate goal of facilitating implementation in appraisal.
- *In relation to crowding:* replace the SP3 bus and other PT games presenting different categorical levels of crowding for rail-based crowding levels presenting occupation levels in terms of pax/m<sup>2</sup> alongside the ability to sit during travelling.
- *Change the number of SP games per person from three to two.* This implies always presenting SP1 with either SP2 (reliability) or SP3 (journey quality) instead of presenting all three types of SP games. This should not influence sampling requirements.
- *Revisit the SP1 designs and consider a mixture of d-efficient designs and the Scandinavian approach.* The Dutch experience shows good promise in relation to dealing with non-trading behaviour.
- *In relation to recruitment:* it is essential to revisit the sampling approach and define all relevant strata such that recruited respondents robustly cover all relevant travel patterns and socio-economic characteristics. Particularly the sampling of shorter distances should receive more attention.
- *In relation to incentives:* consider the extent to which appropriate incentive structures may overcome known self-selection biases of participating in a national VTT study.
- *In relation to sampling targets:* we recommend giving equal importance to bus and other PT as to car and rail in terms of the sampling targets. This slightly increases overall sampling targets.

To address a primary concern associated with the ARUP et al. (2015a) design, i.e. the separation of mode- and user effects, **we recommend collecting additional SP data where respondents are presented with two SP1 type of games respectively covering their primary (or intercepted) mode of transport and a relevant alternative mechanised mode of transport.** Since the same designs can be implemented this requires only minimal additional implementation effort.

The above recommendations all relate to the design and collection of the SP surveys. It goes without saying that despite the ARUP et al. (2015a) providing a strong starting point, **the survey and particularly the SP games need to be subject to a rigorous range of review and testing stages both qualitatively and quantitatively**, as one would expect for a study of such scale and significance. Moreover, **we recommend that 5-10% of the data collection efforts provide an exact repeat of the 2014/15 SP1 design for purposes of continuity.**

The data collected through the SP surveys form the basis for the development of behavioural models explaining the variation in the VTT. In our opinion, the analysis conducted on the ARUP et al. (2015a) data is still state-of-the-art, but there is no requirement to strive to exactly replicate these model specifications. **A new study team and the DfT are encouraged to transparently discuss considerations for (not) adopting specific model features thereby finding a balance between optimising the fit of the model and theoretical considerations regarding the VTT.** New modelling techniques such as

machine learning can be used to inform the specification of the utility functions implemented in the behavioural VTT models.

The translation of behavioural VTTs into appraisal VTTs need a similar transparent conversation between the study team and the DfT. In common with ARUP et al. (2015a) and other national studies, the sample enumeration approach is recommended for this purpose but whether this is applied to the surveyed population or an external dataset, such as the NTS, depends on the representativeness of the collected sample and the extent to which the explanatory variables included in the behavioural VTT models are included in such data. Decisions regarding the appropriate (e.g. distance) weights and the treatment of income are policy-driven but need articulating in the final reporting.

We recommend the DfT to spend at least 6 months on a research project scoping out its potential approach to RP for validation purposes ahead of commissioning the full survey. If active travel is considered an essential part of the main study, we recommend extending this pre-study period by roughly another 6 months enabling the development an appropriate approach to the valuation of travel time savings for active modes. Within this period, also a pre-study into the developing an approach for the valuation of travel times savings for aviation should be considered. These three strands of research require different skill sets and can be commissioned as distinct pieces of work. We recommend all three to take place ahead of the main survey to judge feasibility of including within scope of the main study.

We believe that the options set out above provide the DfT with a robust approach to updating the evidence base for the VTT as central input in its Transport Analysis Guidance. In terms of the main study, we recommend allowing for an overall period of 18-24 months. The project would cover four stages, respectively i) study design, testing & piloting, ii) main data collection, iii) data analysis & development of behavioural VTT models, and iv) translation into appraisal VTT and dissemination and outreach.

Finally, two exploratory areas of research have been identified which are of interest to the wider VTT literature, but not essential for a next national VTT study. These respectively refer to the co-existence of congestion and reliability in the appraisal framework, and the use of long term SP choices. We are confident with the current approach to reliability both from a valuation and a modelling perspective. The roadmap for congestion is, however, less clear and may need additional research impetus to study the intricate relationship between these two dimensions and how reliability and congestion can best be represented to avoid potential double counting. Whilst exploratory studies have found that the VTT derived from long term SP choices are higher than from short term SP choices, there are important questions to be addressed as to how long-term SP choices are measuring the relevant concept of interest, and how this can be integrated into the broader appraisal framework.

In Table 39 below we set out our envisaged timeline to a new national VTT study where the dark green cells indicate the core components of ARUP et al. (2015a) and the light green cells indicate additional features following our review of methods and evidence as described in this report. Our recommendation is to consider the top three rows as part of the main study, and the final two rows as standalone streams of work.

Table 39: Potential timetable of a new national VTT study

|            |                                      | Pre-study                                                        |      | Commissioning | Main study                                                |                 |                              |                                 |
|------------|--------------------------------------|------------------------------------------------------------------|------|---------------|-----------------------------------------------------------|-----------------|------------------------------|---------------------------------|
|            |                                      | M1-6                                                             | M6-9 | M9-12         | M12-M18                                                   | M18-M24         | M24-M30                      | M30-M36                         |
| Main study | Improving ARUP et al. (2015a)        |                                                                  |      |               | Development, review and testing of main survey            | Data collection | Data analysis                | Appraisal outputs and reporting |
|            | Separation of mode- and user effects |                                                                  |      |               | Testing the separation of mode- and user-effects approach | Data collection | Data analysis                | Appraisal outputs and reporting |
|            | Revealed Preference                  | Scoping of RP options using a GPS-based application              |      |               | Development, review and testing of RP survey              | Data collection | Data analysis and validation | Appraisal outputs and reporting |
|            |                                      |                                                                  |      |               |                                                           |                 |                              |                                 |
| Standalone | Active travel                        | Developing and testing of suitable SP approach for active travel |      |               | Commissioning                                             | Data collection | Data analysis                | Appraisal outputs and reporting |
| Standalone | Aviation                             | Setting out requirements and scoping SP options for aviation     |      | Commissioning | Developing and testing of suitable SP approach            | Data collection | Data analysis and validation | Appraisal outputs and reporting |

## 15. References

- ABEILLE, A., PAWLAK, J. & SIVAKUMAR, A. 2022. Exploring the meaning and drivers of personal (Un-)Productivity of knowledge workers in mobile settings. *Travel Behaviour and Society*, 27, 26-37.
- ABEILLE, A., PAWLAK, J., SIVAKUMAR, A., MOLIN, E. & CHORUS, C. G. 2024. What motivates people to work while travelling and what determines the associated productivity? *104th Annual Meeting of Transportation Research Board*. Washington D.C.
- ABRANTES, P. A. L. & WARDMAN, M. R. 2011. Meta-analysis of UK values of travel time: An update. *Transportation Research Part A: Policy and Practice*, 45, 1-17.
- ACCENT & HAGUE CONSULTING GROUP 1999. The value of travel time on UK roads.
- AECOM, ARUP, SIGNIFICANCE & ITS 2023. Freight value of time and value of reliability: Final Report.
- ALGERS, S., DILLÉN, J. L. & WILDERT, S. 1995. The national Swedish value of time study. *Proceedings from the Annual Transport Conference at Aalborg University*, 2.
- ALI, A., DEKKER, T., HESS, S. & CHOUDHURY, C. 2024. Using posterior analysis to predict missing information in passively collected data. *hEART*. Aalto.
- ALONSO-GONZÁLEZ, M. J., CATS, O., VAN OORT, N., HOOGENDOORN-LANSER, S. & HOOGENDOORN, S. 2021. What are the determinants of the willingness to share rides in pooled on-demand services? *Transportation*, 48, 1733-1765.
- ANDERSSON, A. 2024. *Modelling long-distance travel demand by combining mobile phone and survey data*. Doctoral thesis, comprehensive summary, Linköping University Electronic Press.
- ARUP, AECOM, ITS & SIGNIFICANCE 2023. Freight Value of Time and Value of Reliability.
- ARUP & ITS 2017. Programme for maintaining a robust valuation of travel time savings: feasibility study - Final Phase 1 report: List of options.
- ARUP & ITS 2018. Programme for maintaining a robust valuation of travel time savings: feasibility study - Final Phase 2 Report.
- ARUP, ITS & ACCENT 2015a. Provision of market research for value of travel time savings and reliability: phase 2 report.
- ARUP, ITS & ACCENT 2015b. Provision of market research for value of travel time savings and reliability: walk and cycle report.
- ASCHAUER, F., RÖSEL, I., HÖSSINGER, R., KREIS, H. B. & GERIKE, R. 2019. Time use, mobility and expenditure: an innovative survey design for understanding individuals' trade-off processes. *Transportation*, 46, 307-339.
- AXHAUSEN, K., KÖNIG, A., ABAY, G., BATES, J. & BIERLAIRE, M. 2004. Swiss value of travel time savings. *European Transport Conference*. Strassbourg.
- AXHAUSEN, K. W., EHREKE, I., GLEMSER, A., HESS, S., JÖDDEN, C., NAGEL, K., SAUER, A. & WEIS, C. 2014. Ermittlung von Bewertungsansätzen für Reisezeiten und Zuverlässigkeit auf Basis der Schätzung eines Modells für modale Verlagerungen im nicht-gewerblichen und gewerblichen Personenverkehr für die Bundesverkehrswegeplanung. FE-Projekt 96.996/2011 Zeitkosten Personenverkehr. Entwurf Schlussbericht. IVT, ETH Zürich; TNS Infratest.
- AXHAUSEN, K. W., HESS, S., KÖNIG, A., ABAY, G., BATES, J. J. & BIERLAIRE, M. 2008. Income and distance elasticities of values of travel time savings: New Swiss results. *Transport Policy*, 15, 173-185.
- BAEK, K., LEE, H., CHUNG, J.-H. & KIM, J. 2021. Electric scooter sharing: How do people value it as a last-mile transportation mode? *Transportation Research Part D: Transport and Environment*, 90, 102642.
- BANARJEE, I. & KANAFANI, A. 2008. The Value of Wireless Internet Connection on Trains: Implications for Mode-Choice Models. UC Berkeley: University of California Transportation Center.
- BATES, J., POLAK, J., JONES, P. & COOK, A. 2001. The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review*, 37, 191-229.
- BATLEY, R. 2015. The Hensher equation: derivation, interpretation and implications for practical implementation. *Transportation*, 42, 257-275.

- BATLEY, R., BATES, J., BLIEMER, M., BÖRJESSON, M., BOURDON, J., CABRAL, M. O., CHINTAKAYALA, P. K., CHOUDHURY, C., DALY, A., DEKKER, T., DRIVYLA, E., FOWKES, T., HESS, S., HEYWOOD, C., JOHNSON, D., LAIRD, J., MACKIE, P., PARKIN, J., SANDERS, S., SHELDON, R., WARDMAN, M. & WORSLEY, T. 2019. New appraisal values of travel time saving and reliability in Great Britain. *Transportation*, 46, 583-621.
- BATLEY, R. & DEKKER, T. 2019. The Intuition Behind Income Effects of Price Changes in Discrete Choice Models, and a Simple Method for Measuring the Compensating Variation. *Environmental and Resource Economics*, 74, 337-366.
- BATLEY, R., DEKKER, T. & STEAD, I. 2020. Worthwhile Use of Travel Time and Applications in the United Kingdom. *International Transport Forum Discussion Papers*.
- BATLEY, R., WARDMAN, M. & STEAD, I. 2023. New theoretical and empirical insights on valuing business travel time savings using the Hensher Equation.
- BECK, M. J., HESS, S., CABRAL, M. O. & DUBERNET, I. 2017. Valuing travel time savings: A case of short-term or long term choices? *Transportation Research Part E: Logistics and Transportation Review*, 100, 133-143.
- BECKER, G. S. 1965. A Theory of the Allocation of Time. *The Economic Journal*, 75, 493-517.
- BEESLEY, M. E. 1965. The Value of Time Spent in Travelling: Some New Evidence. *Economica*, 32, 174-185.
- BINSUWADAN, J., WARDMAN, M., DE JONG, G., BATLEY, R. & WHEAT, P. 2023. The income elasticity of the value of travel time savings: A meta-analysis. *Transport Policy*, 136, 126-136.
- BJÖRKLUND, G. & MORTAZAVI, R. 2013. Influences of infrastructure and attitudes to health on value of travel time savings in bicycle journeys. *Working papers in transport economics*. Stockholm: Centre for transport studies.
- BLIEMER, M. C. J. & ROSE, J. M. 2024. Chapter 7: Designing and conducting stated choice experiments
- Handbook of Choice Modelling. Cheltenham, UK: Edward Elgar Publishing.
- BÖRJESSON, M. & ELIASSON, J. 2012. The value of time and external benefits in bicycle appraisal. *Transportation Research Part A: Policy and Practice*, 46, 673-683.
- BÖRJESSON, M. & ELIASSON, J. 2014. Experiences from the Swedish Value of Time study. *Transportation Research Part A: Policy and Practice*, 59, 144-158.
- BÖRJESSON, M. & ELIASSON, J. 2019. Should values of time be differentiated? *Transport Reviews*, 39, 357-375.
- BÖRJESSON, M., ELIASSON, J. & FRANKLIN, J. P. 2012a. Valuations of travel time variability in scheduling versus mean-variance models. *Transportation Research Part B: Methodological*, 46, 855-873.
- BÖRJESSON, M., FOSGERAU, M. & ALGERS, S. 2012b. Catching the tail: Empirical identification of the distribution of the value of travel time. *Transportation Research Part A: Policy and Practice*, 46, 378-391.
- BÖRJESSON, M., KOUWENHOVEN, M., DE JONG, G. & DALY, A. 2023. Can repeated surveys reveal the variation of the value of travel time over time? *Transportation*, 50, 245-284.
- BOUNIE, N., ADOUE, F., KONING, M. & L'HOSTIS, A. 2019. What value do travelers put on connectivity to mobile phone and Internet networks in public transport? Empirical evidence from the Paris region. *Transportation Research Part A: Policy and Practice*, 130, 158-177.
- BUCHHOLD, N., DOVAL, L., KASTL, J., MATEJKA, F. & SALZ, T. 2024. Personalized pricing and the value of time: evidence from auctioned cab rides.
- BWAMBALE, A., CHOUDHURY, C. F. & HESS, S. 2019. Modelling departure time choice using mobile phone data. *Transportation Research Part A: Policy and Practice*, 130, 424-439.
- CALASTRI, C., PAWLAK, J. & BATLEY, R. 2022. Participation in online activities while travelling: an application of the MDCEV model in the context of rail travel. *Transportation*, 49, 61-87.
- CARRUTHERS, R. & HENSHER, D. 1976. Resource value of business air travel time. *Modal Choice and the Value of Travel Time*. Clarendon Press, Oxford.

- CELLINA, F., VITTUCCI MARZETTI, G. & GUI, M. 2021. Self-selection and attrition biases in app-based persuasive technologies for mobility behavior change: Evidence from a Swiss case study. *Computers in Human Behavior*, 125, 106970.
- CHOI, S., KIM, D., KO, J., PARK, J. & KO, Y. 2023. Value of travel time savings of autonomous vehicle commuters: a segmented valuation for local and inter-city travel. *International Journal of Urban Sciences*, 27, 645-669.
- COMOUK 2023a. CoMoUK Annual Bike Share Report UK.
- COMOUK 2023b. CoMoUK report on the shared e-scooter trials in England - April 2023.
- CORNET, Y., LUGANO, G., GEORGOULI, C. & MILAKIS, D. 2022. Worthwhile travel time: a conceptual framework of the perceived value of enjoyment, productivity and fitness while travelling. *Transport Reviews*, 42, 580-603.
- CORREIA, G. H. D. A., LOOFF, E., VAN CRANENBURGH, S., SNELDER, M. & VAN AREM, B. 2019. On the impact of vehicle automation on the value of travel time while performing work and leisure activities in a car: Theoretical insights and results from a stated preference survey. *Transportation Research Part A: Policy and Practice*, 119, 359-382.
- DALY, A. & HESS, S. 2020. VTT or VTTS: a note on terminology for value of travel time work. *Transportation*, 47, 1359-1364.
- DE BORGER, B. & FOSGERAU, M. 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.
- DE JONG, G. & KOUWENHOVEN, M. 2020a. Chapter Two - Value of travel time and travel time reliability. In: MOUTER, N. (ed.) *Advances in Transport Policy and Planning*. Academic Press.
- DE JONG, G. & KOUWENHOVEN, M. 2020b. Time Use and Values of Time and Reality in the Netherlands. *International Transport Forum Discussion Papers*.
- DE JONG, G. C. & BLIEMER, M. C. J. 2015. On including travel time reliability of road traffic in appraisal. *Transportation Research Part A: Policy and Practice*, 73, 80-95.
- DEKKER, T., KOSTER, P. & BATLEY, R. 2022. Welfare analysis when income and prices are included in discrete choice models. *7th International Choice Modelling Conference*. Reykjavik, Iceland.
- DEPARTMENT FOR TRANSPORT 2015a. Understanding and valuing impacts of transport investment - Values of travel time savings.
- DEPARTMENT FOR TRANSPORT 2015b. Understanding and Valuing Impacts of Transport Investment: Values of Travel Time Savings. London: Department for Transport.
- DEPARTMENT FOR TRANSPORT 2016. Understanding and valuing impacts of transport investment: Value of travel time savings consultation response.
- DEPARTMENT FOR TRANSPORT 2019. Future of mobility: urban strategy.
- DEPARTMENT FOR TRANSPORT 2022a. TAG unit A1-3 user and provider impacts.
- DEPARTMENT FOR TRANSPORT 2022b. TAG unit A5-1 active mode appraisal.
- DEPARTMENT FOR TRANSPORT 2022c. TAG Unit A.1.3 User and Provider Impacts. In: DEPARTMENT FOR TRANSPORT (ed.) *Transport Appraisal Guidance*. London: Department for Transport.
- DEPARTMENT FOR TRANSPORT 2023a. £2 bus fare cap evaluation: interim report February 2023.
- DEPARTMENT FOR TRANSPORT 2023b. Mobility as a Service: code of practice.
- DEPARTMENT FOR TRANSPORT 2023c. National Travel Survey 2022: Mode share, journey lengths and trends in public transport use
- DEPARTMENT FOR TRANSPORT 2023d. TAG spending objective analysis.
- DEPARTMENT FOR TRANSPORT 2023e. TAG uncertainty toolkit.
- DEPARTMENT FOR TRANSPORT 2023f. TAG unit A5-2 aviation appraisal.
- DEPARTMENT FOR TRANSPORT 2024a. £2 bus fare cap.
- DEPARTMENT FOR TRANSPORT 2024b. DfT Aviation Modelling Suite.
- DEPARTMENT FOR TRANSPORT 2024c. Domestic Transport Usage by Mode.
- DEPARTMENT FOR TRANSPORT 2024d. E-scooter trials: guidance for local authorities and rental operators
- DEPARTMENT FOR TRANSPORT 2024e. NTS 2023: Notes and definitions.

- DEPARTMENT FOR TRANSPORT 2024f. TAG data book.
- DEPARTMENT FOR TRANSPORT 2024g. TAG unit A4-2 distributional impact appraisal.
- DEPARTMENT FOR TRANSPORT 2024h. TAG Unit A5.3 - Rail Appraisal.
- DEPARTMENT FOR TRANSPORT 2024i. TAG unit M2-1 variable demand modelling.
- DEPARTMENT FOR TRANSPORT 2024j. TAG Unit M3.1 Highway Assignment Modelling. *In*: DEPARTMENT FOR TRANSPORT (ed.) *Transport Appraisal Guidance*. London: Department for Transport.
- DEPARTMENT FOR TRANSPORT 2024k. TAG unit M3.2 - Public Transport Assignment Modelling.
- DEPARTMENT FOR TRANSPORT 2024l. TAG unit M5-4 agent-based methods and activity-based demand modelling.
- DEPARTMENT FOR TRANSPORT 2024m. Value for Money Framework. *In*: DEPARTMENT FOR TRANSPORT (ed.).
- DEPARTMENT FOR TRANSPORT 2024n. Zero Emission Bus Regional Areas (ZEBRA) programmes funding amounts
- DESERPA, A. C. 1971. A Theory of the Economics of Time. *The Economic Journal*, 81, 828-846.
- DOUGLAS, B. D., EWELL, P. J. & BRAUER, M. 2023. Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLOS ONE*, 18, e0279720.
- DUBERNET, I. 2019. *Understanding the Value of Travel Time.. Advanced modelling techniques applied to the National German Value of Travel Time and Travel Time Reliability Study*. ETH Zurich.
- DUBERNET, I. & AXHAUSEN, K. W. 2020. The German value of time and value of reliability study: the survey work. *Transportation*, 47, 1477-1513.
- DUBERNET, I., DUBERNET, T. & AXHAUSEN, K. W. 2020. Comparing values of travel time obtained from workplace and short-term decisions. *Travel Behaviour and Society*, 20, 83-90.
- EHREKE, I., HESS, S., WEIS, C. & AXHAUSEN, K. W. 2015. Reliability in the German Value of Time Study. *Transportation Research Record*, 2495, 14-22.
- ENOCH, M. 2018. Mobility as a Service (MaaS) in the UK: change and its implications.
- ESZTERGÁR-KISS, D., TORDAI, D. & LOPEZ LIZARRAGA, J. C. 2022. Assessment of travel behavior related to e-scooters using a stated preference experiment. *Transportation Research Part A: Policy and Practice*, 166, 389-405.
- ETH. 2024. *New Swiss value of times for passenger transport* [Online]. Available: <https://www.ivt.ethz.ch/en/vpl/projects/vss-2017-119.html> [Accessed].
- EUROSTAT 2019. Harmonised European Time Use Surveys (HETUS) 2018 Guidelines.
- FEDUJWAR, R. & AGARWAL, A. 2024. A systematic review on crowding valuation in public transport. *Public Transport*.
- FICKLING, R., GUNN, H., KIRBY, H. R., BRADLEY, M. & HEYWOOD, C. 2009. Productive use of rail travel time and the valuation of travel time savings for rail business travellers.
- FLÜGEL, S. 2014. Accounting for user type and mode effects on the value of travel time savings in project appraisal: Opportunities and challenges. *Research in Transportation Economics*, 47, 50-60.
- FLÜGEL, S., HALSE, A. H., HULLEBERG, N., JORDBAKKE, G. N., VEISTEN, K., SUNDFØR, H. B. & KOUWENHOVEN, M. 2020. Value of Travel Time and Related Factors: Technical Report, the Norwegian Valuation Study 2018-2020.
- FLÜGEL, S. & HULLEBERG, N. 2023. Aversion to In-vehicle Crowding before, during and after the COVID-19 pandemic – an update with post-COVID data from autumn 2022.
- FLÜGEL, S. & MADSLIEN, A. 2020. The Value of Time as a continuous function of travel distance. *TOI*.
- FLÜGEL, S., VEISTEN, K., SUNDFØR, H. B., JORDBAKKE, G. N., HULLEBERG, N. & HALSE, A. H. 2021. The effect of health benefits on the value of travel time savings in active transport. *Journal of Transport & Health*, 21, 101074.
- FLÜGEL, S., WEBER, C., HALSE, A. H. & ELLIS, I. O. 2022. Valuation based on Big Data and revealed preference data - An assessment for Norwegian transport appraisal.



- FOSGERAU, M. 2019. Automation and the Value of Time in Passenger Transport. *Roundtable 176*. International Transport Forum.
- FOSGERAU, M. 2024. Chapter 11: Nonparametric approaches to describing heterogeneity Handbook of Choice Modelling. Cheltenham, UK: Edward Elgar Publishing.
- FOSGERAU, M. & BIERLAIRE, M. 2009. Discrete choice models with multiplicative error terms. *Transportation Research Part B: Methodological*, 43, 494-505.
- FOSGERAU, M., HJORTH, K. & LYK-JENSEN, S. V. 2010. Between-mode-differences in the value of travel time: Self-selection or strategic behaviour? *Transportation Research Part D: Transport and Environment*, 15, 370-381.
- FOSGERAU, M. & MABIT, S. L. 2013. Easy and flexible mixture distributions. *Economics Letters*, 120, 206-210.
- FOSGERAU, M., MCFADDEN, D. & BIERLAIRE, M. 2013. Choice probability generating functions. *Journal of Choice Modelling*, 8, 1-18.
- FOUNTAS, G., FONZONE, A. & OLOWOSEGUN, A. 2023. The Impact of the Cost-Of-Living Crisis on Travel Choices of Scottish Residents: An Exploratory Analysis.
- FOWKES, A. S. 2001. Principles of valuing business travel time savings. University of Leeds.
- FOWKES, A. S., MARKS, P. & NASH, C. A. 1986. The Value of Business Travel Time Savings. University of Leeds.
- FOX, J., Ó SÚILLEABHÁIN, P. & VBRZAIL, W. 2022. Estimating values of time from revealed preference data.
- FRÖHLICH, P., WEIS, C., ERATH, A., VRTIC, M. & AXHAUSEN, K. W. 2013. SP-Befragung 2010 zum Verkehrsverhalten im Personenverkehr. IVT, ETH Zürich.
- FU, M., ROTHFELD, R. & ANTONIOU, C. 2019. Exploring Preferences for Transportation Modes in an Urban Air Mobility Environment: Munich Case Study. *Transportation Research Record*, 2673, 427-442.
- GINKEL, J. V. 2014. The value of time and comfort in bicycle appraisal : a stated preference research into the cyclists' valuation of travel time reductions and comfort improvements in the Netherlands.
- GOLDSZMIDT, A., LIST, J. A., METCALFE, R. D., MUIR, I., SMITH, V. K. & WANG, J. 2020. The Value of Time in the United States: Estimates from Nationwide Natural Field Experiments. *National Bureau of Economic Research Working Paper Series*, No. 28208.
- GREAVES, S., COBBOLD, A., STANESBY, O., SHARMAN, M., JOSE, K., EVANS, J. & CLELAND, V. 2023. Who stays and who plays? Participant retention and smartphone app usage in a longitudinal travel survey.
- GUNN, H. & ROHR, C. 1996. The 1985–1996 Dutch Value of Time Studies. *PTRC International Conference on the Value of Time*.
- HAGUE CONSULTING GROUP 1990. The Netherlands' value of time study: final report.
- HAGUE CONSULTING GROUP 1998. The second Netherlands' value of time study: final report.
- HALSE, A. H., FLÜGEL, S., KOUWENHOVEN, M., DE JONG, G., SUNDFØR, H. B., HULLEBERG, N., JORDBAKKE, G. N. & LINDHJEM, H. 2023. A minute of your time: The impact of survey recruitment method and interview location on the value of travel time. *Transportation*, 50, 1553-1584.
- HARRISON, A. J. 1974. *The Economics of Transport Appraisal*, Routledge.
- HARTWIG, L., GÜHNEMANN, A. & HÖSSINGER, R. 2024. Decomposing mode-specific values of travel time savings with respect to different levels of travel-based multitasking: A revealed preference study. *Travel Behaviour and Society*, 34, 100700.
- HENSHER, D. 1977. *Value of business travel time*, Pergamon Press, Incorporated.
- HENSHER, D. A. 2004. Identifying the Influence of Stated Choice Design Dimensionality on Willingness to Pay for Travel Time Savings. *Journal of Transport Economics and Policy (JTEP)*, 38, 425-446.

- HENSHER, D. A. & BRADLEY, M. 1993. Using stated response choice data to enrich revealed preference discrete choice models. *Marketing Letters*, 4, 139-151.
- HESS, S. 2024. Chapter 14: Latent class structures: taste heterogeneity and beyond  
Handbook of Choice Modelling. Cheltenham, UK: Edward Elgar Publishing.
- HESS, S., DALY, A. & BÖRJESSON, M. 2020. A critical appraisal of the use of simple time-money trade-offs for appraisal value of travel time measures. *Transportation*, 47, 1541-1570.
- HESS, S., DALY, A., DEKKER, T., CABRAL, M. O. & BATLEY, R. 2017a. A framework for capturing heterogeneity, heteroskedasticity, non-linearity, reference dependence and design artefacts in value of time research. *Transportation Research Part B: Methodological*, 96, 126-149.
- HESS, S., MURPHY, P., LE, H. & LEONG, W. Y. 2017b. Estimation of New Monetary Valuations of Travel Time, Quality of Travel, and Safety for Singapore. *Transportation Research Record*, 2664, 79-90.
- HMT 2024. The Green Book.
- HÖSSINGER, R., ASCHAUER, F., JARA-DÍAZ, S., JOKUBAUSKAITE, S., SCHMID, B., PEER, S., AXHAUSEN, K. W. & GERIKE, R. 2020. A joint time-assignment and expenditure-allocation model: value of leisure and value of time assigned to travel for specific population segments. *Transportation*, 47, 1439-1475.
- ILHAM, M. A., FONZONE, A., FOUNTAS, G. & MORA, L. 2024. To move or not to move: A review of residential relocation trends after COVID-19. *Cities*, 151, 105078.
- ITF 2019. What is the Value of Saving Travel Time? Summary and Conclusions. *ITF Roundtable Reports*.
- ITS 2010. Updating Appraisal Values for Travel Time Savings - Phase 1 study.
- ITS 2013. Values of travel time savings for business travellers
- ITS 2020. Congested Values of Travel Time (CVTT) - CVTT Forward Look.
- ITS 2022. Congested Values of Travel Time (CVTT) - Re-analysis of the 2014/15 SP Study.
- JARA-DÍAZ, S. 2024. The value(s) of travel time savings considering in-vehicle activities. *Transportation Research Part A: Policy and Practice*, 184, 104092.
- JARA-DÍAZ, S. R. 2007. Allocation and Valuation of Travel-Time Savings. In: HENSHER, D. A. & BUTTON, K. J. (eds.) *Handbook of Transport Modelling*. Emerald Group Publishing Limited.
- JARA-DÍAZ, S. R., MUNIZAGA, M. A., GREEVEN, P., GUERRA, R. & AXHAUSEN, K. 2008. Estimating the value of leisure from a time allocation model. *Transportation Research Part B: Methodological*, 42, 946-957.
- JARA-DÍAZ, S. & GUEVARA, C. A. 2003. Behind the Subjective Value of Travel Time Savings: The Perception of Work, Leisure, and Travel from a Joint Mode Choice Activity Model. *Journal of Transport Economics and Policy*, 37, 29-46.
- JOHNSON, E. 2023. The cycle of transport poverty: how the cost-of-living crisis is locking disabled people indoors. Available from: <https://www.transportforall.org.uk/news/the-cycle-of-transport-poverty-how-the-cost-of-living-crisis-is-locking-disabled-people-indoors/>.
- JOKUBAUSKAITE, S., HÖSSINGER, R., ASCHAUER, F., GERIKE, R., JARA-DÍAZ, S., PEER, S., SCHMID, B., AXHAUSEN, K. W. & LEISCH, F. 2019. Advanced continuous-discrete model for joint time-use expenditure and mode choice estimation. *Transportation Research Part B: Methodological*, 129, 397-421.
- KAHNEMAN, D. & TVERSKY, A. 2012. Prospect Theory: An Analysis of Decision Under Risk. *Handbook of the Fundamentals of Financial Decision Making*. WORLD SCIENTIFIC.
- KATO, H. 2013. On the Value of Business Travel Time Savings: Derivation of Hensher's Formula. *Transportation Research Record*, 2343, 34-42.
- KATO, H., SANKO, N., ISHIBE, M. & SAKASHITA, A. 2024. Value of travel time savings for leisure trip in autonomous vehicles: Case study from the Tokyo Metropolitan Area. *Transportation Research Interdisciplinary Perspectives*, 24, 101080.
- KESERU, I. & MACHARIS, C. 2018. Travel-based multitasking: review of the empirical evidence. *Transport Reviews*, 38, 162-183.

- KOLAROVA, V. & STECK, F. 2020. Chapter 28 - Estimating impact of autonomous driving on value of travel time savings for long-distance trips using revealed and stated preference methods. In: GOULIAS, K. G. & DAVIS, A. W. (eds.) *Mapping the Travel Behavior Genome*. Elsevier.
- KOLAROVA, V., STECK, F. & BAHAMONDE-BIRKE, F. J. 2019. Assessing the effect of autonomous driving on value of travel time savings: A comparison between current and future preferences. *Transportation Research Part A: Policy and Practice*, 129, 155-169.
- KOUWENHOVEN, M. & DE JONG, G. 2018. Value of travel time as a function of comfort. *Journal of Choice Modelling*, 28, 97-107.
- KOUWENHOVEN, M., DE JONG, G. C., KOSTER, P., VAN DEN BERG, V. A. C., VERHOEF, E. T., BATES, J. & WARFFEMIUS, P. M. J. 2014. New values of time and reliability in passenger transport in The Netherlands. *Research in Transportation Economics*, 47, 37-49.
- KRAUSS, K., KRAIL, M. & AXHAUSEN, K. W. 2022. What drives the utility of shared transport services for urban travellers? A stated preference survey in German cities. *Travel Behaviour and Society*, 26, 206-220.
- KREINDLER, G. 2024. Peak-Hour Road Congestion Pricing: Experimental Evidence and Equilibrium Implications. *Econometrica*, 92, 1233-1268.
- LAVIERI, P. S. & BHAT, C. R. 2019. Modeling individuals' willingness to share trips with strangers in an autonomous vehicle future. *Transportation Research Part A: Policy and Practice*, 124, 242-261.
- LEE, S., KIM, G. C., WU, S. K. & OH, J. 2020. Influence of ICT on Public Transport Use and Behaviour in Seoul. *International Transport Forum Discussion Papers*.
- LU, H., HESS, S., DALY, A., ROHR, C., PATRUNI, B. & VUK, G. 2021. Using state-of-the-art models in applied work: Travellers willingness to pay for a toll tunnel in Copenhagen. *Transportation Research Part A: Policy and Practice*, 154, 37-52.
- LUGANO, G. 2018. Valuing Mobility – the MoTIV Project.
- LYNCH, J., DUMONT, J., GREENE, E. & EHRlich, J. 2019. Use of a Smartphone GPS Application for Recurrent Travel Behavior Data Collection. *Transportation Research Record*, 2673, 89-98.
- MACKIE, P. J., JARA-DÍAZ, S. & FOWKES, A. S. 2001. The value of travel time savings in evaluation. *Transportation Research Part E: Logistics and Transportation Review*, 37, 91-106.
- MACKIE, P. J., WARDMAN, M., FOWKES, A. S., WHELAN, G., NELLTHORP, J. & BATES, J. 2003. Values of Travel Time Savings UK. Leeds, UK: Leeds.
- MALOKIN, A., CIRCELLA, G. & MOKHTARIAN, P. L. 2017. Do Multitasking Millennials Value Travel Time Differently? A Revealed Preference Study of Northern California Commuters. *96th Annual Meeting of the Transportation Research Board*. Washington D.C.
- MALOKIN, A., CIRCELLA, G. & MOKHTARIAN, P. L. 2019. How do activities conducted while commuting influence mode choice? Using revealed preference models to inform public transportation advantage and autonomous vehicle scenarios. *Transportation Research Part A: Policy and Practice*, 124, 82-114.
- MALOKIN, A., CIRCELLA, G. & MOKHTARIAN, P. L. 2021. Do millennials value travel time differently because of productive multitasking? A revealed-preference study of Northern California commuters. *Transportation*, 48, 2787-2823.
- MERKERT, R. & BECK, M. 2017. Value of travel time savings and willingness to pay for regional aviation. *Transportation Research Part A: Policy and Practice*, 96, 29-42.
- METZ, D. 2008. The Myth of Travel Time Saving. *Transport Reviews*, 28, 321-336.
- MEUNIER, D. 2020. Mobility Practices, Value of Time and Transport Appraisal. *International Transport Forum Discussion Papers*. Paris.
- MOKHTARIAN, P. L. 2018. The Times They Are A-Changin': What Do the Expanding Uses of Travel Time Portend for Policy, Planning, and Life? *Transportation Research Record*, 2672, 1-11.
- MOLIN, E., ADJENUGHWURE, K., DE BRUYN, M., CATS, O. & WARFFEMIUS, P. 2020. Does conducting activities while traveling reduce the value of time? Evidence from a within-subjects choice experiment. *Transportation Research Part A: Policy and Practice*, 132, 18-29.

- MOLLOY, J., CASTRO, A., GÖTSCHI, T., SCHOEMAN, B., TCHERVENKOV, C., TOMIC, U., HINTERMANN, B. & AXHAUSEN, K. W. 2023. The MOBIS dataset: a large GPS dataset of mobility behaviour in Switzerland. *Transportation*, 50, 1983-2007.
- MOYLAN, E. K. M., BLIEMER, M. C. J. & HOSSEIN RASHIDI, T. 2022. Travellers' perceptions of travel time reliability in the presence of rare events. *Transportation*, 49, 1157-1181.
- MVA, ITS & TSU OXFORD 1987. The value of travel time savings. *Policy Journals*.
- NATIONAL ACADEMIES OF SCIENCES, E. & MEDICINE 2015a. *Passenger Value of Time, Benefit-Cost Analysis and Airport Capital Investment Decisions, Volume 1: Guidebook for Valuing User Time Savings in Airport Capital Investment Decision Analysis*, Washington, DC, The National Academies Press.
- NATIONAL ACADEMIES OF SCIENCES, E. & MEDICINE 2015b. *Passenger Value of Time, Benefit-Cost Analysis and Airport Capital Investment Decisions, Volume 2: Final Report*, Washington, DC, The National Academies Press.
- NELLTHORP, J. 2023. Estimating the benefits of active travel investment. Report for Active Travel England. Institute for Transport Studies.
- NORDSTRÖM, M. & ENGHOLM, A. 2021. The complexity of value of travel time for self-driving vehicles – a morphological analysis. *Transportation Planning and Technology*, 44, 400-417.
- OJEDA-CABRAL, M., BATLEY, R. & HESS, S. 2016. The value of travel time: random utility versus random valuation. *Transportmetrica A: Transport Science*, 12, 230-248.
- OJEDA-CABRAL, M. & CHORUS, C. G. 2016. Value of travel time changes: Theory and simulation to understand the connection between Random Valuation and Random Utility methods. *Transport Policy*, 48, 139-145.
- OJEDA-CABRAL, M., HESS, S. & BATLEY, R. 2018. Understanding valuation of travel time changes: are preferences different under different stated choice design settings? *Transportation*, 45, 1-21.
- OJEDA-CABRAL, M., SHIRES, J., WARDMAN, M., TEKLU, F. & HARRIS, N. 2021. The use of recovery time in timetables: rail passengers' preferences and valuation relative to travel time and delays. *Transportation*, 48, 337-368.
- ONS 2024. Public opinions and social trends, Great Britain: working arrangements.
- ORTELLI, N., HILLEL, T., PEREIRA, F. C., DE LAPPARENT, M. & BIERLAIRE, M. 2021. Assisted specification of discrete choice models. *Journal of Choice Modelling*, 39, 100285.
- PAGONI, I., BABURAJAN, V., KOUNIADI, D., POLYDOROPOULOU, A. & DE ABREU E SILVA, J. 2022. *A Meta-Analysis of Value of Time for Autonomous Driving*.
- PALMA, D., CALASTRI, C. & PAWLAK, J. 2023. The role of time budgets in models of multi-tasking while travelling: A comparison between the MDCEV and eMDC approach. *Transportation Research Part A: Policy and Practice*, 176, 103796.
- PAWLAK, J. 2020. Travel-based multitasking: review of the role of digital activities and connectivity. *Transport Reviews*, 40, 429-456.
- PAWLAK, J., IMANI, A. F. & SIVAKUMAR, A. 2020. A microeconomic framework for integrated agent-based modelling of activity-travel patterns and energy consumption. *Procedia Computer Science*, 170, 785-790.
- PAWLAK, J., POLAK, J. W. & SIVAKUMAR, A. 2014. Microsimulation-Based Estimation of Value of Employer's Business Traveller's Value of Time: Comparison with the Current Estimation Practices and Implications for the Investment Appraisal. *93rd Annual Meeting of Transportation Research Board*. Washington D.C.
- PAWLAK, J., POLAK, J. W. & SIVAKUMAR, A. 2015. Towards a microeconomic framework for modelling the joint choice of activity-travel behaviour and ICT use. *Transportation Research Part A: Policy and Practice*, 76, 92-112.
- PAWLAK, J., POLAK, J. W. & SIVAKUMAR, A. 2017. A framework for joint modelling of activity choice, duration, and productivity while travelling. *Transportation Research Part B: Methodological*, 106, 153-172.

- PEER, E., ROTHSCHILD, D., GORDON, A., EVERNDEN, Z. & DAMER, E. 2022. Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, 54, 1643-1662.
- PEER, S. & BÖRJESSON, M. 2018. Temporal framing of stated preference experiments: does it affect valuations? *Transportation Research Part A: Policy and Practice*, 117, 319-333.
- PRELIPCEAN, A. C., SUSILO, Y. O. & GIDÓFALVI, G. 2018. Collecting travel diaries: Current state of the art, best practices, and future research directions. *Transportation Research Procedia*, 32, 155-166.
- PUDĀNE, B. 2020. Departure Time Choice and Bottleneck Congestion with Automated Vehicles: Role of On-board Activities. *European Journal of Transport and Infrastructure Research*, 20, 306-334.
- PUDĀNE, B., MOLIN, E. J. E., ARENTZE, T. A., MAKNOON, Y. & CHORUS, C. G. 2018. A Time-use Model for the Automated Vehicle-era. *Transportation Research Part C: Emerging Technologies*, 93, 102-114.
- RAMJERDI, F., RAND, L., SAETERMO, I. A. F. & SAELENSMINDE, K. 1997. The Norwegian Value of Time Study Part I and Part II.
- RICH, J., JENSEN, A. F., PILEGAARD, N. & HALLBERG, M. 2021. Cost-benefit of bicycle infrastructure with e-bikes and cycle superhighways. *Case Studies on Transport Policy*, 9, 608-615.
- RICH, J. & MABIT, S. L. 2016. Cost damping and functional form in transport models. *Transportation*, 43, 889-912.
- RICH, J. & VANDET, C. A. 2019. Is the value of travel time savings increasing? Analysis throughout a financial crisis. *Transportation Research Part A: Policy and Practice*, 124, 145-168.
- ROSSETTI, T. & DAZIANO, R. A. 2024. Crowding multipliers on shared transportation in New York City: The effects of COVID-19 and implications for a sustainable future. *Transport Policy*, 145, 224-236.
- SAMSTAD, H., RAMJERDI, F., VEISTEN, K., NAVRUD, S., MAGNUSSEN, K., FLÜGEL, S., KILLI, M., HALSE, A. H., ELVIK, R. & SAN MARTIN, O. 2010. Den norske verdsettingsstudien - Sammendragsrapport.
- SANDERS, S., DRIVYLA, E., BOURDON, J., BATLEY, R., CABRAL, M., CHINTAKAYALA, P. & SHELDON, R. 2015. Provision of market research for value of travel time savings and reliability Phase 2 Report. 1–267. ed.
- SANDORF, E. D., GRIMSRUD, K. & LINDHJEM, H. 2022. Ponderous, Proficient or Professional? Survey Experience and Smartphone Effects in Stated Preference Research. *Environmental and Resource Economics*, 81, 807-832.
- SANDORF, E. D., PERSSON, L. & BROBERG, T. 2020. Using an integrated choice and latent variable model to understand the impact of “professional” respondents in a stated preference survey. *Resource and Energy Economics*, 61, 101178.
- SANTOS DE SA, A. L., LAVIERI, P. & PAWLAK, J. 2024. Examining the Determinants of Travel Time Use and the Transferability of Activities into Trips.
- SCARPA, R., FRANCESCHINIS, C. & THIENE, M. 2021. Logit Mixed Logit Under Asymmetry and Multimodality of WTP: A Monte Carlo Evaluation. *American Journal of Agricultural Economics*, 103, 643-662.
- SCHEEPERS, E. & HOOGENDOORN-LANSER, S. 2018. State-of-the-art of incentive strategies – Implications for longitudinal travel surveys. *Transportation Research Procedia*, 32, 200-210.
- SCHMID, B., JOKUBAUSKAITE, S., ASCHAUER, F., PEER, S., HÖSSINGER, R., GRIKE, R., JARA-DIAZ, S. R. & AXHAUSEN, K. W. 2019. A pooled RP/SP mode, route and destination choice model to investigate mode and user-type effects in the value of travel time savings. *Transportation Research Part A: Policy and Practice*, 124, 262-294.
- SCHMID, B., MOLLOY, J., PEER, S., JOKUBAUSKAITE, S., ASCHAUER, F., HÖSSINGER, R., GRIKE, R., JARA-DIAZ, S. R. & AXHAUSEN, K. W. 2021. The value of travel time savings and the value of leisure in Zurich: Estimation, decomposition and policy implications. *Transportation Research Part A: Policy and Practice*, 150, 186-215.

- SDG 2016. Mobile connectivity research study.
- SHAH, K., BLOOM, N., BUNN, P., MIZEN, P., THWAITES, G. & YOTZOV, I. 2024. Managers say working from home is here to stay. Available from: <https://cepr.org/voxeu/columns/managers-say-working-home-here-stay>.
- SHELDON, T. L. & DUA, R. 2024. Consumer preferences for ride-hailing: Barriers to an autonomous, shared, and electric future. *Journal of Cleaner Production*, 434, 140251.
- SHIRES, J. D. & DE JONG, G. C. 2009. An international meta-analysis of values of travel time savings. *Evaluation and Program Planning*, 32, 315-325.
- SIGNIFICANCE 2023. Values of Time, Reliability and Comfort in the Netherlands 2022 - New values for passenger travel and freight transport.
- SIGNIFICANCE, VU UNIVERSITY AMSTERDAM & BATES, J. 2013. Values of time and reliability in passenger and freight transport in The Netherlands.
- SMALL, K. A. & ROSEN, H. S. 1981. Applied Welfare Economics with Discrete Choice Models. *Econometrica*, 49, 105-130.
- SOZA-PARRA, J., RAVEAU, S. & MUÑOZ, J. C. 2022. Public transport reliability across preferences, modes, and space. *Transportation*, 49, 621-640.
- STANDEN, C., GREAVES, S., COLLINS, A. T., CRANE, M. & RISSEL, C. 2019. The value of slow travel: Economic appraisal of cycling projects using the logsum measure of consumer surplus. *Transportation Research Part A: Policy and Practice*, 123, 255-268.
- SUNSTEIN, C. R. 2020. Behavioral Welfare Economics. *Journal of Benefit-Cost Analysis*, 11, 196-220.
- SWÄRDH, J.-E. & ALGERS, S. 2016. Willingness to accept commuting time within the household: stated preference evidence. *Transportation*, 43, 219-241.
- SWEET, M. N. 2021. User interest in on-demand, shared, and driverless mobility: Evidence from stated preference choice experiments in Southern Ontario. *Travel Behaviour and Society*, 23, 120-133.
- SYSTRA, IMPERIAL COLLEGE LONDON & DANISH TECHNICAL UNIVERSITY 2015. Value of travel time savings: peer review and audit.
- TABASI, M., RAEI, A., HILLEL, T., KRUEGER, R. & HOSSEIN RASHIDI, T. 2023. Empowering revealed preference survey with a supplementary stated preference survey: demonstration of willingness-to-pay estimation within a mode choice case. *Travel Behaviour and Society*, 33, 100632.
- TJIONG, J., DEKKER, T., HESS, S. & OJEDA CABRAL, M. 2022. The selection of income measures in value of travel time models and their implications for the VTT, its cross-sectional income elasticity and transport appraisal. *Research in Transportation Economics*, 94, 101168.
- TORDAI, D., MUNKÁCSY, A., ANDREJSZKI, T. & HAUGER, G. 2023. The real value of cycling. *Transportation Research Procedia*, 72, 2896-2903.
- TRAFIKVERKET 2024. Analysmetod och samhällsekonomiska kalkylvärden för transportsektorn.
- TRAIN, K. 2009. *Discrete choice methods with simulation*, Cambridge, Cambridge University Press.
- TRAIN, K. 2016. Mixed logit with a flexible mixing distribution. *Journal of Choice Modelling*, 19, 40-53.
- TRAIN, K. & MCFADDEN, D. 1978. The goods/leisure tradeoff and disaggregate work trip mode choice models. *Transportation Research*, 12, 349-353.
- TRAIN, K. & WEEKS, M. 2005. Discrete Choice Models in Preference Space and Willingness-to-Pay Space. In: SCARPA, R. & ALBERINI, A. (eds.) *Applications of Simulation Methods in Environmental and Resource Economics*. Dordrecht: Springer Netherlands.
- TSOLERIDIS, P., CHOUDHURY, C. F. & HESS, S. 2022. Deriving transport appraisal values from emerging revealed preference data. *Transportation Research Part A: Policy and Practice*, 165, 225-245.
- US DOT 2016. Revised Departmental Guidance on Valuation of Travel Time in Economic Analysis.
- VAN CRANENBURGH, S. & KOUWENHOVEN, M. 2021. An artificial neural network based method to uncover the value-of-travel-time distribution. *Transportation*, 48, 2545-2583.
- VAN CRANENBURGH, S., KOUWENHOVEN, M., DEKKER, T., MULLER, J. & DE JONG, G. 2023. Random Valuation models and additive Random Utility Maximisation principles.

- VAN CRANENBURGH, S., WANG, S., VIJ, A., PEREIRA, F. & WALKER, J. 2022. Choice modelling in the age of machine learning - Discussion paper. *Journal of Choice Modelling*, 42, 100340.
- VAN WEE, B. & BÖRJESSON, M. 2015. How to make CBA more suitable for evaluating cycling policies. *Transport Policy*, 44, 117-124.
- VARGHESE, V. & JANA, A. 2018. Impact of ICT on multitasking during travel and the value of travel time savings: Empirical evidences from Mumbai, India. *Travel Behaviour and Society*, 12, 11-22.
- WANG, S., WANG, Q. & ZHAO, J. 2020. Deep neural networks for choice analysis: Extracting complete economic information for interpretation. *Transportation Research Part C: Emerging Technologies*, 118, 102701.
- WARDMAN, M. 2004. Public transport values of time. *Transport Policy*, 11, 363-377.
- WARDMAN, M., BATLEY, R., LAIRD, J., MACKIE, P. & BATES, J. 2015. How should business travel time savings be valued? *Economics of Transportation*, 4, 200-214.
- WARDMAN, M., CHINTAKAYALA, P. & HEYWOOD, C. 2020. The valuation and demand impacts of the worthwhile use of travel time with specific reference to the digital revolution and endogeneity. *Transportation*, 47, 1515-1540.
- WARDMAN, M., CHINTAKAYALA, V. P. K. & DE JONG, G. 2016. Values of travel time in Europe: Review and meta-analysis. *Transportation Research Part A: Policy and Practice*, 94, 93-111.
- WARDMAN, M. & LYONS, G. 2016. The digital revolution and worthwhile use of travel time: implications for appraisal and forecasting. *Transportation*, 43, 507-530.
- WARDMAN, M. & NICOLÁS IBÁÑEZ, J. 2012. The congestion multiplier: Variations in motorists' valuations of travel time with traffic conditions. *Transportation Research Part A: Policy and Practice*, 46, 213-225.
- WEIS, C., VRTIC, M., WIDMER, P. & AXHAUSEN, K. W. 2011. Influence of parking on location and mode choice. A stated choice survey. IVT, ETH Zürich.
- WELSH GOVERNMENT 2022. Ministerial position statement to assist Welsh Transport Appraisal Guidance (WelTAG) 2022 users.
- WHEAT, P. & BATLEY, R. 2015. Quantifying and decomposing the uncertainty in appraisal value of travel time savings. *Transport Policy*, 44, 134-142.
- WHO. 2024. *Health Economic Assessment Tool (HEAT) for walking and cycling* [Online]. Available: [https://www.heatwalkingcycling.org/?&heat\\_locale=en-US&heat\\_lang=en#homepage](https://www.heatwalkingcycling.org/?&heat_locale=en-US&heat_lang=en#homepage) [Accessed].
- WINKLER, C., MEISTER, A. & AXHAUSEN, K. W. 2024. The TimeUse+ data set: 4 weeks of time use and expenditure data based on GPS tracks. *Transportation*.
- WINSTON, G. C. 1987. Activity choice: A new approach to economic behavior. *Journal of Economic Behavior & Organization*, 8, 567-585.
- WSP & MOTT MACDONALD 2019. Congestion Dependent Values Of Time In Transport Modelling.
- WSP & MOTT MACDONALD 2022. Congestion Dependent Values Of Time In Transport Modelling Further Prism Testing.
- WSP, RAND EUROPE & MOTT MACDONALD 2018. Congestion dependent values of time in transport modelling.
- YIN, H. & CHERCHI, E. 2024. Preferences for automated taxis. A comparison between immersive virtual reality and screen-based stated choice experiments. *Transportation Research Part C: Emerging Technologies*, 163, 104628.
- ZANG, Z., BATLEY, R., XU, X. & WANG, D. Z. W. 2024. On the value of distribution tail in the valuation of travel time variability. *Transportation Research Part E: Logistics and Transportation Review*, 190, 103695.
- ZAPATA-DIOMEDI, B. & KROEN, A. 2023. Cost benefit analysis of active modes of transport: Review of international guidance. *Australasian Transport Research Forum 2023*. Perth Australia.