Institute for Transport Studies
FACULTY OF ENVIRONMENT

UNIVERSITY OF LEEDS

Programme for maintaining a robust valuation of travel time savings: feasibility study

Final Phase 1 report: List of options

ARUP

Institute for Transport Studies, University of Leeds

November 2017

ITS

The following people have contributed to this report, listed by firm:

Arup
Margot Finley
Matthew Dillon

Institute for Transport Studies, University of Leeds
Thijs Dekker
Richard Batley
Andrew Daly
James Laird
Peter Mackie
Dan Johnson
Gerard de Jong
Stephane Hess
Greg Marsden
Mark Wardman
Phill Wheat

External reviewers (working for ITS)
Prof. Kay Axhausen (ETH Zürich, Switserland)
Prof. Jonas Eliasson (Royal Institute of Technology, Sweden
Head of Transport - City of Stockholm, Sweden)

Institute for Transport Studies – Document control information

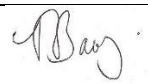
Title	Programme for maintaining a robust valuation of travel time savings: feasibility study. Final Phase 1 report: List of options
Authors	Institute for Transport Studies (ITS), University of Leeds and ARUP
Editor	Thijs Dekker
Internal reviewers	Richard Batley and Peter Mackie
ARUP QA	
External reviewers	Kay Axhausen, Jonas Eliasson
Version number	1.0
Date	10th of November 2017
Distribution	ITS & ARUP team, DfT
Availability	ITS & ARUP team, Iven Stead, Aruna Sivakumar
File	ARUP_ITS_Robust_VTTS_Phase1_Final.docx
Signature	

Table of Contents

1. INTRODUCTION 4
2. THE 2014/15 ARUP AND ITS VTTS STUDY 6
2.1 The 2014/15 framework relative to the 2003 framework 6
3. CHANGES IN THE VTTS OVER TIME 10
4. STREAM 1: REGULAR UPDATING OF THE VTTS USING SECONDARY DATA 13
Option 1.1: Uprating for GDP per capita growth 13
Option 1.2: Adjusting for changing trip and socio-economic characteristics 17
Option 1.3: Applying forecasts on the NTS data for future VTTS values 21
Summary – Stream 1 22
5. STREAM 2: MONITORING CHANGES IN THE VTTS USING EMERGING EVIDENCE AND NEW DATA 23
Option 2.1: Horizon scanning 23
Option 2.2: A monitoring scheme 41
Summary - Stream 2 44
6. STREAM 3: ROBUSTNESS OF THE BASE AND UPDATED VTTS VALUES 45
Option 3.1: Better understanding the 2014/15 confidence intervals 45
Option 3.2: Steps to increase robustness without risk of bias to the estimates 48
Option 3.3: Uncertainty in future VTTS values 49
Summary - Stream 3 51
7. STREAM 4: COMMISSIONING OF STUDIES ON OUTSTANDING AND EMERGING TOPICS 52
8. SUMMARY 53
9. RECOMMENDATIONS FOR NARROWING DOWN THE LONG-LIST 56
10. LIST OF REFERENCES 58
LIST OF ABBREVIATIONS 62

1. Introduction

ARUP and the Institute for Transport Studies (ITS) at the University of Leeds have been engaged by the Department for Transport to undertake a feasibility study into the development of a programme for maintaining a robust valuation of travel time savings (DfT Reference 04/102/). The purpose of this document is to summarise Phase 1 of the study.

The most recent valuation of travel time savings (VTTS) study undertaken by ARUP and ITS in 2014/15[1] has provided the Department with an updated set of national representative values based on state-of-the-art modelling and implementation techniques. The rationale for developing a maintenance programme is that in order to ensure this evidence-base remains fit-for-purpose the development of a maintenance and development programme is required.

The objectives of Phase 1 are to develop a long-list of alternatives for maintaining the robustness of VTTS over time; and to provide recommendations on how this long-list might be narrowed to a short-list of alternatives for more detailed analysis in Phase 2. The final set of options to progress to Phase 2 will be agreed in collaboration with the DfT Project Board.

The project team believe that a maintenance and development programme should be structured along four core streams. The four streams have already been presented to and have been agreed by the Department during Phase 1. The respective streams are:

1. Regular updating of the VTTS using secondary data
2. Monitoring changes in the VTTS using emerging evidence and new data
3. Robustness of the base and updated VTTS values
4. Commissioning of studies on outstanding and emerging topics

Stream 1 assumes that the 2014/15 behavioural framework remains valid over an extended period and that the VTTS only needs to be adjusted for changes in socio-economic and trip characteristics. While preferable to doing nothing at all, Stream 1 ignores potential changes in preferences and thus changes in the VTTS values. Not accounting for the fact that real-world travel behaviour may no longer be in line with the assumed behavioural framework could imply that the VTTS values coming out of the regular VTTS updating programme are no longer fit-for-purpose.

Against this background, **Stream 2** entails a monitoring scheme to detect whether newly emerging evidence is in correspondence with the outcomes of Stream 1. Stream 2 goes into the details of such a monitoring programme, this feeds into **Stream 3** which evaluates the robustness of the 2014/15 VTTS relative to the newly emerging evidence. If and when it is judged that the 2014/15 framework is no longer fit-for-purpose, new data collection efforts and revision of the 2014/15 framework may be deemed appropriate. In order to make sure such efforts are following state-of-the-art

[1] https://www.gov.uk/government/publications/values-of-travel-time-savings-and-reliability-final-reports

data collection and analysis methods, the Stream 2 should also monitor developments in the broader literature. Whereas the first two streams are focused on developing an overall maintenance programme specifically addressing changes in the VTTS over time, **Stream 4** allows the Department to revisit outstanding and emerging topics and thereby refine their approach to specific mode-purpose segments.

The report is structured as follows. Section 2 provides the necessary details of the 2014/15 VTTS study conducted by ARUP and ITS. A key component of Section two is to understand how the 2014/15 study relates to its predecessor, i.e. the Mackie et al. (2003) study. After having introduced the reference study, Section 3 discusses the rationale why the VTTS may change over time. Sections 4 to 7 then work out the separate streams of the programme. Finally, Section 8 provides a summary of the programme bringing the individual elements together and briefly touches upon which topics were considered out of scope for the current study. Section 9 concludes by providing recommendations on narrowing down the long-list for Phase 2.

2. The 2014/15 ARUP and ITS VTTS study

The 2014/15 VTTS study provided a revised set of value of travel time savings estimates for different mode-purpose and distance combinations as well as values of reliability and other journey quality attributes. Deriving the revised WebTAG values required an elaborate study in which new primary Stated Preference (SP) data were collected and jointly analysed using state-of-the-art choice modelling techniques. An Implementation Tool was developed combining the behavioural modelling results with the National Travel Survey (NTS) to derive a national representative set of VTTS measures. Figure 1 below summarises how the 2014/15 VTTS study arrived at the eventual appraisal values.

The task in hand for the current project is to determine the best means of updating the 2014/15 evidence base going forwards ensuring that the framework remains fit-for-purpose. In developing this maintenance and development programme, considerations need to be made with respect to the two underlying datasets, the behavioural model and the interplay between them.

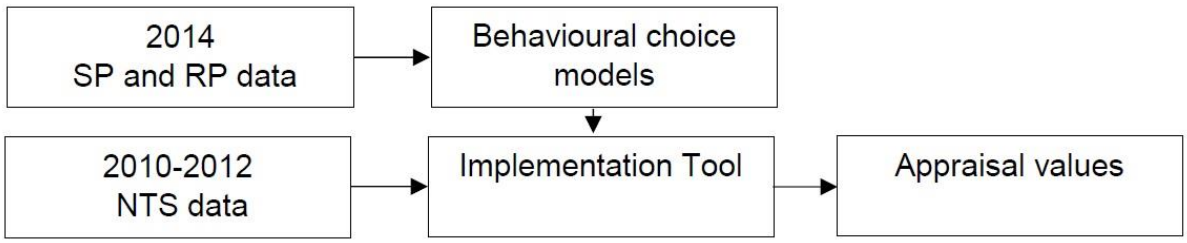


Figure 1 – Graphical summary of the 2014/15 VTTS framework

Before discussing inter-temporal changes in the VTTS, additional details of the 2014/15 framework are provided and contrasted with the preceding Mackie et al. (2003) VTTS framework. We refer to these as being respectively the 2014 and 1994 nationally representative *base values* for specific mode-purpose (and distance) combinations.

2.1 The 2014/15 framework relative to the 2003 framework

Mackie et al. (2003) re-analysed the 1994 Accent Marketing Research and Hague Consulting Group (AHCG,1994) SP survey data for non-work purposes (commute and other non-work).[2] This re-analysis resulted in the following behavioural model:

$$V = \left(\frac{\beta_T}{\beta_C}\right) \cdot \left(\frac{Y}{\bar{Y}}\right)^{\eta_Y} \cdot \left(\frac{C}{\bar{C}}\right)^{\eta_C} \quad (1)$$

[2] Mackie et al. (2003 section 2.1) and Batley et al. (2017) provide a history of VTTS studies in the UK.

In equation (1), *V* represents the VTTS (in pence per minute), $\beta_T, \beta_C, \eta_Y, \eta_C$ are parameters estimated through discrete choice models where the latter two reflect income and cost elasticities respectively and β_T and β_C are generic parameters describing sensitivities to time and cost respectively. The base VTTS values for non-work purposes were therefore treated as a function of income *Y* and travel costs *C*. These factors were both measured relative to a reference level, being $\bar{Y} = 35$ (in £'000 pa) and $\bar{C} = 100$ (in pence) respectively, so that the primary focus of these variables is on how the VTTS values vary in the cross-sectional dimension across different income groups. All monetary values in equation (1) are described in 1994 prices.

In making equation (1) applicable for appraisal purposes, some specific assumptions were implemented:

$$V_{yd} = K \cdot \left(\frac{\beta_T}{\beta_C}\right) \cdot \left(\frac{Y_y}{\bar{Y}}\right)^{\eta_Y} \cdot \left(\frac{D_d}{\bar{D}}\right)^{\eta_C} \tag{2}$$

- V_{yd} was assumed to vary across a range of income *y* and distance *d* bands
- Travel costs were converted into distance using a fixed conversion rate of 13.2 pence per mile. As a result, $\bar{D} = \frac{\bar{C}}{13.2} = \frac{100}{13.2} = 7.58$
- *K* accounts for price inflation, which would only be applied when changing the price year of the base VTTS.

Income and distance were split up into respectively thirteen and ten bands such that a 13x10 matrix arose.[3] Six years of NTS 1995/2000 were used to determine the number of trips that could be attributed to each of the cells (see Mackie et al. (2003)). Specifically, median income levels and distances within each band were calculated to arrive at Y_y and D_d respectively, which can be applied in equation (2) to obtain V_{yd}. These values allowed the derivation of a set of nationally representative appraisal values $\bar{V}$ using both trip and distance weights in equation (3).

$$\bar{V} = \frac{\sum_y \sum_d V_{yd} \cdot N_{yd} \cdot \bar{D}_d}{\sum_y \sum_d N_{yd} \cdot \bar{D}_d} \tag{3}$$

where:

- N_{yd} is the number of NTS trips in the respective distance and income band
- $\bar{D}_d$ weights each trip by distance

At this point, we are able to highlight similarities and differences between the 2014/15 VTTS study and the Mackie et al. (2003) framework. Both frameworks estimate a set of behavioural models using SP data and subsequently combine this with NTS data to arrive at a nationally representative value.

[3] Originally, 21 income and 12 distance bands were explored by Mackie et al. (2003 – p.57), of which certain bands were merged in developing the official appraisal framework.

The newly collected SP data cover a wider range of mode-purpose combinations (see Batley et al. 2017). For example, business values of time were previously derived using the Cost Savings Approach (CSA), but are now also elicited using SP data. Most importantly, the 2014/15 behavioural models account for more than just income and distance effects. The behavioural elements added to the 2014/15 models can best be classified as:

- *Unobserved heterogeneity* – also known as random parameter models
- *Size and sign effects* – as described in De Borger and Fosgerau (2008)
- *Covariates* – accounting for trip and socio-economic characteristics
    - *Elasticities* - distance and time elasticities are now additional to income and cost elasticities, with cost no longer directly proportional to distance[4]
    - *SP effects* – these SP artefacts did not affect the appraisal values
    - *NTS covariates*, such as age, gender, household composition and car ownership
    - *Other covariates*, like the NTS covariates but SP sample averages were taken since these are not present in the NTS data.

In addition to these elements, changes to the functional form of the choice models were made by respectively assuming a multiplicative instead of an additive error structure, by working in a random valuation framework, and by jointly estimating multiple SP games in a single framework (see Hess et al. 2017 for estimation details).

In terms of implementation of the new behavioural models for appraisal purposes, a key difference between the Mackie et al. (2003) and 2014/15 frameworks is that the latter included for each trip included within the 2010-2012 NTS dataset, instead of each income and distance band, an expected VTTS using the required covariates.[5] Note that by working with the expected value, the influence of unobserved heterogeneity was taken into account. Sign effects are also no longer present by making use of a geometric mean approach (see Hess et al. 2017). This means that the values coming out of the 2015 study are in effect middle values between the willingness to pay for travel time savings, and the willingness to accept travel time increases in return for reduced cost. Referring to them as VTTS is common practice in some areas, but is in practice equal to an assumption that the sign effects are a pure survey artefact and that in reality people behave symmetrically to savings and losses.

While sign effects cancel out in the calculations, size effects on the other hand remain present and a value of $\Delta T = 10$ (minutes) has been selected after due consideration and in line with international evidence. Finally, although a weighted average was still applied to derive a nationally representative VTTS measure, the weights now make use of NTS expansion factors (and distance) instead of the number of trips in a respective income-distance band (and distance).

[4] TAG Unit M2 Appendix C.3 still applies distance and income effects to obtain local values of time.

[5] This does not take away that for each income and distance band a me(di)an needs to be defined.

The resulting 2014/15 VTTS formula is captured by equation (4) which is identical to equation 34 in Hess et al. (2017):

$$E(VTT_i(\Delta T)) = \exp\left(\kappa_{SP_x} \cdot a\right) \frac{\exp\left(\kappa_{SP_x} \cdot b\right) - 1}{\kappa_{SP_x} \cdot b} \nu_{SP_x,VTT} \prod_m z_{mi}^{\lambda} \prod_n \zeta_n^{z_{ni}} |\Delta T|^{\kappa_{SP_x} - 1} \quad (4)$$

where:

- $\exp\left(\kappa_{SP_x} \cdot a\right) \frac{\exp\left(\kappa_{SP_x} \cdot b\right) - 1}{\kappa_{SP_x} \cdot b}$ represents the mean base VTTS, similar to $\left(\frac{\beta_T}{\beta_C}\right)$ but taking account of the random heterogeneity
- $\nu_{SP_x,VTT}$ represents SP game multipliers which are normalised to one for the reference SP1 game. These terms are essential when valuing reliability and crowding effects.
- $\prod_m z_{mi}^{\lambda} \prod_n \zeta_n^{z_{ni}} |\Delta T|^{\kappa_{SP_x} - 1}$ represents a set of VTTS multipliers, including income and cost elasticities similar to $\left(\frac{Y_y}{\bar{Y}}\right)^{\eta_Y} \cdot \left(\frac{D_d}{\bar{D}}\right)^{\eta_C}$

Sample enumeration on 2010-2012 NTS is subsequently used to derive a nationally representative appraisal value for specific mode-purpose segment *s:*

$$\overline{VTT}_S(\Delta T) = \frac{\sum_{i \in S} w_i \cdot l_i \cdot E(VTT_i(\Delta T))}{\sum_{i \in S} w_i \cdot l_i} \quad (5)$$

Where the summation is taken over all the trips i in the segment *s* and

- w_i represents the NTS expansion factor operating like N_{yd}.
- l_i represents a trip distance and operates like the distance weighting measure $\bar{D}_d$, but is now observed at the trip-level rather than using a median distance at the band level.

In summary, the 2014/15 framework made significant progress in determining the behavioural framework by using state-of-the-art choice modelling techniques. It also made use of the NTS data at a more disaggregate level to arrive at a set of *base VTTS values*. Despite this methodological progress, the concept of using SP data to obtain a behavioural model and combining the SP models with representative NTS data was also present in the Mackie et al. (2003) framework and we can therefore draw upon experience in assessing the feasibility of options for the envisaged updating programme. However, a key distinction of the 2014/15 framework, is that it has more detailed data requirements relative to the preceding Mackie et al. (2003) framework.

3. Changes in the VTTS over time

Economic theory informs us that the value of non-work travel time savings is a ratio of the marginal utility of transferring time spent travelling to leisure and the marginal utility of income (Mackie et al., 2001). Considering (inter-temporal) changes in the VTTS, it is important to understand the nature of these two marginal utility terms. The marginal utility of income is generally assumed to diminish with increasing income due to the notion of decreasing marginal utility of consumption. We therefore expect those with higher incomes to have a higher non-work VTTS, ceteris paribus. A similar theoretical relationship does not exist between income and the marginal utility of transferring time spent travelling to leisure; we are dependent on empirical evidence to support even the basic proposition that leisure is a normal good. Journeys which are more arduous for whatever reason (lack of comfort, tighter time budget, more crowded, etc.) will, however, be associated with a higher marginal disutility of travel time and thereby also have higher non-work VTTS, ceteris paribus (see also Mackie et al. 2003 sections 6.5-6.8).

For business trips, economic theory informs us that the productivity (either for the business or for the personal use of the traveller) of travel time saved is critical to the valuation of Business VTTS. It is again the relative productivities between time spent travelling and time spent at work or at leisure, which are relevant.[6] This philosophy underlies different models for business time savings – see Wardman et al (2015) for a review of alternative theoretical models. All theoretical models are intrinsically linked to the wage rate. The role of other factors including, but not limited to, the productive use of time whilst travelling, and whether the time saved can be used to produce useful output, are expected to have a strong influence on the actual VTTS. The 2014/15 study found strong evidence that the VTTS for business trips was significantly lower than that expected from the CSA at short distances, but similar to the CSA at long distances[7]. The data underlying the model was not able to explain the behavioural reasons for this, but the model is consistent with the broader business VTTS evidence base and was considered robust. The implication of the model is that when travel time saved from 'short distance' trips is transferred to other uses its productivity does not increase by much, whereas at 'long distances' the relative productivities of time are much greater.

How VTTS changes over time therefore is a function of how all the different parameters within the non-work and business VTTS models change over time. The overall growth rate in VTTS is therefore a confounding of many 'micro-growth' factors possibly having complex interdependencies. These include but are not limited to:

[6] The productivities mentioned here include, besides work productivity, the arduous circumstances, such as comfort, referred to under non-work travel affecting the direct utility of business travel.

[7] It should be noted that 'reference time' and 'reference cost' drive this change and not distance per se.

Non-work:

- How socio-demographic characteristics of the travelling public change over time (including how the time budget constraint binds, as well as ages of travellers, income (via the marginal utility of income), etc.).
- How technology continues to influence the productive use of time whilst travelling.
- How travelling comfort and other quality characteristics of the journey change.
- How trip lengths change (as longer trips are associated with higher VTTS).
- How preferences change.

Business:

- How wages change for the travelling workforce.
- Whether travelling time becomes more (or less) productive for work.
- Whether there are changes in the manner that saved time can be directed into additional work.
- Whether there are changes in the manner that saved time can be directed into non-work activities.
- Personal preferences of the business traveller.

These lists of 'micro-growth' factors are non-exhaustive and the relevance of each of the factors mentioned for the business VTTS depends on the perspective adopted whether the business VTTS is the outcome of a utility maximising consumer deciding on a business trip and thereby also internalising the personal welfare of travel time savings, or based on a marginal productivity approach based using either the CSA or the 'Hensher-equation'.

The challenge that emerges in the context of the present study is that changes in the VTTS need to be positioned relative to the 2014/15 framework. Some of the above factors driving changes in the VTTS over time may be a result of changes in the behavioural model, e.g. when preferences or technologies change, and would call for new data collection efforts to identify the resulting changes in the VTTS. Other factors may be of less concern as they would not affect the behavioural framework. It may be assumed that ageing, for example, implies these people will adopt the preferences of the people who were in that age category before. This would imply that the behavioural model would still be relevant and new NTS data and expansion weights could be used to update the base VTTS.

The default approach of the Department has been to uprate the VTTS for growth in Gross Domestic Product (GDP) per capita over time using equation (6). Prior to 2004 the Department implemented a unit income elasticity, i.e. $\epsilon_{GDP} = 1$, which was updated for private travel based on the meta-analysis of Wardman (2004) on the VTTS in public transport to 0.8. A subsequent meta-analysis by Abrantes and Wardman (2011) covering public transport and car VTTS values in the UK found a range of GDP elasticities close to one. As a result, the Department then reverted to an income elasticity of 1 in WebTAG. A more recent meta-analysis (Wardman et al. 2016), using a broader European evidence base, confirmed the use of an income elasticity around

1. Laird et al. (2013) in their peer review, however, noted that the elasticity found by Abrantes and Wardman (2011) is sensitive to the studies that have been included, with the more recent studies driving the increase in elasticity from 0.8 to 1.0. There appeared, for example, to be no growth in the VTTS between 1994 and 2003. This appears to be in contrast with findings by Wardman et al. (2016) who tested for different income elasticities for different decades but found these differences not to be significant, except for a small and positive pre-Euro period.

$$V_t^* = V \cdot \left(\frac{GDP_t}{GDP_{1994}}\right)^{\epsilon_{GDP}} \tag{6}$$

Some repeat studies looking at the VTTS at multiple points in time have been carried out in a few countries, which makes it possible to investigate the changes in the real VTTS over time (Gunn, 2001; Tapley et al., 2007; Börjessson et al., 2012; Kouwenhoven et al., 2017). The overall picture, especially for public transport, is not very consistent (ranging from reductions in the VTTS over time to increases that exceed the real income growth). For car transport, the picture is not so mixed and the studies show increases in the VTTS, but generally by less than the income change. One of the reasons stated in these studies why the VTTS could increase less over time than might be expected based on income growth is the increase in the use of information and communications technology that makes the journey itself more productive or pleasant (and thus reducing the VTTS). On the other hand, the use of econometric models that better capture the extreme right tail of the VTTS distribution could lead to higher VTTS (Börjesson et al. 2012a; Significance et al., 2013).

Fosgerau (2005) argues that a unit income elasticity is a natural expectation to have since the simplest micro-economic models find that the VTTS is equal to the marginal after-tax hourly wage. There are, however, many theoretical extensions (i.e. model specifications) in which, and empirical reasons why, the inter-temporal and cross-sectional income elasticities can differ from unity. One of the reasons could be that studies use before- instead of after-tax income. Also, the absence of the wage rate in the theoretical definition of the non-work VTTS plays a role. Other societal and technological developments increasing or decreasing the VTTS, which correlate with growth in the GDP per capita, could similarly wrongfully attribute changes in the VTTS to income growth and thereby produce an inter-temporal income elasticity different from unity. As such, we do not believe that a unit income elasticity should be treated as ‘correct’ or should be treated as an upper bound without improving our understanding on what the empirically estimated parameter represents.

4. Stream 1: Regular updating of the VTTS using secondary data

Assumption: The 2014/2015 VTTS behavioural models remain fit-for-purpose

As established in the previous section, there are many factors that may induce inter-temporal changes in the VTTS. Most importantly, the 2014/15 behavioural framework remains fit-for-purpose as long those factors do not affect the assumed behavioural model. The latter assumption has implicitly been adopted in the Department's default approach of adjusting the base VTTS values for growth in GDP per capita. This assumption will also be taken forward in Stream 1 of the proposed updating programme. The options described in Stream 1 are therefore aimed at adjusting the 2014/15 values over time given the 2014/15 behavioural model. By assuming that the behavioural models remain fit-for-purpose, there is no need to collect new SP and Revealed Preference (RP) data, but there is a need for secondary data sources e.g. to identify income growth.

Option 1.1: Uprating for GDP per capita growth

Option 1.1 continues the Department's default approach of adjusting the VTTS for annual growth in GDP per capita using a pre-determined income elasticity (see equation 6).[8]

WebTAG Unit 1.3 provides future VTTS values for the next 60 years. The forecasted values are based on long term annual income growth of approximately 2% per year such that the VTTS also grows at 2% per year. Figures are, however, regularly corrected for actual income growth and medium term income growth levels are gradually adjusted to reach the long term income growth levels. As a result, the recommended VTTS figures in WebTAG are updated on an annual or semi-annual basis. Clearly any changes to the default long run GDP per capita growth assumptions recommended by OBR, the Green Book etc would need to be adopted in transport sector appraisal practice including the VTTS forecasts.

Option 1.1 is common practice in most European countries, although using alternative values for the income elasticity parameter. The Guide to cost-benefit analysis of investment projects (European Commission, DG-REGIO, 2014) writes on such escalation factors:

"*The real value of work time is directly related to the real wage rate. Thus, it will grow with the projected wage rate, which is typically assumed to equal the growth in GDP per capita. The economic literature suggests escalating value of time for future years across the time horizon based on a default inter-temporal elasticity to GDP per capita growth of 0.7 to 1.0. This elasticity is expected to vary very little across market segments and to be stable over time. The value of non-work time is not related to the wage rate and as such there is no theoretical justification for linking it to wage rate*

[8] The default approach additionally relies on the assumption that the makeup of the travelling population and their trips remains constant.

growth. However, its value is related to income and any changes in income will affect that value. Studies in the UK and the Netherlands have indicated elasticity of value of time with respect to income of approximately 0.5 to 0.8. It is generally recommended that value of both work and non-work time be treated as increasing over time in proportion to GDP per capita, unless there is local evidence to the contrary. For the sake of prudency, it is however recommended to use the lower elasticity values illustrated above: 0.7 and 0.5 for, respectively, work and non-work time. If HEATCO values are adopted as a last resort, the use of lower elasticity is recommended. In line with the use of constant prices, the inflation effect must not be taken into account for escalation".

In the values of time and safety that RAND Europe and CE Delft (2004) recommended for the European Investment Bank (EIB), a GDP per capita elasticity of the VTTS of 1 was used for passenger and freight transport. HEATCO (2006) recommended an elasticity of 0.7 here (with a sensitivity test at 1). For passenger transport, JASPERS (a common unit for project assistance of the EU, EIB and EBRD) uses a GDP per capita elasticity of 0.5, referring to the above-mentioned guide to Cost-Benefit Analysis (CBA). In CBA in The Netherlands, for passenger transport, a real income elasticity of the VTTS of 0.5 is used. This is based on the national and international literature, but especially on the comparison of the 1988 and 1997 VTTS study in The Netherlands by Gunn (2001), which of course omits more recent information.

A key issue that therefore needs to be addressed and continuously revisited is what the most appropriate value for ϵ_{GDP} is. In the present context, ϵ_{GDP} is the only parameter through which all sources of inter-temporal change in the VTTS can be operationalised. It therefore needs to be established whether the parameter should be interpreted as a pure income effect or whether it represents a mixture of effects. The literature (e.g. Wardman 2001; Laird et al. 2013) acknowledges that, despite including a number of control variables in the meta-analyses on which the current value for ϵ_{GDP} is based, the income effect cannot be entirely separated from other factors and should therefore not be interpreted as a pure income effect.[9] Laird et al. (2013) furthermore highlight that the current evidence base for the inter-temporal income elasticity parameter is inconclusive as the value of parameter is sensitive to the set of studies included. Börjesson et al. (2012) provide further evidence that the income elasticity is not constant but increases with income.

We would therefore **recommend** additional research on the appropriateness of the current assumption where $\epsilon_{GDP} = 1$ from both a theoretical and empirical perspective. The meta-analyses that have been used to estimate the current elasticity parameter are based on a combination of longitudinal and cross-sectional data and only a limited set of studies have looked at the inter-temporal change in the VTTS in a single case. The nature of these elasticities will also not be purely longitudinal. The elasticity based

[9] Laird et al. (2013) p22: "*Any chosen inter-temporal elasticity is likely to be a mixture of a pure income effect and other trends over time which cannot easily be separated.*"

only on UK evidence by Abrantes and Wardman (2011), will be more of a longitudinal nature than the one estimated on European evidence (Wardman et al. 2016).

When considering the assumptions on ϵ_{GDP}, it might occur to the Department to adopt a different approach across different journey purposes along the recommendations made by the European Commission. That is to say, adopting a different income elasticity for work and non-work trips due to difference in the theoretical relationship between income growth and growth in the VTTS for these purposes. Similarly, the VTTS for commercial vehicles should grow in line with wages of commercial drivers – which may not grow in line with GDP per capita. For example, in the Netherlands and in JASPERS, freight VTTS is estimated using the real change in the costs of providing freight transport services.

At this stage, we do not **recommend** the DfT to adopt an income elasticity parameter that varies over other dimensions, such as time or income levels, since the supporting evidence is inconclusive. Naturally, this ought to be part of future research in this area.[10]

Implications of Option 1.1 for the updating programme

In summary, adjusting the VTTS for GDP per capita over time (Option 1.1) is considered a feasible option. Additional research is recommended to better understand what is captured by this income growth correction and what the best evidence base is to identify the inter-temporal income elasticity. Selecting the correct income elasticity should, however, be accompanied by accurate GDP growth forecasts. Finally, future GDP growth expectations need to be continuously monitored.

[10] One of the reviewers pointed out that the link with between the income elasticity and the discount rate needs to be considered. Indeed, the Ramsey equation, as adopted by HM Treasury, includes a (unit) income elasticity and a GDP income growth term. The discount rate, however, falls beyond the control of the DfT. A common discount rate is applied across government departments. Moreover, the income elasticity term in the Ramsey equation refers to the marginal utility of consumption (income), whereas in the VTTS the interpretation taken on the income elasticity may be broader than representing the marginal utility of income.

Actions needed to develop Option 1.1 for implementation:

- Review the interpretation of the inter-temporal income elasticity parameter
- Review the evidence behind the inter-temporal income elasticity parameter
- Explore the potential for applying different GDP growth factors across different journey purposes
- Review the consideration of the stability of the characteristics of the travelling population and its trips
    - This is implicitly included in Option 1.2

Additional actions needed when implementing Option 1.1:

- Adjust the VTTS for actual and changing growth forecasts over time
    - Connect with Option 2.1.1 exploring alternative future scenarios
- Apply inflation correction when the price year changes

Actions needed in Phase 2 of this study:

- Academic consultation on the interpretation of the inter-temporal income elasticity parameter and the corresponding evidence base.

Option 1.2: Adjusting for changing trip and socio-economic characteristics

In 2013, the Department departed from its default approach (Option 1.1). That is, besides accounting for growth in GDP per capita, new NTS data covering the years 2008-2010 were used to re-evaluate equation (3) for non-work purposes. Hence, changes in trip distance and the relative income distribution in the population of interest were also taken into account. Laird et al. (2013) provided a peer review of this procedure and concluded that this updating framework put forward by Bates (2008) was implemented correctly.[11]

Option 1.2, as presented in this report, corresponds to the above re-basing practice. It is important to note that:

- assuming the behavioural model remains fit-for-purpose in combination with;
- using new nationally representative NTS data;

will change the base year of the resulting VTTS to the newly assumed base year. For the 2013 update, the base year changed from 1994 to 2010.

When changing the base year, some assumptions need to be made, in particular in relation to the income, cost, distance and time elasticities which are measured relative to a base level (see equations (2) and (4)). The key question is whether and how these base levels should be adjusted over time. In the original studies, these base values represented mean (or median) income, travel time, travel cost and distance in the population.

Adjusting the denominators over time would imply that only *changes* in the distribution of the respective variables in the population would be taken into account. Increases in, for example, average trip characteristics will therefore not be taken into account although such effects have been shown to be of influence in the 2014/15 study. Such effects could alternatively be captured by additionally adjusting for GDP per capita growth (as described in Option 1.1), which would require an adjustment of the GDP growth income elasticity (most likely <1) since certain effects currently controlled for by the income elasticity in Option 1.1 will now be (partially) captured through updating of other control variables through the NTS. This again calls for improving our understanding of what the GDP growth income elasticity represents.

Not adjusting the denominators over time would imply that the impact of income growth and changes in trip characteristics on the VTTS would be taken into account directly. A correction that does need to be implemented is that inflation corrections need to be implemented ensuring that (base) price and income levels are presented in the value of the newly adopted base year. Not adjusting the denominators implicitly assumes that the cross-sectional income, cost, time and distance elasticities can simultaneously be interpreted as inter-temporal elasticities, an assumption that is often contested in the literature. This is therefore a clear area of future research.

[11] Bates (2008) did some early analysis on the effect of updating using new NTS trip patterns and he did not get the growth found by Wardman in the meta-data. Therefore, he concluded that it needed to be applied together with the Departments default approach.

It appears that in the 2013 updating procedure mentioned above, the Department controlled for GDP per capita income growth (Option 1.1) whilst maintaining the same set of denominators (after controlling for inflation). Even if the GDP growth factor only represents an income effect, the Department could have exposed itself to the risk of double-counting. Simply put, with part of the population shifting to higher income bands whilst not adjusting the income denominator, this income effect is already reflected in equation (2), possibly making equation (6) redundant.

In the 2014/15 framework, other population and trip characteristics were introduced in addition to the elasticity terms (see equation (5)). These control variables were not measured relative to a base value and any change in them will therefore directly affect the nationally representative VTTS function. In effect, it is assumed for example that people moving into a new age cohort will adopt the preferences of those who had previously in that age cohort (i.e. behavioural parameters are held constant). Again, the question remains as to how such changes relate to the GDP growth factor.

Most control variables used in the 2014/15 behavioural framework are directly observable in the NTS. However, this does not apply to all of them, in which case averages based on the SP data were used for appraisal purposes in the 2014/15 study. This includes the percentage of time spent in light and heavy congestion, but also the frequency of the trip, seat reservation effects and whether the costs were covered by the employer. Adjustment of these non-NTS covariates over time therefore needs to be considered.

The discussion above makes clear that although new NTS data becomes available at regular intervals and can be implemented in the behavioural framework relatively easily (practicalities will be discussed below), certain delicate decisions need to be made on adjustment of the base levels and whether all possible sources of changes in the VTTS have been captured. The outcome of this process determines whether additionally controlling for GDP per capita growth is required, most likely using a smaller income elasticity if the base year of the VTTS were adjusted at the same time.

Mackie et al. (2003) made use of six years of NTS data, respectively 1995-2000, in establishing their base VTTS for non-work purposes. In 2013, these six years were replaced by three years covering 2008-2010. The 2014/15 study also made use of a three-year time frame, 2010-2012. We **recommend** that when re-basing the VTTS a three-year period should be maintained to ensure national representativeness of the sample.

Re-basing the VTTS can be done on an annual basis by replacing the most distant year by a newer year, e.g. replacing 2010 with 2013 data. As a result, Option 1.2 can be implemented on an annual basis, however, certain data modifications would need to be implemented from the source NTS data. When this is combined with additional corrections for GDP growth, it may be more feasible to implement new NTS data on a less frequent basis, e.g. every three or five years.

It needs to be stated that by moving the base year of the VTTS further away from the year in which the SP (or RP) data were collected, the risk increases that the behavioural framework is no longer fit-for-purpose. This highlights the importance of running a monitoring scheme (i.e. Stream 2) parallel to Stream 1.

In the 2014/15 study several changes to the NTS data needed to be made in order for the data to be made applicable for the behavioural framework and in the calculation of trip weights needed for equation (5). Such adjustments would need to be made every time new NTS data are prepared for the Implementation Tool. A few data modifications are listed below:

- Uplift the weights on public transport modes because of missing cost information, based on the number of observations removed.
- Stage cost figures need uplifting for inflation. These figures will need to be consistent with any other inflation corrections implemented. In 2014/15 the inflation factors were taken from https://www.gov.uk/government/statistics/gdp-deflators-at-market-prices-and-money-gdp-march-2015-quarterly-national-accounts
- Car fuel costs are currently calculated based on 2011 figures and parameters – these will need adjusting over time (see WebTAG Table A 1.3.13)
- Income figures need to be adjusted over time:
    - There is a need to calculate medians of household income categories.
    - McClements weightings are applied to household income to create equivalent income figures and these weights might change over time.
    - Personal income, wage and hours (separately for part-time and full-time) figures are calculated based on average hours worked, wages and incomes in each income category. Adjustment of these figures is needed over time.
    - Finally, derived representative personal and household income values are uplifted for real GDP growth and inflation with respect to the base year.

As a result, a few days of work would be needed in order to prepare the updated NTS data for the Implementation Tool. Most of the actual adjustments can be automated, but the respective data sources need to be checked every time. An alternative challenge is posed when NTS, or other data sources, change the definitions of the respective variables such that conversion factors may need to be developed.

Implications of Option 1.2 for the updating programme

In summary, Option 1.2 is considered a feasible option from a practical perspective. New NTS data can be fed through the Implementation Tool and to re-base the VTTS possibly in combination with GDP growth corrections. However, decisions would need to be made in relation to a number of issues, so as to ensure that the Department captures all relevant changes in the VTTS and avoids double-counting.

The procedure described in Option 1.2 applies to uprating to a past year. A forecast procedure for the next 60 years, e.g. using Option 1.1 is still required.

Actions needed to develop Option 1.2 for implementation:

- When updating non-work VTTS we feel that it is appropriate to use as much available information as possible. This would suggest using the latest version of NTS data to update non-work VTTS for current socio-demographic and trip making characteristics.
- The difficulty with doing this is that the observed inter-temporal growth in VTTS does not distinguish between the different sources of growth. Furthermore, at this point in time we do not understand whether the finding by Bates (2008) still holds – that is whether changes in socio-demographics and trip making over time are not sufficient to impose a significant influence on the VTTS. This may be model specific and with a different behavioural model this finding may not hold.
- We therefore suggest using historic NTS data to estimate historic VTTS (e.g. for 1994, 2002 and 2010), but using the 2014/15 behavioural model. This would allow comparison of the resultant growth rate to the observed meta-analysis figure. From this one could make a recommendation as to whether the observed aggregate growth rate in non-work VTTS from the meta-analysis should be applied unadjusted or should be adjusted. If it were to be adjusted, the analysis may be able to shed some light on how it should be adjusted.
- There is also scope to better understand changes in business VTTS over time in two ways. First, undertake more detailed analysis using NTS data of how trip patterns for business users have changed over time. Second, separate out inter-temporal growth in the VTTS of business trips from non-work trips. This would need to control for trip distance – as business VTTS is sensitive to trip distance (through reference time and reference cost in the behavioural model).

Additional actions needed when implementing Option 1.1:

- New NTS data and corresponding data preparations
- Decisions on how non-NTS data control variables change over time
- Additional research on how changes in the VTTS as a result of new NTS data relate to Option 1.1 This may call for a change in the income elasticity parameter.

Actions needed in Phase 2 of this study:

- Extend the academic consultation listed in Option 1.1 to allowing for additional control factors in addition to GDP growth, i.e. through the NTS.

Option 1.3: Applying forecasts on the NTS data for future VTTS values

Option 1.2 proposed the adjustment of the base year of the VTTS over time using new NTS data possibly in combination with an additional correction for income growth. Practical applications of the VTTS through WebTAG, however, requires future values that can be used for the next 60 years. Option 1.2 would recommend predicting these values using the same income growth path as would be used under Option 1.1. Thus, new NTS data would only be relevant when changing the base year. Note that the GDP income elasticity applied might be different between updating the base year in Option 1.2 and respectively uprating the future VTTS values. This is a direct result of the GDP income growth factor possibly capturing some of the effects reflected by the NTS data (see Option 1.2. for a discussion on double-counting).

In Option 1.3, future VTTS values would be derived using forecasted 'NTS data'. The benefit of such an approach is that future changes in the VTTS can be attributed directly to particular sources, such as an ageing population, instead of a generic GDP per capita growth term. For certain NTS variables, such as age, gender and household composition, reasonable forecasts may exist. For other variables, such as the number and type of self-employed, the number of part-time workers and particularly certain trip characteristics may be associated with larger degrees of uncertainty. Forecasting the non-NTS variables would likely prove to be even more challenging. Considerations that applied to Option 1.2 also apply to Option 1.3 making this altogether a very challenging option. In discussions with the Department it was therefore decided that Option 1.3 would introduce too much uncertainty and is currently not deemed to be feasible and therefore not explored further.

If such forecasts would be deemed feasible in future years, we would see it necessary for these forecasts to be consistent with trip and socio-economic forecasts made through, e.g. the National Transport Model (NTM) or the National Trip End Model (also known as TEMPRO). A key question is whether this is possible, since TEMPRO only provides forecasts up to 2051.

Summary – Stream 1

Stream 1 provided three options to account for inter-temporal changes in the VTTS. The assumption made across all three options is that the behavioural framework, as estimated on the 2014/15 SP data, remains fit-for-purpose. A key conclusion is that inter-temporal changes in the VTTS are not well understood and it is hard to attribute these changes to specific factors. Option 1.1, which adjusts the VTTS solely for growth in GDP per capita and is a continuation of the current baseline, is considered a feasible option. However, the interpretation and size of the income elasticity parameter deserves revisiting. Option 1.2 adjusts the base year of the VTTS by using new NTS data. This could potentially be a feasible option when its connection with GDP per capita growth is better understood. Additional research would therefore be required and specifically the approach of Bates (2008) updating as implemented by the Department in 2013 (Laird et al. 2013) needs revisiting. Option 1.3, which further extends Option 1.2 to future years and relies extensively on forecasts, was not considered to be feasible due the large degree of uncertainty associated with forecasting many individual variables. Irrespective of which adjustment procedure is selected, the validity of this framework is likely to diminish over time. The behavioural model was based on 2014 SP data and the more time passes the less likely the behavioural framework matches with current preferences.

5. Stream 2: Monitoring changes in the VTTS using emerging evidence and new data

Stream 1, as described in Section 4 of this report, assumed the behavioural framework remained fit-for-purpose. This assumption may hold in the short term, but is unlikely to hold for the next twenty years. It is therefore essential that alongside Stream 1 a monitoring scheme is developed that tests whether preferences are changing in society and whether a new large-scale data collection effort is required. Waiting twenty years before commissioning such a new study, as was done before the 2014/15 update, is not advisable.

Stream 2 is structured as follows. First, Option 2.1 bundles a series of *horizon scanning* activities allowing the Department to:

i) develop a picture on societal and technological developments that may drive preferences and the VTTS;
ii) observe changes in the VTTS in newly emerging empirical evidence;
iii) identify whether new state-of-the-art data collection methods are being adopted in the academic, policy and grey literature;
iv) identify whether new state-of-the-art analytical methods are being adopted in the academic, policy and grey literature.

In particular, activities three and four allow the Department to commission a new state-of-the-art collection and analysis efforts allowing for better elicitation and estimation of the VTTS when deemed necessary without needing a lengthy consultation process.

Second, alongside relying on emerging evidence, we **recommend** that the Department develops its own monitoring scheme verifying the continuing validity of the 2014/15 behavioural framework.

Option 2.1: Horizon scanning

Option 2.1 bundles a series of independent horizon scanning activities which are described in more detail in the following subsections.

Option 2.1.1: Behavioural and technological change

Option 2.1.1 provides a general reflection on societal and technological developments that may induce changes in the VTTS and the need to consider data collection to understand potential changes. Option 2.1.1 draws on two main sources of information:

1. The ‘Travel in Britain in 2035: Future scenarios and their implications for technology and innovation’ report released in 2016 for Innovate UK.
2. The Commission on Travel Demand, which has solicited, through a call for evidence and expert events, evidence and views on reasons for and future directions of changes in travel demand.

Both bodies of work draw on academic and grey literature. There is no ‘right’ or ‘definitive’ answer about what the future will look like. Other futures exercises may

come to different conclusions. However, these two substantive and recent pieces of work provide key elements closely related to inter-temporal changes in preferences.

The study conducted for Innovate UK (Rohr et al., 2016) considered a range of different technological developments which can be anticipated out to 2035 and, through a process of expert workshops, developed three contrasting scenarios which featured the technologies to differing degrees and in different combinations.

Driving Ahead – This scenario assumes that fully autonomous vehicles have become normal by 2035 and that this helps to transform a number of aspects of society such as healthcare and retail as well as contributing to GDP growth.

Living Local – This scenario puts digital substitution for travel to the fore with advances in ICT, user apps and the Internet of Things making telepresence more normal. This is coupled with a shift in environmental attitudes towards living more locally and a social norm against long distance travel.

Digital Divide – This scenario has lower than trend rate GDP growth and rising income inequalities which are also reflected in rising inequality in access to ICT services. There is an extensive peer to peer and sharing economy in this scenario.

Table 1 below identifies some of the key aspects of each of the scenarios that might be relevant to thinking about vehicle miles travelled. The implications for freight transport have largely been set aside, although the value of travelling for retail is discussed. The next three subsections collate the information from the table into three generalised topics for the Department to consider in relation to changes in the VTTS. The supporting evidence on these topics is drawing on the work of the Commission.

2.1.1.a Automation

One of the key areas of interest is the shift to automated vehicles (AVs). This will likely happen in a phased manner and will have different implications for VTTS:

1) Increasing opportunities to relax whilst remaining in supervisory mode during motorway driving (sometimes referred to as Level 3 automation 12);
2) Fully autonomous motorway and inter-urban road driving with supervised urban driving (referred to as Level 4); and
3) Fully autonomous driving everywhere (Level 5).

It would seem sensible to consider the extent to which 1) impacts on journey quality. Intuitively one would expect a lower discomfort penalty for driving under these conditions. An initial look at the literature suggests that the focus of studies to date has been on understanding the willingness to pay for the technology (e.g. Daziano et al., 2017). Little research was found on how the technologies were subsequently interwoven into the driving experience and therefore whether or not they had or might be expected to have an impact on journey experience and therefore VTTS. Cyganski et al. (2015) looked at four different use cases to try and develop an understanding of

[12] https://www.sae.org/news/3544/

the perception of the impacts of automation. They note however that "*the lack of empirical examination and the inclusion of important aspects about the usage context, constraints and perceived benefits make predictions difficult.*"

Table 1: Travel in Britain in 2035 – Scenario implications

Driving Ahead	***Living Local***	***Digital Divide***
Full AVs (Level 5) available from mid 2020s	Autonomous vehicles are not widespread	Full AVs in place but only at very high end of market
Full AVs assumed to allow people to live further from work with more productive commutes	Road pricing has become sophisticated and effective	Far fewer people work in regular employment
Scheduling algorithms mean shared vehicle commutes commonplace	Telework finally reduces amount of time people spend at work and commute journeys	For those in stable employment most work is done remotely.
Working hours are longer as well as commute	Younger people have a decreasing tolerance for 'unnecessary travel'	30% of population over 65 due to low immigration and low birth rates
Business meetings largely conducted in person	Small but growing minority with strong preferences for local sourcing and living with on-line for other things	Not enough resources for remote telemedicine to achieve widespread adoption
Less active travel and public transport use	More home working and community hubs means fewer central offices	Much greater individual consumer incentives for activities via apps
Full AVs used for last-mile deliveries for growing on-line sector	Rapid rise in on-line sales supported by increased personalisation of experience	Retail sector not doing well. On-line growing and much greater peer to peer to reduce costs
Active leisure time has decreased but disposable income is high	Leisure time increasing with trend to fewer but longer holidays	Low preference for sharing small vehicles so conventional public transport continues
Full AVs are increasingly shared and specialised rather than owned (in urban areas)	Preference for rail vs air on environmental grounds	Some cities have cordon pricing
Users summon vehicles to a given schedule	Climate legislation has made it necessary to levy road pricing fees	

For 2) the extent to which additional benefits accrue might depend significantly on how easy people find it to work whilst in motion, particularly in a car rather than on a train. There are other differences in comfort and opportunities to move around between a car and a train also which make the direct potential substitution of 'working on the move' on a train to in a car too simplistic. Whilst easy to understand, this situation may be difficult for many people to imagine and so some more experimental trials might be desirable. This could be done through simulator based studies – although the simulator used would need to have a realistic sense of motion cues (rather than being

fixed based) to get some realistic feeling for this. It may also be possible to look at proxy data for those people being chauffeured or using a taxi.

From the perspective of a CBA, not just the changes in the VTTS resulting from increased automation will be of importance, but also the demand implications (i.e. forecasts) resulting from changes in road capacity, speeds; and adoption rates of automated vehicles. A lower disutility of travel time for car trips will not only lead to a lower VTTS for car trips, but also increased demand for car trips (partly at the expense of other modes), and the CBA needs to account for both effects.

Full automation for all tasks seems further into the distance but the Innovate scenarios have this in place by the mid-2020s and so discussion of whether this is part of the near term VTTS programme is necessary.

2.1.1.b Integration of services

The next set of innovations to consider is the shift towards Mobility as a Service (MaaS). It is important to acknowledge that there is not one agreed vision of what MaaS is and how it will be achieved (TSC, 2016). However, in its most complete form, people choose a mobility service provider, pay a regular fee for access to a basket of mobility services (from hire cars to bike share) and have a tailored 'optimal' service offered to them for each journey. The user simply types in an origin and destination and then selects from the choices given. MaaS is seen as being capable of replacing car ownership for many, particularly in more urban areas.

Figure 2 below shows the space in which the MaaS proposition could unfold from the current position of multi-modal journey planners and integrated ticketing systems.

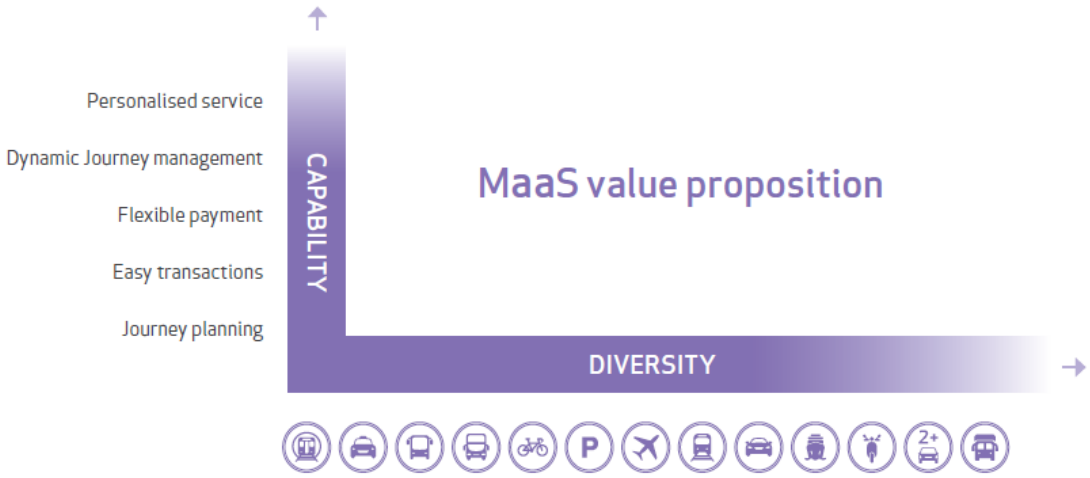


Figure 2: Transport Systems Catapult MaaS mapping (TSC, 2016, p10)

The changes which more complete forms of MaaS bring. Some key differences with implications for VTTS include:

1) Ease of transaction and payment will change and will likely be hidden in a monthly bundle making the trade-offs between costs of journey and journey times far less transparent.
2) The services are potentially on-demand and the integration of services to provide options is all done behind the scenes. This seems likely to impact on how the whole journey is seen. Lateness will not be relative to current forms of timetable but an algorithm's estimate across the whole journey. Lateness penalties may become less relevant within journeys as the MaaS system is likely to re-calculate and optimise within journey (e.g. replacing a late bus leg with an Uber or Lyft type service)
3) Waiting time will be increasingly certain as, with innovations such as Uber and more recent public transport tracking apps, the arrival of the connecting service is known to the user.
4) Reliability of journeys would also need to be understood from a door to door perspective rather than a more unimodal approach as adopted today. In reality, there would be more of a mixed approach required as some journeys would continue to be largely unimodal but others much more multimodal and integrated through the platform rather than the user. This would require overcoming some of the issues with different measurement traditions currently applied to different modes (Marsden et al., 2014).
5) The preferences of users will be able to be tracked by inferring trade-offs between journey times and interchange points. They may also be able to be tested through incentives and offers.

It is worth noting that full versions of MaaS often also include a presumption that, at some point, Level 5 shared autonomous vehicles will form part of the solution. This would further complicate consideration of how journey time is treated relative to in-vehicle time or waiting time.

As with automated vehicles, we foresee that MaaS will evolve in phases and it should be possible to infer shifts in, for example, interchange penalties on the VTTS as information on services improve as simplicity and certainty gradually improve. The increasing integration of services will, however, also increase the challenge of matching RP and SP outcomes when real-world prices become more bundled in smartcards for multi-modal journeys whereas SP choices are aimed at unbundling.

2.1.1.c Categories of Journey Purpose and societal developments

The notion of a business value of time, commute value of time and leisure value of time rely, to some degree on a clear distinction existing between the categories. Of course, previous work includes some blending of effects with both relaxation and work being possible on train journeys for work for example. There is however, growing recognition that the purposes of trip categories are becoming increasingly blurred.

Commuting: The Department has commissioned work by Imperial College into commuting. In the evidence to the Commission on Travel Demand, the Department

reported that "*the relative decline in commuting trips has been greater than the decline in overall trips, with commuting trips falling from 7.1 journeys per worker per week in 1988/92 to 5.7 journeys per worker per week in 2013/14*" (DfT, 2017). The evidence pointed to an increase in trip chaining and, therefore a decline in "traditional home-work commutes" as well as a rise in those not having a usual fixed workplace. This is identified as likely to accelerate in the Digital Divide Scenario. More people are also able to work from home which makes assessing trade-offs on an individual journey compared with a bundle of journeys relevant. All of these factors serve to complicate the derivation of a VTTS specific for commuting.

These changes are happening slowly but are continuing over time. In the Live Local scenario there is a major acceleration of home and remote working. In addition, where full automation is possible (Driving Ahead) then work is assumed to be possible on the move, which renders the distinction between home and work less meaningful for those in some sectors. Journey time may remain a factor where there is a strong face-to-face meeting culture (Driving Ahead) but would perhaps be less important where virtual presence is the norm (Live Local) where journey comfort factors might become relatively more important.

The blurring of the meaning of the commute can be explored through context dependent investigations of the population today. It may also be possible to infer whether there is a blurring of business and commute or leisure and commute with the advances of technology by examining what happens on public transport. This however will need to be tempered with understanding the quality of the journey experience in many public transport peak hour trips.

Retail: Since 1995/97 there has been a 27% decrease in shopping trips per capita and a 19% decrease in distance travelled. Much, but not all, of this has happened over a period where there has been a rapid rise in on-line retail. It has been accompanied by a significant rise in LGV traffic, although only a fraction of this can be attributed to last-mile retail logistics. A key point to raise here is that shopping has become much more fragmented over this 20-year period. The acts of browsing, comparing, buying and receiving goods have always been able to be done separately (e.g. through multiple trips or mail order magazines and with bulky delivery or standard postal delivery of goods). However, the norm is shifting quite rapidly from goods being purchased by a visit to the store to goods being purchased away from the main stores for delivery at different points in time and in a variety of places.

Much as with commuting, the question then becomes whether what is classed as a shopping trip today is the same as before. There may be different motivations for physically going shopping and different trade-offs between the costs of going shopping versus the costs of receiving goods through on-line purchases. Whilst shopping is just one of a basket of leisure journeys it is an important one. If inferences are to be made about changing values of time it might be necessary to more fully understand what is underlying those changes (Zhou and Wang, 2014).

Ageing Population: Advances in healthcare are contributing to an ageing population. The National Infrastructure Commission reviewed the ONS population projections (see Figure 3 for core projection) and found that "*whereas 59% of the increase in population over the last 25 years has been in the working age population, 63% of the projected increase for the next 35 years is in the over 65s. Even in the high migration variant [i.e. an influx of younger immigrants] this is still 50%.*" (NIC, 2017, p18). The anticipated forecast range is for between 23% and 28% of the population to be over 65 by 2050.

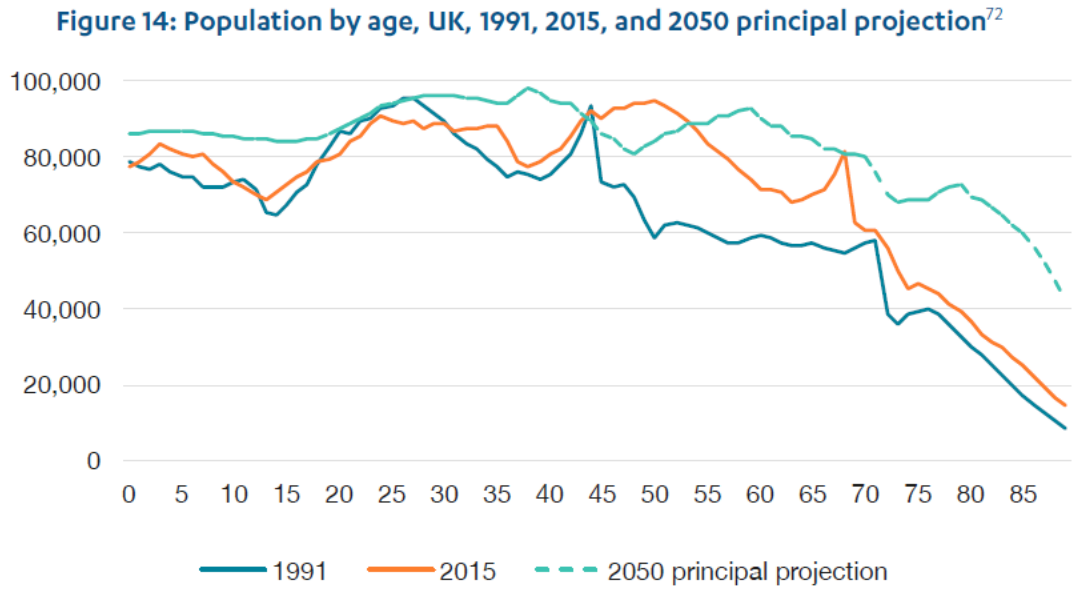


Figure 3: ONS core projections of population in 2050 (NIC, 2017, p18)

Conditions will of course change over that period. Indeed, the 2016 ONS projections have revised downwards net inward migration figures and slowed the rate of growth in life expectancy. However, the headline trends remain unchanged (ONS, 2016). It can be anticipated that people will work longer and retire later. People will live healthier lives for longer. This will particularly be associated with increases in leisure travel which is seen to be a key determinant of active ageing amongst older people (Fox et al, 2017). It is also likely that many people will experience mobility restrictions for longer periods of time in their lives. Many of the Innovate UK scenarios included substantial use of automated vehicles for independent mobility or telemedicine to provide care.

A key question to consider, which becomes more important as this group grows in share of the population, is whether the motivations of the travel of older people are the same as those of people of working age. For example, research as part of the DEMAND Centre has shown that there is great value placed on getting out and being seen out. This may be one, amongst several, of the reasons that this age group makes so many more shopping trips than that of younger age groups. The ability to share experiences of places visited (near and far) is part of demonstrating an on-going vibrancy (Fox et al., 2017). There could, at least for some journeys, be significant

intrinsic value in the journey itself as well as the activity, which may make it necessary to consider new ways of eliciting the preferences around such journeys. It is also understood that not travelling is a source of loneliness and isolation and so there will also be some potentially important lower limit thresholds where journeys are much more valued.

Changing Preferences: Much has been made of the declines in driver license holding and travel amongst people aged 18-35 (the Millenials). In particular, the decline amongst young males has been quite rapid. One of the key debates has been the extent to which this decline has been the result of economic factors or a change in preferences. McDonald (2017), in her evidence to the Commission on Travel Demand, suggests that roughly half of the decrease in Millennial travel is explained by traditional economic and demographic variables. The remainder, it is argued may be attributed to unobserved factors and changing preferences. However, Bastian (2017) argues that there is no need to separate out changing preferences from changing economic circumstances as the relationships are complex and it is difficult to assign causality. To that end, it is necessary to consider to what extent preferences are stable across cohorts and whether the change in behaviour of younger travellers is something which will persist throughout the life course. This can be explored by looking at the stability of preferences within and across cohorts over time but with greater attention to other potentially important contextual variables.

The latter notion is closely related to the core assumption underlying Stream 1. If the change in behaviour persists throughout the life course, we are observing a structural break and new behavioural data needs to be collected to ensure the new VTTS values capture these changing preferences. As mentioned under 2.1.1.a, from the perspective of CBA, the impact on travel demand is at least as important as the impact on valuations of travel time components.

Implications of Option 2.1.1 for the updating programme

The takeaway message from Option 2.1.1 is that society and technology are changing over time and this inherently affects the way in which we travel. Unavoidably, these developments will affect the VTTS. It is therefore important for the Department to be aware of the main developments and considers the speed at which these developments are happening, especially when providing guidance on VTTS values for the next sixty years. Of course, such developments are associated with a large degree of uncertainty, but we see significant value in making assumptions (or developing alternative scenarios) on the implications for the VTTS, including the appropriateness of maintaining the current level of segmentation. A simple correction of the GDP growth elasticity could already make a large difference.

There is a risk here in talking about structural breaks and tipping points since most of the developments happen gradually. We would rather be looking, or making assumptions, at the extent to which deployment paths are more or less rapid. If schemes are reliant on assumptions which are vulnerable down the line (i.e. when the

benefits beyond 15 years do matter significantly) then there is a much greater (or less) risk associated with that investment if rapid adoption scenarios emerge.

The updating programme should thus ensure that the Department is aware of these developments as often covered by non-VTTS reports, such as Travel in Britain in 2035 and the Commission on Travel Demand. The Department can subsequently commission more detailed work (e.g. through Stream 4) investigating the possible implications for the VTTS. We would suggest this part of the monitoring process to be thematically focused, e.g. 'Automation', 'MaaS' etc., and aimed at providing answers to the following questions:

1. What is the current state of play?
    - For automation, this would refer to the current state of technology and level of automation currently embedded (and used) in the vehicle fleet. For more societal developments, such as flexible working hours, this question is aimed at capturing the current attitudes in society towards such arrangements.
2. What necessary steps are needed to make it feasible at a national scale?
    - For automation, this refers to the technical barriers, but also legal, that need to be overcome for such technology can be implemented. For societal developments, this refers to the institutional arrangements that need to be put in place. Corresponding timelines are essential in defining whether these are short-, medium- or long-term developments
3. How will this development influence travel behaviour and are forecasts available on the significance of this development?
    - This is the key stage where alternative frameworks, such as the Innovate UK study, and scenarios need to be contrasted in terms of the assumptions made on travel behaviour (e.g. car ownership, vehicle miles travelled etc.), adoption rates (e.g. share of population working as self-employed or flexible hours) and implications for travel demand.
4. How are the identified changes related to the VTTS?
    - The third question, but also the discussion in 2.1.1.a, made clear that travel demand may change because of automation, but that for CBA purposes both the demand and the valuation effect are important. Where it is particularly likely that technological improvements, e.g. in automation and MaaS, will influence the disutility of travel time, flexible working hours might be limited to demand effects although the quality of the journey might improve due to reduced crowding levels. We foresee that answers to this question are conditional on the identified scenarios in the different frameworks.

No actions needed in Phase 2 of this study in respect of Option 2.1.1

Option 2.1.2: The role of meta-analysis in the monitoring programme

Option 2.1.2 assess how newly emerging evidence can be used to validate whether the 2014/15 behavioural framework remains fit-for-purpose, as assumed in Stream 1. A natural candidate for assessing emerging evidence is meta-analysis. Meta-analysis has already played a significant role in determining the inter-temporal GDP growth elasticity, but we believe it can also be used as part of the monitoring programme.

The latest large meta-analysis on the VTTS that was purely based on UK evidence is Abrantes and Wardman (2011). This includes 1749 values from 226 studies between 1960 and 2008. In a project for the EIB, this dataset was combined with an international (largely European) dataset on VTTS from Shires and de Jong (2009). For this EIB study, the evidence base for both the UK and continental Europe were extended to 2011. The new dataset, after accounting for double counting, consisted of 3109 VTTS values from 389 studies, of which 1862 values and 233 studies are from the UK.

This new database was analysed twice. Firstly, in the original study for the EIB, which was reported in Wardman et al. (2012), and secondly in Wardman et al. (2016). The main difference between the two studies is that the meta-regression for EIB used GDP per capita in market exchanges rate (MXR), whereas the later work used GDP per capita in purchasing power parities (PPP). In our view, PPP is the more correct approach for an international comparison, since it corrects for differences in purchasing power between countries. The distinction between PPP and MXR does not affect the comparability of studies over time.

The meta-analyses cover a large period and allow for a more detailed study of inter-temporal developments in the VTTS. The meta-regression in Abrantes and Wardman (2011) did not include time trend variables; the only explicitly time-based explanatory variable was GDP per capita, which gave a GDP per capita elasticity of the VTTS of 0.900. The distribution over time of the number of UK studies included is as follows: fewer than 10% are pre-1980. After that we have 20-30 studies in every 5-year period, except for 1990-1995 with 68 studies. Moreover, the early studies have fewer values per study than the later studies. The database for the EIB and Wardman et al. (2016) studies has 2% of the values pre-1980, 17% in 1981-1990, 42% in 1991-2000 and 39% in 2001-2011. Again, these studies have no decade-specific dummies or dummies for other clusters of years[13] and they do not have the year (since the first study) as an explanatory variable either (time trend), but they do include a time-varying GDP per capita term. Notably, the income elasticity was 0.721 (for MXR) in the EIB study and 1.031 (for PPP) in Wardman et al. (2016).

[13] Shires and de Jong (2009) did include dummies for specific 5-year periods and found that these were not significant for commuting, but for employer's business they obtained higher values before 1990 and in the period 2000-2005 relative to the base of 1990-1999 (a period of relatively fast GDP growth). For leisure travel the time period dummy for 2000-2005 is significant and positive, indicating an increase in the leisure VTTS after 1999.

The meta-analyses cited above have been used to calculate VTTS for three journey purposes (commuting, business and other) and for car, train, bus and air transport. Wardman et al. (2016) also distinguishes between urban free flow, urban congested and inter-urban free flow for car (by purpose). These distinctions are possible because the meta-regression includes coefficients for journey purpose and for mode used. Other coefficients in the model relate to specific time attributes (relative to in-vehicle time), mode valued in the SP, numéraire used in the study, SP presentation method, SP replications, choice context, data type and country. These variables could also be used to calculate different values of travel time savings, but since these are not really journey characteristics they do not offer added value for segmentation by journey characteristics. Mode and purpose could in future estimations be interacted with time period to test whether there are specific trends for specific mode and purpose combinations as long as sufficient observations are available.

Implications of Option 2.1.2 for the updating programme

The last comprehensive meta-analysis for the UK alone covered the period until 2008. Our **first recommendation** would be for the Department to update the UK database and the corresponding estimation of the meta-analysis using studies for the period 2009-present, including the values from the 2014/15 study. At the same time, additional sample statistics and study characteristics from currently included studies can be collected.

Our **second recommendation** would be to continuously extend the database, e.g. on an annual basis. A soon as a sufficient number of studies (say at least 25 studies for the period 2015-2019) in the UK have become available, two things can be done:

1) Compare the (average) VTTS values for the new period with the range provided by i) the meta-model for the UK as well as ii) the VTTS from the 2014/15 study. For the VTTS from the meta-regression, we can calculate the resulting VTTS for a new year by inserting the inputs for the UK for that year. For getting a future year benchmark from the 2014/15 study, it would be best to look at both the original value (only corrected for inflation) and at the upscaled value on the basis of the methods discussed in Stream 1 (e.g. correcting for real income growth by employing a GDP per capita elasticity to the VTTS). These analyses can be complemented by calculating confidence intervals around the official VTTS and around the meta-model results for the UK.

   Wardman et al. (2016) includes a comparison of the implied VTTS from the meta-model for the Netherlands, Germany, France and the UK with values from the latest official national VTTS studies for these countries which were all published after 2011 (see Table 12). The conclusion was that the degree of correspondence between the meta-model values and official values was acceptable. The above exercise thus provides two good benchmarks to contrast newly emerging evidence and to statistically test whether the 2014/15 behavioural framework remains fit-for-purpose.

2) Include the new VTTS values and explanatory factors for the new period in the meta-analysis dataset and estimate a new meta-regression that includes either:
    a. a time period dummy for the new period relative to the old period, or
    b. different GDP per capita coefficients for the new and preceding periods.

    These two options can be combined. If the time period dummy is found significant, it is likely that there has been a structural change in the VTTS not associated with income, which would suggest doing new data collection and analysis (especially given that in the past the time period dummies were not significant). If only the income coefficient has changed, adjustment of the upscaling factor may be sufficient.

In Wardman et al. (2016), 60% of the observations are from the UK. There seems to be much more VTTS-related research in the UK than in any other country in Europe. A lot of this has only been reported in the grey literature, e.g. consultancy reports for train operators or for updating PDFH. Our impression is that this difference is not only due to the bigger data collection efforts carried out in the UK for doing meta-analysis in this field, but also that this topic is studied much more frequently in the UK than elsewhere. For specific segmentations and journey characteristics, our **third recommendation** is to also include non-UK data since the UK evidence base may be considered limited. If one wants to test interactions between period on the one hand and mode and purpose on the other hand, many observations will be needed and again a wider evidence base than the UK may be deemed appropriate. Keeping the non-UK dataset up-to-date increases the sample size by a factor of about 2/3.

Actions needed to develop Option 2.1.2 for implementation:

- Apart from the above mentioned recommendations, the meta-analytical framework is ready for implementation

Additional actions needed when implementing Option 1.2.1:

- Besides collecting and analysing new VTTS evidence in and outside of the UK, no further actions are required.

Actions needed in Phase 2 of this study:

- Scope an update of the current UK based VTTS meta-analysis to present
    - Data collection and estimation
- Scope annual data collection efforts for newly emerging evidence as part of the updating programme. No new SP/RP data collection needed.
    - This includes assessing the new evidence for quality
- Scope option 1): contrasting the new evidence against the official and VTTS values
    - Possibly an annual exercise
- Scope option 2): updating the meta-analysis
    - Most likely once every 3-5 years
- Scope the latter three items with and without using non-UK evidence.

Option 2.1.3: New RP data collection methods

Initial attempts, from Beesley (1965) onwards, to estimate VTTS relied on RP data. Partly this was because SP methods did not become current until 20 years later, but partly it was because classical micro-economists believed that only RP methods allowed valid inferences to be drawn about consumers' preferences. Nevertheless, the small size of the RP data sets available in that period meant that VTTS estimates were subject to considerable estimation error.

The development of SP methods, gradually focussing on Stated Choice (SC), offered greatly improved confidence limits, even when these were calculated more correctly by taking account of the correlation between multiple responses from a single individual. A SC methodology has become standard and applied in successive studies in western Europe, in particular in the UK. This is currently accepted as best practice.

Nevertheless, concerns remain about the validity of the SC approach. Centrally, the economists' concern about the validity of non-RP methods has been generalised into a worry about the hypothetical nature of SC and other stated preference methods. More specifically, experience with VTTS estimation using SC data has revealed a wholly unwelcome dependence of the result on the design of the experiment, particularly on the size and sign of the time differences that survey respondents are asked to value. The development of Prospect Theory suggests that these issues may apply in the 'real world' and not just in the hypothetical realm of the SC experiment, but the extent to which this is true remains unknown. Certainly, size and sign differences in VTTS are unacceptable to governments seeking to appraise transport policy and projects.[14]

In this context, it is worth revisiting the decision to abandon RP estimation of VTTS in the UK context. The original issue that RP data volumes were too small to make reasonably accurate estimates has been overtaken by the collection of much larger home interviews and other RP data sets. These describe mode, destination and other everyday choices made by travellers and are routinely used to estimate large-scale travel demand forecasting models; good estimates are generally made of the time and cost coefficients, despite the correlation in the relevant variables. From these models, VTTS estimates can be inferred, which are currently used only to test the validity of the models, by comparing the inferred VTTS with accepted (SC based) estimates, though the SC estimates typically cover fewer modes and have a less detailed purpose specification. It is worth considering whether these inferred RP values could be also used as independent VTTS estimates.

[14] See Daly et al. (2014) for a review of governments' practices in this respect. In practice, sign differences are eliminated by averaging the gain and loss values (often, a geometric average) though this is an assumption that is difficult to justify. However, size differences cannot be eliminated at all easily and an assumption of a specific amount of savings needs to be made; this is frequently set at 10 minutes, despite evidence that time changes brought about by transport policy are usually considerably smaller than this.

In considering RP data for use in this context, initial attention focusses on the large-scale home interview surveys that have been conducted in many countries to gain insight into current travel behaviour and as a basis for modelling work. In these surveys, it seems reasonable to assume that the reported choices (i.e. mode and destination) are reliable, but that the description of the alternatives that the traveller was facing are a cause for concern.

One aspect of this concern is the individual traveller's view of the alternatives. Which alternatives were 'available', i.e. could 'in principle' alternatives be considered? Which alternatives were actually considered? Did the traveller have a good assessment of the characteristics of the alternatives, in particular of the time and cost that would be spent in using them? There is some research in this area, which suggests that some of the concerns might be justified, but further work would be needed to reach conclusive findings. However, it may also be questioned whether it is necessary to consider these issues at all: they do not form part of a classical economic analysis of a market, which works on the basis that consumers behave 'as if' they have perfect information. At least it seems we are moving closer towards a situation of perfect information with travel times from Google Maps nowadays being accurate for most modes, and travel costs can also be obtained accurately by automatic queries to e.g. public transport web pages.[15]

A second aspect of concern is the measurement of the characteristics of the alternatives and in particular of the time and cost. Early RP data collection relied on travellers reporting of the attributes of the alternatives, but this is easily shown to be unreliable, subject to rounding and self-justification, in particular for unchosen alternatives. Large-scale modelling relies on transport networks to deliver the required attribute data; this may be acceptable for forecasting, the main application of the models, as the forecasts will also be based on transport networks, but concerns remain when we infer VTTS from the specific time and cost coefficient values. It should be clear that neither reported nor calculated values exactly represent perceived or 'true' values of the required attributes. Also in this context, the use of Google APIs could prove useful in future studies to verify information and add obtain information on the non-chosen alternatives with supplementary information being drawn from railway pricing sites. Note that such information will still be associated with some degree of measurement error. This would need to traded off against the advantages of the choices being "real" choices.

The accuracy of time and cost measures derived from transport networks has been the subject of a small number of recent studies that use an approach based on latent

[15] Inferring travel costs remains a challenging tasks and explicit assumptions need to be put in place. For example, what costs are considered for car journeys, e.g. only fuel cost or also parking costs? Who pays for these car costs and how are these distributed across car passengers? Similarly, for public transport seasonal passes form a challenge where the question should be asked if making this trip was foreseen when buying the pass. These issues, however, to a very large extent also apply in SC data, e.g. in terms of what people actually consider as driving costs.

variables that seems very promising. In these studies, it is postulated that the true values of the attributes are unknown, but that network and reported values give 'indicators' of these true values. By setting up model systems that explicitly postulate that both network and reported values differ both systematically and randomly from the true values, as well as taking account of observed choices, it is possible to describe the degree of error in the indicators and to estimate a choice model from which VTTS can be inferred from the parameter estimates based on the 'true' time and cost values. A current as yet unpublished study at the Royal Institute of Technology (KTH) in Stockholm (Lorenzo et al. 2017) is taking a broader approach than that of the previous studies and has achieved very promising intermediate results from home interview data collected in Stockholm. It is hoped to take the study further to derive good VTTS estimates from the RP data being used.

This latent variable approach could be applied for national VTTS estimation in any country that has a national travel survey and transport networks to generate time and cost values for the journeys reported. In the UK, the NTM relies on information of this type and the latent variable approach could in principle be applied to drive good VTTS estimates. However, one would want the Stockholm approach to be developed further and to appear in the peer-reviewed literature before applying it on that scale.

The latent variable approach applies data from home interview surveys and transport networks. One could also think of other data sources, whether automatic or semi-automatic, such as GPS-based data. Some initial studies yielding VTT estimates have been made of route choice by HGVs in Britain and of travellers' mode choice in Italy (Hess et al., 2015, and Calastri et al., 2017, respectively), but again one would want further development of the methods before applying them for practical VTTS estimation on a national scale.

Implications of Option 2.1.3 for the updating programme

Option 2.1.3 makes clear that traditional travel survey data, but also new data sources becoming available in this era of 'Big Data', such as smart-ticketing or GPS-based data, combined with new modelling approaches are expected to revive interest in RP data collection methods in the near future. The wealth and volume of information included in such sources is promising. The applicability for national VTTS purposes, however, comes with several considerations:

- National (and representative) coverage needs to be ensured.
- Data collection methods may vary across modes and possibly across geographic areas (GPS for car; tap-in and tap-out for public transport with different data collection efforts being used in different cities). This makes the development of a consistent methodology across modes, journey purposes and quality attributes challenging.
- In most cases, additional data on non-chosen modes, price data, crowding and reliability levels and respondent info, needs to be collected.

These developments are, however, in too early a stage for adoption in a national VTTS study. We therefore **recommend** that the Department should not to take any action at this stage apart from keeping up to speed with the ongoing developments in the academic literature. If the Department takes the stand that it should take a leading role in developing the collection of such datasets, we think it should be ensured that such data can be used for multiple purposes other than the VTTS in order to justify the associated costs.

Actions needed to develop Option 2.1.3 for implementation:

- We recommend the Department to annually scan the literature identifying whether:
  - new VTTS estimation methods based on conventional RP data are becoming competitive with current SC-based estimations
  - VTTS estimation based on novel RP-based data collection is becoming the new state-of-the-art.
  - SC data collection efforts are being improved further.
- Ideally, when moving towards an RP-based evidence base a validation framework is introduced contrasting RP and SP based VTTS values.

Actions needed in Phase 2 of this study:

- No actions needed in Phase 2 of this study in respect of Option 2.1.3

Option 2.1.4: New SC data collection and analysis techniques

We believe that studies should "move with the times" and that it is not wise to maintain a given survey approach or modelling approach just to ensure consistency. Any future large-scale study should thus have the opportunity to "start from scratch" and use current best academic thinking, which would possibly include moving away from an approach where different valuations come from different games. This, however, together with new modelling approaches, will likely again lead to changes in results.

The 2014/15 behavioural framework and data collection efforts reflected the state-of-the-art at that particular moment in time. Section 2.1 which contrasted the 2014/15 framework with the Mackie et al. (2003) framework highlights the significant amount of progress that was made with respect to the analytical framework.

Both reviewers indicated that current SC data collection methods are still associated with significant concerns and that RP data collection methods, as discussed in Option 2.1.3, should therefore not be disregarded. This should, however, not prevent improvements to SC data collection efforts. We list a few areas that deserve more attention. First, the conventional VTTS SC experiment, as used in the UK, focuses on short-term route choice decisions facing travellers with trade-offs between travel time and costs, whereas appraisal VTTS measures are used to evaluate long-term infrastructure investments. Alternative approaches could look into longer term labour or housing market choices, an approach used for example in the German national study. Second, self-selection can be a significant issue in current intercept methods, but also in collecting respondents through internet access panels (e.g. Kouwenhoven et al. 2014). Such issues can possibly be avoided by implementing SC studies in the NTS or other panel surveys (e.g. Weis et al. 2017), but they would come at significant setup costs and increase the response burden.

In terms of analysis, we do not expect that over the next five to ten years the choice modelling field will introduce ever more flexibility in the choice models. However, we do expect that the methodologies used to estimate the VTTS, including the underlying behavioural effects, will be further refined and improved. A prime example is the use of latent variable models, as discussed in Option 2.1.3, where measurement error in the reported journey characteristics are accounted for in the RP data.

Implications of Option 2.1.4 for the updating programme

Option 2.1.4 is very closely related to Option 2.1.3 in the sense that it aims to keep the SC knowledge base of the Department up-to-date. Option 2.1.4 focuses on SC data collection and analytical techniques instead of RP methods as covered by Option 2.1.3. We therefore recommend that the Department, in addition to scanning

the literature for developments in RP data collection efforts, also to keep up-to-date with the state-of-the-art in the analysis of (stated) choice data.

Actions needed in Phase 2 of this study:

- No actions needed in Phase 2 of this study in respect of Option 2.1.4

Option 2.2: A monitoring scheme

A key challenge faced by the 2014/15 study was that twenty years had passed since the collection of the behavioural data. At the same time, major changes in survey techniques had occurred, as well as important developments to modelling methodology, even if the gap in time here was smaller given the re-estimation on the original data by Mackie et al. (2003). Any differences in the VTTS and related measures between the 2014/15 and 2003 studies could thus be a result of changes in travel patterns and sensitivities, changes in design approaches, and/or changes in modelling approaches.

To better understand changes in the VTTS and its related components over time that are not related to changes in survey and modelling techniques, it is important that until a new large-scale VTTS study is commissioned, there is a regular monitoring programme in place relying on evidence from new behavioural data but using the same sampling approach, survey and modelling techniques as used in the 2014/15 study. In turn, the monitoring framework would allow any new large-scale VTTS study to give an indication for what share of any changes in values observed compared to 2014/15 are due to changes in methodology as opposed to changes in behaviour.

We **recommend** that the monitoring framework collects additional behavioural data on a yearly basis. In a given year, the efforts will be focussed on two out of the four modes of transport whilst covering all corresponding journey purposes and quality characteristics.[16] The recommended framework thus ensures that the values for each mode-purpose combination and quality characteristics will be validated every other year. This is of course a high frequency of updating and the sample sizes obtained will thus by necessity be smaller than with less frequent updates. A sample size of between 10% and 20% of the original sample would be a lower limit to make this exercise viable. However, it will allow for deeper insights into if and how values change over time.

In terms of empirical work, we **recommend** that any new data is *added* to the existing estimation sample. The behavioural models are then estimated on the combined data. Estimation on the new data alone would not be viable given the small sample size. Joint estimation will allow analysts to understand the shift in values obtained by adding in the new data. The inclusion of shift parameters in the behavioural model, e.g. in the baseline value of time, accounts for potential differences between the original and added data. In practice, one would expect shift parameters to be included on the key drivers of the VTTS. If shift parameters were imposed on all explanatory variables, one would effectively be estimating the model only on the new data which, again, would not be viable. Good practice would suggest to conduct a range of sensitivity tests varying the explanatory variables with which the shift parameters are associated. As a minimum, we would recommend a shift parameter to be associated with the

[16] The monitoring scheme does not cover walk and cycling. It is considered that an appropriate methodology of collecting such data does not exist to date and this is therefore an issue better placed in Stream 4.

baseline VTTS. Such time dummies would pick-up intertemporal changes in the VTTS, but would assume the influence of the explanatory variables remains constant. The significance levels for the shift parameters provide an indication of the validity of the assumption underlying Stream 1, i.e. stability of the behavioural model. These significance levels would need to be judged against the smaller sample size.

Option 2.1.1 highlighted that inter-temporal changes in the VTTS can occur gradually and adding behavioural data for a single year may not reveal significant changes. It is therefore **recommended** that every six or eight years, i.e. after collecting 30-40% of the original sample, a robustness check is done by estimating the model using only the new behavioural data and contrast this against the 2014/15 VTTS model and values. One could question the validity of combining 6-8 years of data, but the purpose here is to identify trend changes relative to the 2014 data as opposed to estimating a new unbiased VTTS estimate. The model would be still able to capture such effects.

In making the behavioural data comparable over time, the journey cost and income levels would need to be adjusted to the same price years. Adjustment of other variables is not required. The dataset would thus pick up both cross-sectional and intertemporal variation in income. We would therefore **recommend** this as an excellent opportunity to better understand the relationship between these two elasticities. This could provide valuable information in relation to the discussions put forward in Options 1.1 and 1.2. Sensitivity testing of the behavioural parameters should, however, not be restricted solely to the baseline VTTS and (or) the income elasticity. To avoid the confounding of inter-temporal income effects with other trend variables, the analyst should include time dummies and thereby possibly pick up influences of technological change induced by e.g. automation and MaaS.

A benefit that arises from collecting new behavioural data on an annual basis is that, as long as the 2014/15 behavioural model is considered fit-for-purpose, the behavioural models can be re-estimated regularly (as recommended by Wheat and Batley 2015). The first benefit that arises is that confidence intervals are expected to become narrower over time. Additionally, the re-estimation provides the opportunity for re-basing when combined with new NTS data (Option 1.2) and thereby reduces the exponentially growing impact of uncertainty associated with the GDP per capita income elasticity parameter (See Option 3.3).

Finally, the Department has expressed interest in monitoring whether changes in the VTTS happen at a different pace across geographical areas. The 2014/15 study made use of a set of geographical variables also captured by the NTS. These variables cover trips in the following areas London-London; London-non-London; intra-urban; inter-urban (20+) and rural. In the end, control variables were only put in place for London-London and rural trips. It was not verified whether the sampling for the SC data, despite being nationally representative, was representative for each geographic area. It could be that the geography effect therefore correlates with other socio-economic characteristics. If the Department has a particular interest in studying changes in the VTTS across different geographical areas, we **recommend** that the monitoring

framework develops specific sampling strategies for the different geographic areas of interest. Namely, it is required that the sampling has enough coverage in the different areas as well as maintains its national representativeness. Most likely this would require collected sample sizes to be larger than the 10-20% of the original VTTS study.

A related consideration that needs to be made is that if the Department acknowledges that the VTTS varies by geographic area, the practical implications for treating long distance travel in appraisal would be unclear. That is to say, from a practical point of view we consider this to be an option, but we have our reservations on the implementation.

Implications of Option 2.2 for the updating programme

Option 2.2 has set out a framework for monitoring changes in the behavioural model over time and thereby validates the main assumption underlying Stream 1. The proposed monitoring programme consists of a high frequency (annual) data collection and estimation effort validating the 2014/15 behavioural model against the newly collected data through joint estimation without making changes to the original method.

Actions needed to develop Option 2.2 for implementation:

- The monitoring framework needs to be agreed together with the Department
    - Frequency, sample size etc.
- The Department needs to consider its position on (inter-temporal) geographical variation in the VTTS.

Additional actions needed when implementing Option 2.2:

- SC designs, surveys and choice models can be implemented again.

Actions needed in Phase 2 of this study:

- Conduct Delphi surveys with market research companies on the possibilities and feasibility of running such a high frequency monitoring programme.

Summary - Stream 2

With respect to developing a programme to maintain the robustness of the Departments VTTS framework, Stream 2 provides two levels of recommendations.

First, Options 2.1.1; 2.1.3 and 2.1.4 recommend that the Department should not take a step back and wait until it is decided in an ad-hoc fashion that a new large-scale VTTS study needs to be conducted. Instead, we recommend that the Department should pro-actively consider the high-level technological and societal changes that may drive changes in the VTTS beyond what would be captured by Stream 1 framework (Option 2.1.1). Simultaneously, we recommend that the Department should stay up to speed with the state-of-the-art RP and SC data collection and data analysis (Options 2.1.3 and 2.1.4) techniques. The latter two options allow for such a new large-scale study to remain of the best standards.

Second, Options 2.1.2 and 2.2 both recommend a framework to test whether changes in the VTTS are occurring and to what extent it is necessary to initiate a new large-scale test. Option 2.1.2 relies on newly emerging evidence, i.e. VTTS studies, in the UK and in Europe that allow identification of changes in the VTTS beyond what can be expected from Stream 1 and a meta-analysis framework. Option 2.2 is more involved in the sense that it requires the Department to frequently sample new SC data, using the 2014/15 framework, and re-estimate the behavioural models whilst controlling for inter-temporal differences. Despite being more involved, Option 2.2 does potentially allow the delay of the next large-scale VTTS survey. By having an evidence base that remains relevant to date and which can possibly be combined with new NTS data (Option 1.2) re-basing options and reductions in confidence intervals are benefits that may outweigh the additional costs.

6. Stream 3: Robustness of the base and updated VTTS values

Whereas Stream 1 of the updating programme aims to move forward the 2014/15 behavioural model, Stream 2 aims to provide evidence on inter-temporal changes in the VTTS through the use of meta-analyses and regularly collecting new behavioural data. In contrasting Streams 1 and 2 a challenge emerges and there is a need to assess the degree of confidence by which the outcomes of Stream 1 can be taken forward, should be combined or adjusted using Stream 2 evidence or whether a case can be made for conducting a new large-scale study. Against this background, Stream 3 aims to provide guidance on the robustness of the results and the degree of confidence the Department can place on base year and future VTTS values.

Stream 3 is divided into three options respectively looking into:

- A better understanding the sources of uncertainty
- Introducing steps to increase the robustness of the VTTS values
- Understanding uncertainty in future VTTS values

Option 3.1: Better understanding the 2014/15 confidence intervals

The background to Stream 3 is that the 2014/15 VTTS values are considered to be associated with relatively wide confidence intervals. Wide confidence intervals can be problematic in implementation and this issue thus provides a clear justification for looking further into the robustness of the estimates from the behavioural models. Here, we use the same definition as in Wheat and Batley (2015) for 'base year VTTS estimates'. In particular, this is the NTS weighted estimate of VTTS for a specific mode-purpose segment emerging from the behavioural model in the base year. For the Mackie et al. (2003) study the base year was 1994. For the 2014/15 study the base year was set to 2010.

When comparing uncertainty in the base VTTS, it is important to note what sources of uncertainty are quantified. Table 2 highlights that the 2014/15 VTTS estimates include an additional source of uncertainty within the reported confidence intervals.[17] It should therefore be no surprise that the 'new' base VTTS estimates are associated with wider confidence intervals than the 'old' VTTS estimates. The 2014/15 work, however, established that the contribution of data uncertainty to the total uncertainty was marginal compared to parameter uncertainty. The core focus of this option will therefore be on the issue of parameter uncertainty.

Before engaging in a more detailed discussion that the 2014/15 studies have been subject to a range of robustness checks and thereby satisfy the following criteria:

- The models, analytical and sampling techniques used are supported by credible, peer reviewed academic literature.
- Econometric models used have been subject to rigorous quality assurance so as to minimise the scope for error.

[17] Only the first two sources of uncertainty are relevant in determining the confidence intervals in the base year.

Table 2: Sources of uncertainty in the VTTS

Source of Uncertainty	Quantified in Wheat and Batley (2015) using the 2003 Study	Quantified in the 2014/15 study
i. Parameter uncertainty arising from estimation of the discrete choice model	✓	✓
ii. Data uncertainty associated with the weighting of the base VTTS there is a degree of uncertainty as to whether the NTS sample is truly representative of the population.	×	✓
iii. Data uncertainty associated with the forecasting of future income levels.	×	×
iv. [for appraisal values in years after the base year] Parameter uncertainty arising from estimation of the meta-model used to provide the GDP elasticity for uprating the base VTTS.	✓	×*
* Not quantified, but can be quantified without significant effort		

A number of key points arise in better understanding the degree of parameter uncertainty associated with the 2014/15 relative to the Mackie et al. (2003) model:

1. The confidence intervals (i.e. standard errors) reported in the 2014/15 study are obtained through analytical calculations, using the delta method (see Daly et al. 2012). The same rigorous and theoretically correct approach was not used for the old confidence intervals presented in Mackie et al. (2003).
2. The error measures reported in the 2014/15 study make use of robust rather than classical standard errors. Robust standard errors, as obtained from the sandwich estimator, are known to be larger than classical standard errors as they account for some level of misspecification of the model from which the estimates are obtained. A direct effect of using robust standard errors is therefore that wider confidence intervals are obtained relative to relying on classical standard errors. Robust standard errors have not been used by Mackie et al. (2003).
3. In addition to using robust standard errors, the calculation of standard errors in our models directly recognises the repeated choice nature of the data. That is, each respondent provided answers to multiple choice scenarios. At the time the previous official values were produced, techniques such as bootstrapping and jack-knifing were used for this purpose and this should theoretically account for the same correlation across choices. Mackie et al. (2003) on p17 refer to previous work by AHCG (AHCG 1994) making use of jack-knifing, but explicitly mention not controlling for this issue when re-estimating the 1994 data.

The methodological discrepancies listed above may have led the general expectation of narrower confidence intervals in the 2014/15 work. To understand the implications of the above discrepancies in methodology, we **recommend** the Department to re-estimate the 2014/15 specification using the classical standard errors methodology allowing for a fairer comparison and assessment of the size difference.

In what follows, we move beyond the method used to calculate the standard errors and evaluate the departures from the previous studies, in terms of design and modelling (see Section 2 of this report for more detail)

The 2014/15 study incorporated a number of important methodological improvements:

1. The use of efficient experimental designs should in theory lead to smaller standard errors and there is no a priori reason to expect that it would have led to increased standard errors and wider confidence intervals. A small investigation could be carried out to compare the priors used in the generation of the designs with the final estimates obtained from the models. The use of very poor priors could in theory have negatively affected the efficiency of the designs. We, however, expect this to be of minor effect since specific designs were developed for very specific mode and journey purpose combinations for specific distance bands.
2. The fact that all valuations, including journey characteristics, are obtained through joint estimation rather than from separate game specific models should in theory also lead to smaller standard errors as more data is available per person. The advantage of joint estimation is that it reduces inconsistencies in estimates across the different games, for example in key metrics such as income elasticities. However, recent academic work (Hess et al. 2017a) has highlighted the differences in valuations that can be obtained for the same component in different experimental games and has suggested that future studies should instead rely on a single game including all components. It is thus conceivable that larger standard errors are obtained when “imposing” this consistency across games for shared components. A simple test would be to estimate game-specific models using a specification that is consistent with the specification from the joint models. However, it should be noted that this exercise would likely lead to differences in key shared components, such as income elasticity, across experiments.
3. There is also a belief held by some that more complex models, such as used in the 2014/15 study, will lead to larger standard errors. However, it is not some universal truth that when more parameters are included for the same number of observations, the statistical significance of each individual effect goes down as there may for example exist negative correlations between the parameters. It is of course possible to test this assumption by simplifying the models, but this needs to then be weighed against the possibility of the estimates from simpler models being biased.

Implications of Option 3.1 for the updating programme

Option 3.1 has highlighted that confidence intervals associated with the base values and the size difference of these confidence intervals between the 2003 and 2014/15 study can be potentially be attributed to i) a number of methodological elements related to calculating the standard errors and ii) the differences in the functional form between the 2014/15 model. If the Department is interested in further decomposing the confidence intervals (i.e. standard errors) Phase two of this study can scope some of the mentioned options out. However, there are no direct implications for the updating programme specified in Streams 1 and 2.

Actions needed in Phase 2 of this study:

- No actions needed in Phase 2 of this study in respect of Option 3.1

Option 3.2: Steps to increase robustness without risk of bias to the estimates

Acknowledging that the 2014/15 confidence intervals are wider than expected, the question arises whether actions can be taken to reduce the width of the confidence intervals for the base year VTTS. The purpose of Option 3.2 is to identify possible steps potentially increasing the robustness of the results.

Two options arise in this context:

1. Further data collection, as proposed in Option 2.2, is without doubt the most reliable approach for improving robustness without the risk of introducing bias. It should be noted that the sample size of the 2014/15 study was small relative to the target population size and ambition of the study; and also when comparing the study with similar recent efforts, for example in Singapore (Hess et al. 2017b). Any estimation with additional data will of course need to account for the fact that the two samples will have been collected at two different points in time.
2. Reducing the number of segmentations will lead to increased sample size per model without increasing cost. This could lead to efficiency gains as the differences across segments potentially apply only to some of the components and could then be accommodated through within-model segmentations. Two subcomponents arise here:
    a. Purpose segmentations: this seems a viable approach in the case of commute and other non-work, but likely not in the case of EB where more significant differences in specification arise.
    b. Mode segmentations: the differences in valuations across the four modes are so substantial that combining the data across modes would not seem an advisable approach.

Since the primary goal of the maintenance framework is to move the 2014/15 framework forward over time, we **recommend** that Option 2.2, i.e. the development of

a regular monitoring framework, is the only feasible option to increase the robustness of the base year VTTS. Indeed, Option 2.1.1 suggested that there may be ongoing developments that suggest that the adopted level of segmentation may not remain appropriate over time. Such decisions should, however, be taken when the evidence emerging from Options 2.2 and 2.1.2 (meta-analysis) indicate that values from different mode-purpose segments are no longer significantly different from each other.

As mentioned in Option 2.2, by extending the evidence base on an annual basis through the proposed monitoring framework, the Department will not only be able to identify changes in the VTTS, but also re-base the VTTS regularly and reduce the width of the confidence intervals when such changes are not present.

Implications of Option 3.2 for the updating programme

In short, Option 3.2 provides additional arguments to further explore the potential of developing a regular monitoring framework collecting new behavioural data using a framework that is identical to the 2014/15 study. Besides improving our understanding of dynamics in the VTTS, it will also contribute to increasing the accuracy of the VTTS measures.

Actions needed in Phase 2 of this study:

- No actions needed in Phase 2 of this study in respect of Option 3.2

Option 3.3: Uncertainty in future VTTS values

Options 3.1 and 3.2 have not covered the fourth source of uncertainty as listed in the Table in Option 3.1, i.e. uncertainty in the GDP per capita income elasticity parameter. The latter type of uncertainty was central to the work of Wheat and Batley (2015). It stipulated that the uncertainty induced by applying the departments default VTTS adjustment approach over time (see Option 1.1.) grows exponentially over time and rapidly outweighs the estimation uncertainty associated with the behavioural model.

Figure 4 below is reproduced from Wheat and Batley (2015). It shows that for the purpose of appraisal VTTS, the uncertainty in the base VTTS is of minimal impact once 15-20 years have elapsed from the base year. This can be seen from the confidence intervals associated with a base VTTS which has three times the uncertainty (measured by width of confidence interval in the base year) in the base year (green lines) than the same base VTTS point estimate but with less uncertainty associated. Effectively after 15 years (2030 in the figure), the green and red boundaries are very similar. This is because the uncertainty introduced from uprating via an estimated GDP elasticity dominates. This is of importance since most transport schemes have benefits occurring at least 10-15 years from the base year. A key recommendation of Wheat and Batley (2015) was therefore to re-sample VTTS data on a regular basis, possibly with smaller sample sizes, to re-base the VTTS and thereby reset the confidence intervals.

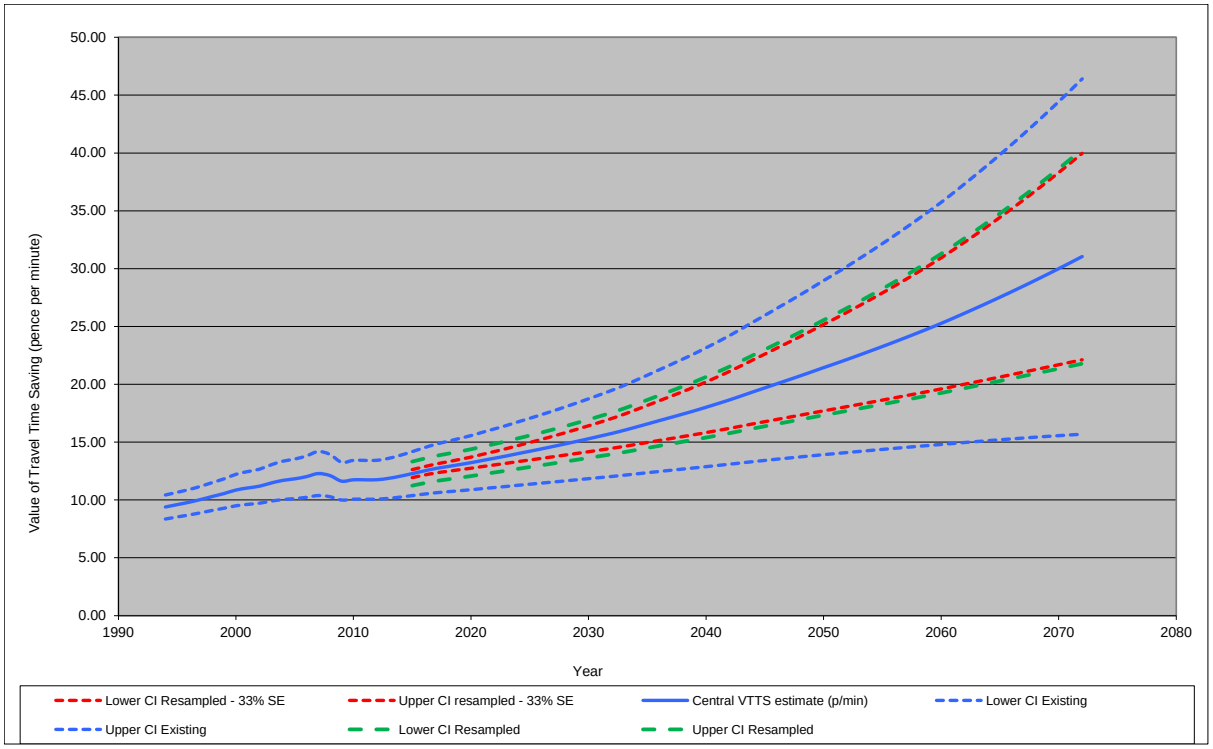


Figure 4: Reproduced from Wheat and Batley (2015): Figure 5: 95% confidence intervals for VTTS for the commuting trip purpose including 'what if' analysis scenarios (2010 prices)

On the one hand, this is good news since it implies that implications of the increased width on the 2014/15 confidence intervals surrounding the base year are only marginal 10-15 years from now when Option 1.1 is maintained. On the other hand, having high uncertainty in the base year does still translate into future years and should be avoided where possible. In essence, the conclusions from Wheat and Batley (2015) have not changed despite the change in the analytical framework to arrive at the base values.

The **recommendation** following from the above is therefore to explore whether the uncertainty introduced by the inter-temporal income elasticity parameter can be avoided (or at least minimised). Option 1.2 provides a framework to do so if it can be argued that the cross-sectional effects picked up by implementing new NTS data capture all relevant changes in the VTTS thereby avoiding or limiting the role of the GDP per capita escalation factor. In combination with the new data collection put forward in Option 2.2, there is the possibility that such re-basing can even reduce the base confidence intervals over time as put forward in Option 3.2. Nevertheless, since Option 1.2 still requires uprating of future VTTS values through GDP growth, the funnel-shaped uncertainties will persist.

Implications of Option 3.3 for the updating programme

Option 3.3 does not introduce new dimensions to the updating programme, but uses Wheat and Batley (2015) to highlight that there are good reasons to further explore Options 1.2 and 2.2 in order to limit the degree of uncertainty in future VTTS values. identical to the 2014/15 study.

Actions needed in Phase 2 of this study:

- No actions needed in Phase 2 of this study in respect of Option 3.3

Summary - Stream 3

Stream 3 takes a different approach to (inter-temporal changes in) the VTTS and provides guidance on the degree of accuracy associated with base and future VTTS values resulting from the 2014/15 behavioural model and the various options discussed in Streams 1 and 2. A key part of providing guidance on confidence intervals is obtaining a better understanding on how modelling techniques and model specification affect standard errors. Option 3.1 provided a set of possible tests that could be conducted on the existing behavioural data. Option 3.2 then discussed how confidence intervals of the base VTTS can be refined. Option 3.3 continued this discussion by highlighting the importance of re-basing for limiting the width of the confidence intervals on future VTTS values. The latter two options provided key arguments to further scope the work put forward in Options 1.2 and 2.2.

7. Stream 4: Commissioning of studies on outstanding and emerging topics

Finally, the Department may wish to revisit specific topics or mode-purpose combinations to resolve outstanding or contentious issues from the 2014/2015 study. Similarly, the Department may wish to investigate and validate emerging topics as indicated in Stream 2. This is an essential step to challenge the existing framework and validate new developments. Although closely related to the updating programme, these studies are expected to be undertaken on an ad-hoc basis. Some potential topics are listed below. It was not part of the scope of Phase I to provide recommendations on the feasibility and scoping of these options.

Option	Label	Description
4.1	Business value of time	- Revisit the work while travelling element (unsatisfactory results) - Revisit the decision to use employee vs. employers' values (challenge the existing framework)
4.2	Walking and cycling	- Are suitable data collection and analysis methods emerging to estimate these values (uncovered areas)?
4.3	deltaT	- The current appraisal VTTS are based on time savings of ten minutes. The sensitivity and implications of the VTTS to the size of this 'deltaT' need to be better understood.
4.4	Temporal change	- Understanding the observed change in the commute vs. other non-work values vs. 2003.
4.5	Reliability	- Understanding the observed change in values of reliability for car vs. 2003.
4.6	Distance effects	- Understanding the strength of the distance effect in the overall values
4.7	Boarding, frequency and interchange penalties in rail	- Unbundling the VTTS for rail passengers in journey and service quality characteristics.
4.8	Other PT crowding	- The 2014/15 Other PT crowding multipliers were considered conservative relative to UK guidance.
4.9	Incremental approaches	- Testing of new SP models and alternative segmentations (also in light of robustness criteria in Stream 3)

8. Summary

The intention of the Department to develop a maintenance programme for the VTTS which goes beyond the default inter-temporal adjustment for GDP per capita growth (as described in this report by Option 1.1) can be classified as ambitious. It signifies that the Department is willing to take (and maintain) a leading role in the development of national VTTS studies. Namely, also in other western European countries Option 1.1 represents the default option to adjust the VTTS over time.

The 2014/15 VTTS study provided the Department with a set of revised set of VTTS values for a variety of mode-purpose combinations and improved our understanding how the VTTS varies with socio-economic and journey characteristics. However, little is currently understood about how the VTTS changes over time. We see the proposed maintenance programme as a prominent opportunity to increase the general understanding of inter-temporal changes in the VTTS in the UK.

This report has described how such a maintenance programme could be formulated, and provided a long list of solutions that can be taken forward to further develop the maintenance programme. The long list of options was structured along four main streams which are summarised in more detail below.

In developing the structure of the maintenance programme, we started in Section 2 by providing an appreciation of the latest 2014/15 VTTS study and how it relates to its predecessor. Section 2 highlighted the methodological contributions that were made in the 2014/15 study, but also that the Mackie et al. (2003) study already made use of a combination of a behavioural framework and NTS data to arrive at a set of appraisal VTTS values. In developing the maintenance programme, we could therefore build on past experience. Section 3 put forward a number of causes for inter-temporal changes in the VTTS, which until the present day are all accommodated for by the GDP per capita growth factor. After having set the context in Sections 2 and 3, the structure of the maintenance program was then presented in Sections 4-7.

The focus of **Stream 1** (Section 4) is on adjusting the VTTS over time whilst maintaining the 2014/15 behavioural framework. Three options are available to the Department of which the two already existing options were considered feasible. Option 1.1 refers to the use of GDP growth escalation factors as currently used by the Department. Option 1.2 refers to the framework described by Bates (2008) and as implemented by the Department in 2013 (Laird et al. 2013). Option 1.2 feeds new NTS data into the behavioural framework and thereby arrives at a new base VTTS. It should be noted that Options 1.1 and 1.2 may potentially be combined, as long as double-counting for the sources of change in the VTTS is avoided. Stream 1 highlights that additional research is, however, recommended on which sources of inter-temporal change in the VTTS are captured by either option. With respect to future VTTS values, e.g. 60 years from now, it is recommended that the Department continues its existing practice and only accounts for forecasted GDP growth and not for expected changes in socio-economic and journey characteristics as captured by the NTS data. Namely,

Option 1.3 was considered, in agreement with the Department, as unfeasible due to high degrees of uncertainty.

Stream 2 (Section 5) is needed to judge whether the 2014/15 framework remains fit-for-purpose. Stream 2 is split into two parts. Firstly, Options 2.1.1, 2.1.3 and 2.1.4, are classified as horizon scanning activities. Here, the Department is recommended to track, by making use of the social sciences, how socio-economic and technological developments may impact the VTTS in the near future and thus affect the 2014/15 behavioural framework (Option 2.1.1). Option 2.1.3 recommends the Department to keep track of new data collection efforts, primarily Revealed Preference, which may surpass the conventional use of Stated Choice surveys in collecting new behavioural data for VTTS estimation. Currently, Revealed Preference methods are judged not to be developed enough to replace Stated Choice surveys, but promising developments are taking place. Similarly, Option 2.1.4 recommends the Department to keep track of new SC data collection and analyses techniques. The latter two options are best to go hand in hand and are aimed at ensuring that if and when a new large-scale study is necessary that it will, again, be commissioned according to the state-of-the-art.

Secondly, Option 2.1.2 and Option 2.2 are aimed at empirically validating the outcomes of Stream 1. In doing so, Option 2.1.2 relies on emerging empirical evidence being incorporated in the well-developed meta-analytical frameworks on the VTTS in the UK and Europe. The meta-analytical framework gives an independent source of information against which the outcomes of Stream 1 can be statistically tested. For this framework to remain relevant, regular updates of the meta-model(s) need to be made when sufficient new evidence has become available. Option 2.2 allows for a more direct empirical testing of the temporal stability of the 2014/15 behavioural framework. It recommends that the Department should collect on an annual basis new Stated Choice data using the identical methodology and models that were used in the 2014/15 study. As such, pure temporal variation in the VTTS can be picked up. Since only sample sizes of 10-20% of the original data are recommended to be collected, it is recommended that the data are jointly analysed with the original dataset whilst controlling for shifts in key VTTS parameters. A benefit of this approach is that, as long as the behavioural framework is judged to be appropriated, the updated models are expected to be applicable alongside Option 1.2, possible resulting in narrower confidence intervals.

Stream 3 is aimed at increasing the Department's understanding of the robustness of the base year and future VTTS values. Option 3.1 respectively provides a number of suggestions to further decompose the differences in the width of the confidence intervals between the 2014/15 and Mackie et al. (2003) appraisal values. Option 3.2 then identifies possible steps potentially increasing the robustness of the results and provides additional support for further developing Option 2.2. In similar fashion, Option 3.3 evaluates the robustness of future VTTS values. Building on Wheat and Batley (2015), Option 3.3 highlights the importance of frequently re-basing the base year

VTTS to limit the width of the confidence intervals on future VTTS values. Again, this provides support to further scope Options 1.2 and 2.2.

Finally, **Stream 4** of the maintenance programme allows the Department to commission research on outstanding and emerging topics. A couple of potential topics were listed such as the use of employers' or employee's for the Business Values of Time. A topic that was deemed out of scope for Stream 4 is *congested values*. Although this is an interesting topic to be included in Stream 4, it is already part of the 'Congested Values of Time in Transport modelling' (DfT Reference: 04/102/33) study. The goal is that these studies in Stream 4 can be used to shape the approach for the next large-scale VTTS study. The proposed maintenance programme should provide the Department with the tools to judge when the next large-scale VTTS study is necessary.

9. Recommendations for narrowing down the long-list

The presented structure of the programme for maintaining a robust valuation of travel time savings in the UK together with the options described in this report form the long-list of alternatives for maintaining the robustness of VTTS over time. The purpose of this section is to reflect on how this long-list was developed and to provide recommendations on which elements should be taken forward in Phase 2 of the research project for more detailed analysis and scoping.

The structure of the proposed programme largely emerged from reviewing current practice of the Department towards updating the VTTS over time through WebTAG. This led to the creation of Stream 1, where Options 1.1 and 1.2 reflect the Department's current default approach and an adjustment procedure which has been applied once before respectively (see Laird et al. 2013). Moreover, Option 1.2 also echoes through what the Department defined as the 'new baseline' in the original call for research. Where Option 1.2 only covers adjustment of base year VTTS values, but not future values, we looked into extending the methodology to future years in Option 1.3. The degree of uncertainty associated with the latter option was judged too large that we recommend not to take it forward in Phase 2.

Recommended actions from Stream 1 to take forward in Phase 2:

- A priority identified by Phase 1 of the research project is that the Department's default approach of uprating the VTTS using a unit GDP per capita growth elasticity (Option 1.1) should be revisited both from a theoretical and empirical perspective.
- Closely related to the above, we recommend the Department to study how Options 1.1 and 1.2 are linked and whether additional factors to income can be applied to adjust the base year VTTS over time.
- These two recommendations can be taken forward in Phase 2 through academic consultation.

Stream 2 was then designed to challenge the key assumption underlying Stream 1, where it was assumed that the behavioural model remains fit-for-purpose. Challenging this assumption, and being able to identify those changes, requires set of horizon scanning activities (Option 2.1) and an actual monitoring programme (Option 2.2.). Stream 2 covers a wide range of activities of which we consider the following two the most urgent and therefore recommend taking these forward in Phase 2:

Recommended actions from Stream 2 to take forward in Phase 2:

- Scoping of the updating and re-estimation programme of the UK and European meta-analyses for validating the outcomes of Stream 1 (Option 2.1.2).
- Scope out the monitoring programme presented in Option 2.2 and conduct Delphi surveys with market research companies on the possibilities and feasibility of running such a high frequency monitoring programme.

Stream 3 was put into place to consider the robustness of the base and updated VTTS values arising from Streams 1 and 2, particularly Options 1.2 and 2.2. Robustness was specified as one of the key evaluation factors in the research call of this project. Stream 3 provides further support for the above recommendations to further develop Options

1.2 and 2.2 in Phase 2. One of the key arguments for this is given by the Wheat and Batley (2015) work stressing the importance of frequent rebasing of the VTTS. The robustness implications should therefore be revisited once the properties of the referred options have been worked out in more detail in Phase 2.

Stream 4 was deemed to be necessary to give the Department the opportunity to revisit specific outstanding or contentious issues from the 2014/15 study. Since it was not part of the scope of Phase I to provide recommendations on the feasibility and scoping of these options, no actions have to be taken forward in Phase 2.

10. List of references

Abrantes, P.A.L. and Wardman, M.R. (2011) Meta-analysis of UK values of time: an update, Transportation Research A, 45 (1), 1-17.

Accent and Hague Consulting Group (AHCG): The value of travel time on UK roads. Report to Department for Transport (1999)

Arup, ITS Leeds, Accent: Provision of market research for value of time savings and reliability. Phase 2 report to the Department for Transport (2015). https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/470231/vtts-phase-2-report-issue-august-2015.pdf

Bastian A. (2017) Explaining Trends in Car Use, PhD thesis KTH Stockholm. http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1141730&dswid=5059

Bates J. (2008) Value of Time Calculations. *Report to Department for Transport*. Report dated 23th December 2008.

Batley R; Bates J; Bliemer M; Börjesson M; Bourdon J; Cabral MO; Chintakayala PK; 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 (2017) New appraisal values of travel time saving and reliability in Great Britain, *Transportation*, 1-39.

Beesley ME (1965) The value of time spent travelling: some new evidence. *Economica* 32:174–185

Börjesson, M., Fosgerau, M. and Algers, S. (2012) On the income elasticity of the value of travel time, Transportation Research A 46, 368-377.

Borjesson, M, Fosferau, M. and Algers S. (2012a) Catching the tail: empirical identification of the distribution of travel time. *Transportation Research Part A*, 46(2), 378-391.

Calastri, C., Hess, S., Choudhury, C., Daly, A. and Gabrielli, L. (2017) Developing mode choice models using semi-ubiquitous data sources: a case study using a smartphone based GPS survey from Italy, accepted for special issue of *Transportation Research Part B*.

Cyganski, R., Fraedrich, E. and Lenz, B. (2015) Travel-time valuation for automated driving: A use-case-driven study, Paper presented at the 2015 T*ransportation Research Board Annual Meeting*.

Daly, A., Tsang, F. and Rohr, C. (2014) The value of small time savings for non-business travel, in *Transport Economics and Policy*, Vol. 48, 2, May.

Daziano, R.A., Sarriasb, M. and Leard, B. (2017) Are consumers willing to pay to let cars drive for them? Analyzing response to autonomous vehicles, *Transportation Research Part C: Emerging Technologies*, 78, 150-164

De Borger, B., Fosgerau, M.: The trade-off between money and travel time: a test of the theory of reference-dependent preferences. *J. Urban Econ*. 64(1), 101–115 (2008).

DfT (2017) Evidence to the Commission on Travel Demand, http://www.demand.ac.uk/wp-content/uploads/2017/03/09-EC1-Department-for-Transport-01.pdf

European Commission, Directorate General for Regional and Urban Policy (2014) Guide to cost-benefit analysis of investment projects, Economic appraisal tool for Cohesion Policy 2014-2020, EC, Brussels.

Fox,E., Hitching, R., Day, R. and Venn, S. (2017) Demanding distances in later life leisure travel, Geoforum, 82, 102-111

Fosgerau, M. (2005) Unit income elasticity of the value of travel time savings, Paper presented at the European Transport Conference 2005.

Gunn, H.F. (2001) Spatial and temporal transferability of relationships between travel demand, trip cost and travel time, Transportation Research E, 37(2/3), 163-159.

HEATCO (2006) Bickel, P, A. Hunt, G. de Jong, J. Laird, C. Lieb, G. Lindberg, P. Mackie, S. Navrud, T. Odgaard, J. Shires and L. Tavasszy - HEATCO Deliverable 5: Proposal for harmonised guidelines, IER, Stuttgart.

Hess, S., Daly, A., Dekker, T., Ojeda Cabral, M., Batley, R.: A framework for capturing heterogeneity, heteroskedasticity, non-linearity, reference dependence and design artefacts in value of time research. *Transp. Res. Part B* 96, 126–149 (2017).

Hess, S., Börjesson, M. & Daly, A.J. (2017a), Critical Appraisal of Use of Simple Time-Money Trade-Offs for Appraisal Value of Time Measures, *96th Annual Meeting of the Transportation Research Board*, Washington, D.C.

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*, accepted for publication, February 2017.

Hess, S., Quddus, M.A., Rieser-Schüssler, N. & Daly, A.J. (2015), Developing Advanced Route Choice Models for Heavy Goods Vehicles Using GPS Data, T*ransportation Research E*, 77, 29-44.

Kouwenhoven, M. de Jong, G., Koster, P., van den Berg, V., Verhoef, E., Bates, J. and Warfemius, P. (2014) New values of time and reliability in passenger transport in The Netherlands, *Research in Transportation Economics*, 47, 37-49.

Kouwenhoven, M., Börjesson, M., Daly, A. and Jong, G.C. de (2017) Changes in the value of travel time over the years, Paper presented at the ICMC, Cape Town.

Laird, J.J., J. Bates and P.J. Mackie (2013) Peer Review of Proposals for Updated Values of Travel Time Savings. *Report to Department for Transport*. Report dated 20th October 2013.

Lorenzo, J.M., Daly, A. and Börjesson M. (2017) Measuring errors with latent variables in transport models. *hEART Conference, Tel-Aviv*.

Mackie, P.J., Jara-Dıaz, S. and Fowkes, A.S., 2001. The value of travel time savings in evaluation. *Transportation Research Part E: Logistics and Transportation Review*, 37(2), 91-106.

Mackie, P.J., Wardman, M.R., Fowkes, A.S., Whelan, G.A., Nellthorp, J., Bates, J.J.: Value of travel time savings in the UK. Report to Department for Transport (2003). http://eprints.whiterose.ac.uk/2079/2/Value_of_travel_time_savings_in_the_UK_protected.pdf

Marsden, G., Wardman, M.A., Shires, J.D. and Strauss, A. (2014) 'Integrated Information for Integrated Transport (IIfIT)', Report to Transport Systems Catapult

McDonald, M. (2017) 'Evidence to the Commission on Travel Demand', http://www.demand.ac.uk/wp-content/uploads/2017/03/26-EC1-McDonald.pdf

NIC (2017) 'The Impact of Population Change and Demography on Infrastructure Demand', National Infrastructure Commission, London.

ONS (2016) National Population Projections: 2016-based statistical bulletin. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/nationalpopulationprojections/2016basedstatisticalbulletin

RAND Europe and CE Delft (2004) Value of time and value of safety guidelines for transport projects, final report, Report prepared for the EIB, TR-227-EIB, RAND Europe, Leiden.

Rohr C., Ecola L., Zmud, J., Dunkerley, F., Black, J. and Baker E. (2016) Travel in Britain in 2035: Future scenarios and their implications for technology innovation. https://www.rand.org/content/dam/rand/pubs/research_reports/RR1300/RR1377/RAND_RR1377.pdf

Shires, J.D. and Jong, G.C. de (2009) An international meta-analysis of values of travel time savings, Evaluation and Program Planning 32(4), pp.315-325.

Significance, VU University, John Bates Services, TNO, NEA, TNS NIPO and PanelClix (2013) Values of time and reliability in passenger and freight transport in The Netherlands, Report for the Ministry of Infrastructure and the Environment, Significance, The Hague.

Tapley, N, Wardman, M.R., Gunn, H.F. and Hyman, G. (2007) Inter-temporal variations in values of time in Great-Britain, Paper presented at the European Transport Conference.

TSC (2016) 'Mobility as a Service: Exploring the opportunity for Mobility as a Service in the UK', Transport Systems Catapult, Milton Keynes.

Wardman M. (2001) Inter-temporal variations in the value of time. Working paper, Institute for Transport Studies, University of Leeds, UK.

Wardman, M. (2004) Public Transport Values of Time. *Transport Policy* 11, pp.363-377.

Wardman, M., Batley, R., Laird, J., Mackie, P. and Bates, J., 2015. How should business travel time savings be valued? *Economics of transportation*, 4(4), pp.200-214.

Wardman, M.R., Chintakayala, P., Jong, G.C. de and Ferrer, D. (2012) European wide meta-analysis of values of travel time, Paper prepared for EIB, ITS Leeds.

Wardman, M.R., Chintakayala, P., Jong, G.C. de (2016) Values of travel time in Europe: review and meta-analysis, *Transportation Research A*, 94, 93-111.

Weis, C., Kowald, M., Axhausen, K., Balać, M. and Mathys, N. (2017). The Swiss national stated preference study on transport 2015. Paper presented at the European Conference on Mobility Management, Maastricht, Netherlands, May 31 - June 2.

Wheat, P., Batley, R. (2015) Quantifying and decomposing the uncertainty in appraisal values of travel time savings. *Transp. Policy* 44, 134–142.

Zhou, Y., & Wang, X. (2014). Explore the relationship between online shopping and shopping trips: An analysis with the 2009 NHTS data. Transportation Research Part A: Policy and Practice, 70, 1-9

List of abbreviations

AHCG	Accent Marketing and Research and Hague Consulting Group
AV	Automated Vehicle
CBA	Cost Benefit Analysis
CSA	Cost Savings Approach
DfT	Department for Transport
EBRD	European Bank for Reconstruction and Development
EIB	European Investment Bank
EU	European Union
ITS	Institute for Transport Studies
GDP	Gross Domestic Product
MaaS	Mobility as a Service
MXR	Market Exchange Rate
NTS	National Travel Survey
OBR	Office for Budget Responsibility
PPP	Purchasing Power Parity
RP	Revealed Preference
SC	Stated Choice
SP	Stated Preference
TSC	Transport Systems Catapult
VTTS	Value of Travel Time Savings
WebTAG	Transport Analysis Guidance