Institute for Transport Studies
FACULTY OF ENVIRONMENT

UNIVERSITY OF LEEDS

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

Final Phase 2 Report

Arup

Institute for Transport Studies, University of Leeds

March 2018

ARUP 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

Institute for Transport Studies – Document control information

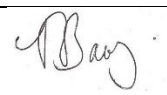
Title	Programme for maintaining a robust valuation of travel time savings: feasibility study.
Authors	Institute for Transport Studies (ITS), University of Leeds and Arup
Editor	Thijs Dekker & Margot Finley
Internal reviewers	Richard Batley
Arup QA	Matthew Dillon
Version number	1.0
Date	26th of March 2018
Distribution	ITS & Arup team
Availability	ITS & Arup team
File	Arup_ITS_Robust_VTTS_Phase2_Revised_final_260318.docx
Signature	

Table of Contents

1. Introduction

This document is part of the larger project '*Programme for maintaining a robust valuation of travel time savings'* as commissioned by the Department for Transport. In Phase 1 of the project, Arup and ITS Leeds provided a possible structure for such a programme including a long-list of options of recommended activities. At the end of Phase 1, it was decided that a limited number of options would be worked out in more detail in Phase 2 of the project. The aim of this document is to give the Department for Transport recommendations on which options to take forward, and information on the recommended activities and corresponding costings that would feed into the eventual invitations to tender (ITT).

The document provides further scoping work on Options 1.1; 1.2; 2.1.2; 2.1.3 and 2.1.4 of the Phase 1 report. The first two of the aforementioned options describe regular procedures to adjust the value of travel time savings (VTTS) and related valuations of reliability etc. over time without collecting new survey data in the proposed programme. Option 2.1.2 describes a potential role for meta-analysis in the VTTS updating programme. Options 2.1.3 and 2.1.4 shed further light on the development of revealed and stated preference data collection methods respectively.

The document presents:

- Section 2: Options 1.1 and 1.2
- Section 3: Option 2.1.2
- Section 4: Option 2.1.3
- Section 5: Option 2.1.4
- Section 6: Summary and recommendations

In all sections, the necessary background information is presented alongside a description of the required work and an associated budget.

2. Options 1.1 and 1.2 – uprating the VTTS over time

2.1 Background

The core assumption underlying Options 1.1 and 1.2, as described in the Phase 1 report, is that the 2014/2015 VTTS behavioural models remain fit-for-purpose.[1] As such, the key question is how the 2014/15 VTTS values should be adjusted over time.

Option 1.1 reflects the Department's default option and at present adjusts the base VTTS values for growth in GDP per capita over time using a unit income elasticity. Any changes to the default long run GDP per capita growth assumptions, such as recommended by the Office for Budget Responsibility (OBR) and the Green Book, should be adopted in transport sector appraisal practice, including the VTTS forecasts.

[1] A detailed description of Option 1.1 and 1.2 can be found in the Phase 1 report.

As such, future work on Option 1.1 is limited to determining the extent to which the assumed unit income elasticity remains the best assumption. As put forward in the Phase 1 report, the use of Option 1.1 is the current state of practice in most European countries where the implemented income elasticities range between 0.5 and 1.

Option 1.2 is aimed at enriching the Department's default approach to uprating the VTTS. Option 1.2 aims to combine the adjustment for GDP per capita income growth over time with changes in the travelling population and changes in the characteristics of the trips they are making. In practice this would lead to regularly feeding new NTS data into the behavioural model (i.e. the VTTS Implementation Tool), thereby adjusting the base year VTTS, possibly in combination with an additional income growth adjustment. The adjusted base year VTTS would then be projected for future years using a standard income elasticity parameter.

2.2 **Requirements**

2.2.1 Objectives

The objectives of the activities below are twofold.

First, there is a need to determine the most appropriate inter-temporal income elasticity parameter. In this process, it should be considered whether the income elasticity parameter should (or should not) vary across journey purposes.

Second, there is a need to determine how the Department's default approach of controlling for GDP per capita income growth can be combined with changes in travel patterns as captured by the NTS.

*2.2.2 **Work required***

The first objective is clearly connected with Option 1.1 and the work required can be split up into several separate activities. First, the theoretical connection between the VTTS and growth in GDP per capita should be established.[2] This is informed by the academic literature. Fosgerau (2005) and Jara-Diaz and Rosales-Salas (2017) form good starting points for this exercise. A key question is whether or not a unit elasticity relationship between the wage rate and the VTTS is theoretically justified. The goal of this review is the development of a recommended link between GDP per capita and other non-work journey purposes. The outcomes from this review should guide the empirical work described below.

[2] As referred to in the Phase 1 report on p.13, growth in GDP per capita is typically assumed to represent growth in the wage rate, which is directly related to the real value of work time and suggests a unit income elasticity of the VTTS with respect to income when adopting the cost-savings approach. When adopting variations of the cost-savings approach, through the Hensher equation, or WTP values for business trips the notion of a unit elasticity is no longer supported by micro-economic theory and largely becomes an empirical matter. Nevertheless, a connection between income and the VTTS persists, also for non-work purposes. By adopting a WTP approach for all journey purposes in WEBTAG after the 2014/15 study, the marginal utility of income allows linking growth in the VTTS to growth in income.

Having established the theoretical connection between income growth and the VTTS, the next activity required is reviewing the empirical approaches and evidence identifying the inter-temporal income elasticity parameter. Identifying the inter-temporal income elasticity parameter requires observing variation in the VTTS over time. Potential data sources are VTTS studies in one location collected at different moments in time. Börjesson et al. (2012), for example, use data from the 1994 and 2007 Swedish value of travel time savings studies. The difficulty here is that income growth is concurrent with other sources of change. Moreover, the comparability of these datasets over time is limited as there are many potential sources of noise present, making it challenging to pick up a pure income elasticity. This is particularly the case when applying meta-analysis for this purpose due to the use of a large set of heterogeneous source studies.[3] Specific attention should also be paid to how the studies treat income (gross or disposable) since not accounting for the progressive nature of income taxes can affect the income elasticity parameter. It would be beneficial for the work to draw on other evidence from the literature on the inter-temporal income elasticities of other valuations (not just time valuations) and to review the empirical relationship between the inter-temporal and cross-sectional income elasticity. Daly and Fox (2012) and Swärdh (2008) provide good references in relation to the existing empirical evidence on the inter-temporal income elasticity.

Together, the above two tasks should translate to a recommended empirical approach to identify the best possible inter-temporal income elasticity for the different journey purposes. This should include recommendations for the data and estimation approach to be used. An indicative costing of the data collection and estimation work is included in Section 2.3.1.

While the first objective is mainly concerned with adjusting the base VTTS over time in line with income growth, the second objective is to adjust the base VTTS for changes in the travelling population and their travel characteristics. In principle, feeding new NTS data into the behavioural model on an annual basis, or once every three years, would be a relatively straightforward undertaking, although a limited number of data preparation modifications, as highlighted on page 19 of the Phase 1 report, would need to be conducted before the new data can be fed into the tool that would estimate the resulting VTTS (i.e. the VTTS Implementation Tool). Additionally, assumptions would need to be made on how non-NTS control variables change over time. 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.

The most important exercise under objective two (Option 1.2) is a repeat of Bates (2008). In the context of the Mackie et al. (2003) behavioural model (see Section 2.1 of the Phase 1 report for more details), Bates (2008) relied on new NTS data to adjust the number of trips attributed to the relevant income and distance bands. As such, this

[3] The meta-analysis approach is currently used to determine the inter-temporal income elasticity for the Department.

adjusted the trip and distance weights in the enumeration of nationally representative appraisal values.[4] Minor modifications, e.g. correcting for inflation, were implemented to the VTTS in a specific income-distance band. The updated appraisal values were then additionally uprated using the standard GDP escalation factor. Bates (2008) and also Laird et al. (2013) found that changes in the NTS over time play only a small role, although this conclusion may have changed by having adopted the new 2014/15 behavioural model.

A replication of Bates (2008) can either take form of a back-casting exercise relying on past NTS data, or a forecasting exercise relying on more recent NTS data.[5] The NTS data for the target base year would then be combined with the (backward or forward) GDP escalation factor to derive the combined change in the VTTS. The relative contribution of the adjustment in NTS data and the escalation factor can be determined by making both adjustments in isolation.

A limitation of repeating the Bates (2008) study is that there is no validation of the resulting VTTS measures. Validation requires having access to a VTTS estimates in the target year of the back- or forecast, preferably obtained from a comparable behavioural model based on comparable source (i.e. stated preference) data. Currently, the best available evidence for such a validation exercise is the survey responses to the 1994 national VTTS survey, in combination with the relevant NTS data (e.g. 1993-1995). This is a challenging exercise since the national VTTS survey has changed significantly over time and there are many sources of potential change in the VTTS going beyond the changes in the travelling population.[6] Options for validation would significantly improve if Option 2.2 of the Phase 1 report, i.e. a rolling survey collecting new survey data on a regular basis, were implemented.

Based on such a validation exercise, one could develop recommendations as to whether the inter-temporal income elasticity derived under Option 1.1 should be combined with new NTS data; or whether either should be implemented in isolation. Combined with Option 2.2 more options become available to evaluate whether such adjustment exercises are appropriate at all; or whether a new large data collection exercise might be appropriate after several years.

A key issue that requires further scrutiny in the context of Bates (2008) is double counting. The Phase 1 report provided more detail on this issue on pages 6-9 and 17, and we summarise the key issues here. First, it has been acknowledged that the empirical estimate of the GDP per capita escalation factor does not represent a clean income effect, but inevitably captures other changes in the VTTS, including those in

[4] Repeating the Bates (2008) exercise in the context of the 2014/15 VTTS Implementation Tool is slightly more involved since the Implementation Tool relies on individual trips (and characteristics of the traveller) instead of only adjusting the number of trips in each income and distance band.

[5] The 2014/15 VTTS Implementation Tool model makes use of 2010-2012 NTS data.

[6] As a case in point, Wardman (2001) refers to earlier work by Gunn making inter-temporal comparisons of the VTTS in the Netherlands and the UK.

the travelling population and their trip characteristics. New NTS data would represent such effects and hence introduce the risk of double counting. Second, income is a key variable in the 2014/15 behavioural model acknowledging that the VTTS varies across cross-sectional income categories. Increases in income would thus directly increase the VTTS in the Implementation Tool. If additionally, we uprate the resulting appraisal VTTS for GDP per capita growth there is a risk of double counting. There is the possibility to counteract the second source of double counting by adjusting the denominators associated with the cross-sectional income, cost, time and distance elasticities in the Implementation Tool over time (see equation (2) in the Phase 1 report).[7] This would then only capture changes in the distribution of cross-sectional income in the travelling population over time. Unfortunately, there is no theoretical guidance on this issue and therefore we recommend conducting a sensitivity analysis using both the original and adjusted denominators to study the influence on the VTTS as part of the above validation exercise.

2.3 Timescales and budget

2.2.1 Identifying the inter-temporal income elasticity

In relation to the first objective, we have identified three strands of work.

1. Theoretical review of the connection between income growth and the VTTS.
2. Review of the empirical evidence on the inter-temporal income elasticity parameters.
3. Recommending a preferred approach for identifying the inter-temporal income elasticity for the UK VTTS for different journey purposes.

In costing the above three steps, and in Section 2.3.2, we assign a number of working days to each of the activities which is subsequently translated into pounds using a blended consultancy day rate of £1,000 (2018 prices, excluding VAT).

Based on our experience, for the theoretical review we envisage that around 10 person-days will be required and an additional 2 person-days for reporting will be needed. These 12 person-days thus translate into £12K.

The amount of work for the empirical review is somewhat higher as a more diverse literature needs to be reviewed on a variety of elements. In total around 18 person-days will be needed, including around 3 person-days for reporting; adding up to £18K.

The development of a recommended approach for the UK VTTS would require a review of potential datasets matching up with the outcomes of the first two steps. A total of 5 person-days would be needed for this, giving a total of £5K.

In all, activities 1-3 add up to £35K. This is an approximate cost for developing the framework and empirical approach, based on our professional experience. Conducting the additional analyses and preparing the datasets to estimate the inter-temporal

[7] The interest is not so much in the value of the cross-sectional elasticities, but in the associated denominators used in the VTTS Implementation Tool. In estimation the denominator is irrelevant, it is not in implementation.

income elasticity would cost an additional £20K. Note that all costs in this section exclude VAT.

2.2.2 Adjusting the baseline VTTS with NTS data

In relation to the second objective, we can again identify three strands of work:

1. Regular adjustment of the base year VTTS using new NTS data
2. Repeating the Bates (2008) exercise testing the importance of NTS changes.
3. Sensitivity tests for double counting

The 2014/15 VTTS values are based on three years of NTS data. We suggest that new NTS should be fed into the VTTS Implementation Tool every three years. The data preparation involves preparing the data, adjusting prices for inflation and changes in fuel costs. We expect that this preparation would take about six person-days and an additional two person-days would be required to feed the data into the Implementation Tool. In total, the first strand of work would cost around £8K every three years (2018 prices, excluding VAT, and using the same day rate as previously). If this task were automated then this would save on some of the costs, albeit at the expense of additional start-up costs to set up the automation. This costing includes making assumptions about adjusting non-NTS data and finding alternatives when changes to the NTS data collection are made.

The effort required to repeat the Bates (2008) exercise is a little more involved as it requires sourcing old (or new) NTS data to update the appraisal VTTS to a new base year, including the escalation procedure accounting for growth in GDP per capita. Validating these updated values would additionally require estimating the 2014/15 behavioural model on old (or new) UK VTTS data and subsequently adjusting the Implementation Tool based on the new behavioural parameters to derive a set of comparable appraisal values for validation purposes. For validation based on the 1994 survey and sourcing the corresponding NTS data, this would account for approximately twenty person-days of work, including estimation, implementation and reporting. This would add up to a total of around £20K (£14K SP estimation and £6K in preparing NTS data, 2018 prices, excluding VAT, and using the same day rate as previously). Note that as part of a potential future rolling survey, the estimation costs would be much lower since the survey is designed to match the current model specification. The latter approach would be our recommended approach.

If the Department would simply wish to compare the relative magnitude of Option 1.1 and Option 1.2 on the change in the VTTS, i.e. a direct repeat of Bates (2008) without validation, only NTS data would need to be sourced and implemented into the Implementation Tool which would reduce the cost slightly to approximately £7K (£1K for preparing the Implementation Tool and conducting the comparisons and £6K for NTS data collection, again at 2018 prices, excluding VAT, and using the same day rate as previously). This, however, lacks the validation of the full exercise. For the reported sensitivity tests in the third strand, one should expect additional costs of approximately one day at £1K.

3 Option 2.1.2 – meta-analysis

3.1 Background

Over the past 20 years, there have been a number of meta-analyses of values of time in the UK.[8] These studies have all been published in leading international journals.

The first meta-analysis of UK values of (in-vehicle) time was reported by Wardman (1998) as a result of a discussion with the Department for Transport as to whether the second national value of time study (AHCG et al., 1999) included a review of the literature, given that there had been numerous studies reporting value of time estimates since the first national study of the early 1980s.

The second study (Wardman, 2001) was undertaken primarily to extend coverage to a wide range of values other than in-vehicle time. It covered values of walking time, waiting time, access time headway, congested travel time, departure time shifts, parking search time and late time.

The third study (Wardman, 2004) was undertaken as part of the Mackie et al. (2003) study that formed the basis of revised DfT value of time recommendations making use of the AHCG et al. (1999) data. The focus of the meta-analysis was on public transport values of time and how these related to car time valuations, since the AHCG et al. (1999) study only covered car. The other purpose was to extend the years covered to support the estimation of a more robust GDP elasticity.

The fourth study had the ambitious aim of updating the value of time meta-analysis and also updating a price elasticity reported by Wardman and Shires (2003) and extending the meta-analysis approach to time elasticities. This would then allow the consistency of time valuations, price elasticities and time elasticities to be explored. In the event, the meta-elasticity studies were completed and published but the consistency aspect was not undertaken due to limited availability of resources. The value of time meta-analysis was reported in Abrantes and Wardman (2011).

The most recent study, conducted for the European Investment Bank, extended coverage to the whole of Europe and also covered Schedule Delay and Reliability Ratio measures of reliability in addition to Late Time (Wardman et al., 2016). The meta-analyses studies discussed above summarised in Table 1, with the resources directly available to each study indicated.[9,10]

[8] There have also been other such studies; for example, Zamparini and Reggiani, (2007) and Shires and de Jong, (2009).

[9] In each of these studies, we understand that the resources available were far less than that were needed. In one sense, these studies were a 'victim of their own success'. Identifying more studies than expected meant greater data assembly costs. Any future study should not dis-incentivise search and assembly in order to meet budgetary constraints. We would estimate, based on accounts provided to us, that the funding covered around a quarter to a third of the time expended on search, data assembly and checking, analysis and reporting, but excluding the time involved in preparing journal papers.

[10] Two recent studies might serve as useful evidence bases here. An ongoing meta-analysis of international cross-elasticity evidence undertaken by ITS Leeds in conjunction with TOI Oslo spent of the order of 40 days identifying,

Table 1: Summary of Meta-Analysis Studies: Coverage and Resources

Study	Funder	When	Resource	Variables	Coverage	Obs	Years
I	DfT	1997	£6k	IVT	UK	444	1980-1996
II	DETR, OPRAF and PDFSS	2000	£14k	IVT, Congested, Free Flow, Walk, Wait, Access, OVT, Headway, Displacement, Parking Search, Late Time, Interchange Penalties	UK	1116	1980-1996
III	DfT	2003	10 days	IVT, Walk, Wait, Access, OVT and Headway	UK	1167	1963-2000
IV	DfT	2008-09	£25k[11]	As II except Interchange	UK	1749	1963-2008
V	EIB	2011-12	£30K	As IV plus SDE, SDL and Reliability Ratio	Europe	3109	1963-2011

3.2 **Requirements**

3.2.1 The original data

The databases underlying the meta-analyses reported in Table 1 are not readily available and would therefore need to be reconstructed from the original source studies. This would require a significant amount of additional effort, and would also introduce the risk that the original meta-analytical models cannot be replicated due to lack of access to the source data, and researchers making alternative interpretations when entering the relevant data (and necessary conversions) into the database.

In scoping the work below, we only take account of the activities required to extend the existing database. In the final section, we do provide an indication of the costs involved in recreating the data underlying meta-analysis V, i.e. the Wardman et al. (2016) data.

3.2.2 Objectives

As identified at the end of Option 2.1.2 in the Phase 1 report, a set of activities is required to make the meta-analysis an integral part of the robust VTTS monitoring programme. Below, we discuss the key tasks involved in this process and where possible separate between UK and European evidence.

assembling and checking a data set of around 1100 cross-elasticities. The data collection element of a recent review of diversion factors undertaken by RAND and SYSTRA for DfT, which identified a sample of 934 largely UK observations, involved around 25 person days. The RAND Europe and CE Delft (2004) study for the EIB had had 32 days for data collection, meta-model estimation and reporting on passenger and freight VTTS and VOR. The 2012 meta-analysis that Significance did for EIB on values of safety had 23 days for data collection, meta-model estimation and reporting.

[11] This budget was spread across meta-analyses of price elasticities and of time elasticities in addition to updating the value of time data set.

3.2.3 ***Work required***

The key tasks are summarised as follows:

- Update the UK-only and European VTTS meta-analysis databases to 2018.
- Undertake regular data collection efforts for newly emerging evidence.
- Contrast the new evidence against the official VTTS values.
- Regular re-estimation of the meta-analysis models using the new data.

In the remainder of this section, we pay specific detail to updating the existing databases to 2018 and briefly discuss the remaining tasks.

The updating task can be split up into three subtasks:

1. 'Mop-up' of previously unidentified or unobtained pre-2011 studies.
2. Revisiting the explanatory variables and collecting additional attributes.
3. Adding new valuation evidence over the 2011-2018 period.

'Mop-up'

Of course, not all relevant studies undertaken before the 2008 cut-off of Study IV or the 2011 cut-off of Study V will have been identified. Search methods are now better and we would expect this to identify further pre-2011 studies. In addition, some 'resistance' to providing evidence to support extended meta-analyses has been experienced by researchers in the past. We might expect support from the Department in the process to unlock studies prior to 2011 that were not available, particularly those conducted for local and regional government bodies. The Department might also be able to unlock good quality grey literature from other government organisations in Europe.[12]

'Retro-fit'

The second subtask of 'retrofitting' is envisaged to look back at the data that was collected from the set of pre-2011 source studies. The world of stated choice survey design and analysis for value of time estimation has developed significantly over the period of the various meta-analyses discussed above. A weakness of previous meta-analyses is that the explanatory variables collected then only cover issues of a methodological nature to a limited extent, such as how SP exercises are designed and the estimation procedures used. Such aspects have, however, become more important over time.

Going forward, it would be essential to account for these factors, through doing a 'retro-fit' by re-visiting the old source studies and collect the relevant explanatory variables where possible. This would then make the existing dataset compatible with what should be collected in any future updates to the meta-analysis.

[12] Whilst a subjective assessment, but based on extensive experience where data identified has exceeded expectations, we would not be surprised if such a 'mop-up' identified upwards of 250 valuations.

A second element of 'retro-fit' relates to adding variables that were not previously assembled. Income has a significant impact on values of time, yet there are two deficiencies with the existing data sets in this regard:

- Often, local income data attached to the values of time in the meta data-sets are not available (this is an issue that referees of submitted journal papers pick up on).
- Frequently, there has been no collection of within-study values of time segmented by income group, even though the assembled meta data-sets do contain within study valuations according to distance, purpose, mode and a range of other influential factors.

This would, however, require revisiting the data of the source studies and constructing such measures, which is likely to be a challenging exercise.

A third element of 'retro-fit', closely related to the first, is new variables. These include crowding, more granularity on congested time and interchange penalties.

We would therefore recommend an option to retro-fit the current data with details that would now be essential to understanding the value of time and to better understand income variations.

'Update 2011-2018'

The final subtask of updating the existing databases requires an extensive search of the published, unpublished (or grey literature) to identify all the relevant studies and register all the relevant valuations that have emerged after 2011.

'*Regular updating, estimation and contrasting*'

A natural extension of the above subtask is to undertake the same data collection efforts on an annual basis keeping the database up-to-date. When sourcing the new studies, the valuations can be directly contrasted by the corresponding VTTS values coming from i) the uprated official VTTS values and ii) the predictions of the meta-analytical model. The comparisons should give the first indications on whether changes in the VTTS (or related factors) are occurring. Together these form the second and third task of the work.

The new evidence covers the UK and mainland Europe over the period 2011-2018. We might expect identification of more valuations per year on the following grounds:

- More efficient search processes.
- Official backing and support.

Referring to Table 1, study IV in the years 2000-2008 identified on average only 6 studies per year, yielding on average around 50 additional observations per year across all attributes. Study V added material from mainland Europe. Over the same period, around 9 additional studies per year were identified and around 70 observations per year.

If the 2000-2008 experience was repeated over the period 2011-2018, we might expect across the UK and mainland Europe around 15 studies per year and 120 observations. This would imply 120 studies and around 1000 new observations. We take this to be a lower bound given the points made above.

We would recommend an ongoing monitoring process that identifies studies. Then every 3-5 years, depending on the volume of evidence – at least 25 new studies, there would be a joint update of the database and analysis. The advantages of assembling the data in one go is that it is more efficient and avoids inconsistencies. Moreover, additional analysis (i.e. estimation – the final task) is only relevant when sufficient new evidence is collected.

3.3 **Timescales and budget**

The Phase 1 report already highlighted the importance of meta-analysis in the proposed VTTS updating programme. Namely, validating the regular updates of the VTTS coming out of Stream 1 is through contrasting them with trends in newly published VTTS studies or by collecting new data. Meta-analysis offers a relatively cost-effective way of doing so by analysing trends in the VTTS over time from newly published studies in the UK and mainland Europe.

3.3.1 Rebuilding the original database

A significant proportion of the evidence, particularly the earlier evidence, was sourced from unpublished literature. We understand that it was awareness of the large emerging grey literature of the late 1980s and early 1990s, undertaken for organisations such as British Rail and national and local governments, which stimulated the initial meta-analysis. Much of this early evidence is available only in paper form. Matters have changed considerably in more recent years, with soft-copy reports available and a 'more enlightened' attitude to the publication of research reports by government bodies and industry alike. This material is much more accessible.[13]

Our estimates of the costs of re-assembling the already identified literature to replicate that in the most recent UK meta-analysis (Study IV) from scratch are approximately:

- Sourcing and delivering hard copy evidence: 5 person-days
- Coding and Inputting hard copy evidence: 15 person-days
- Sourcing soft copy evidence: 5 person-days
- Coding and Inputting soft copy evidence: 30 person-days
- Data checking, adjusting for inflation and cleaning: 2 person-days

[13] The most recent UK meta-analysis (Study IV) covers 226 studies. Of these 21% were operator unpublished reports, 8% were unpublished academic reports and 49% were unpublished government reports. The extension by largely European evidence, but with some additional UK evidence, in study V, largely involved readily accessible published material.

Sourcing, coding and inputting the European material and small amount of additional UK material of Study V would be of the order of 20 person-days. In monetary terms that would account for £57K for the UK database and £77K including European evidence (2018 prices, excluding VAT, and using the same £1,000 day rate as previously).[14]

We should point out two important factors relating to the resource requirements:

1. The data assembly was not undertaken by junior staff. Although 'literature reviews' are often assigned to junior staff, we would counsel against this as false economy.
2. Our experience is that significant inefficiencies, not to mention inconsistencies, can arise if data assembly is spread across a large team, suggesting that a smaller team is preferable. This might impact upon the time taken (weeks / months elapsed) to deliver the results.

*3.3.2 **Resourcing of updating and maintaining the meta-analyses***

Section 3.2.3 described four tasks in relation to the role of meta-analysis in the VTTS maintenance programme. Updating the UK-only and European meta-analyses, the latest update dates from 2011 for the European database, is the most time-consuming task in the list. Below, we an indication of the resources associated with this and the other tasks described in Section 3.2.3.

Updating the relevant datasets comprises three subtasks; i) mop-up of pre-2011 studies; ii) retro-fitting of methods, coverage and new values of studies already included in the dataset and iii) adding post 2011 studies to the dataset. Based on experience, we assume that 2/5 (40%) new studies will be UK-based and 3/5 (60%) European based. The difference in volume is considered when making cost estimates of the different subtasks of assembling the relevant database in Table 2.

With respect to the retro-fitting option, sourcing the data needs to be done irrespective of which of the three options is taken forward. We have therefore provided an overall search cost estimate for the retro-fit option.

Having the cost estimates in hand of the data assembly, we can now provide cost estimates for the four tasks specified at the start of Section 3.2.3, and link back to the recommendations as presented in the Phase 1 report. In Table 3, we combine the data assembly costs, as reported in Table 2, with the costs of estimating the required models on the improved data.

[14] The day rates are generally set by the framework under which the ITT will be distributed, but are likely to change over time. Hence, we have adopted a uniform day rate of £1,000.

Table 2: Resources going forward for UK and non-UK studies in data assembly (2018 prices, exc VAT)

	UK-only (40% of European)	European (including UK)
Mop-up		
- Search	- £1.6K	- £4K
- Assemble	- £3.2K	- £8K
- Checking	- £0.8K	- £2K
Retro-fit		
- Search	- £4.8K	- £12K
Retro-fit 1: methods		
- Assemble	- £4K	- £10K
- Checking	- £0.8K	- £2K
Retro-fit 2: coverage		
- Assemble	- £4K	- £10K
- Checking	- £0.8K	- £2K
Retro-fit 3: new values		
- Assemble	- £3.2K	- £8K
- Checking	- £0.4K	- £1K
Update		
- Search	- £4K	- £10K
- Assemble	- £7.2K	- £18K
- Checking	- £0.8K	- £2K
Ongoing: (every 3-5 years)		
- Search	- £3.2K	- £8K
- Assemble	- £6K	- £15K
- Checking	- £0.8K	- £2K

Table 3: Costing of alternative recommendations (2018 prices, exc VAT)

	UK-only	European (including UK)
Updating 2011-2018:		
a) Mop-up + update	a) Total: £33K	a) Total: £59K
• Data assembly	• *£18K**	• *£44K*
• Estimation and reporting	• *£15K*	• *£15K*
b) Mop-up + update + retro-fit	b) Total: *£51K*	b) Total: *£104K*
• Data assembly	• *£36K**	• *£89K*
• Estimation and reporting	• *£15K*	• *£15K*
Ongoing (every 3-5 years):		
c) Collect new data, no specification search	c) Total: *£15K*	c) Total: *£30K*
• Data assembly	• *£10K*	• *£25K*
• Estimation and reporting	• *£5K*	• *£5K*
d) Collect new data and specification search	d) *Total: £25K*	d) *Total: £40K*
• Data assembly	• *£10K*	• *£25K*
• Estimation and reporting	• *£15K*	• *£10K*
e) Options 1.1 and 1.2 fit-for-purpose test	e) Total: *£3K*	e) Total: *£3K*
* Rounding applied on data assembly costs from Table 2.		

With respect to updating over the 2011-2018 period, Option 2.1.2 of the Phase 1 report recommended two approaches:

a) Update the data over the 2011-2018 period and mop-up any missed studies.
b) Extend a) with the retro-fitting option.

For both options, the required resources are reported in the second and the third column of Table 3. It is assumed that analysis and reporting activities are comparable across the alternative approaches. These account for an additional 15 person-days for analysis and reporting are required. Together, option a) would cost around £33K at the specified day rate for the UK-only option. If the same exercise would be done for the European data, the total costs would account for around £59K. Adding the retro-fitting option to both specifications increases the cost by £18K and £45K to £51K and £104K, respectively (2018 prices, excluding VAT, and using the same £1,000 day rate as previously).

The second recommendation from the Phase 1 report on Option 2.1.2, considers maintaining the meta-analysis database by regularly collecting new data from 2018 onwards. The frequency depends on the number of emerging studies, but will most likely turn out at once every 3-5 years. Ideally, this would be combined with the 'Update 2011-2018' option to avoid a gap in the database.

We propose two variants. The first variant collects the new data and re-estimates the existing meta-analysis model without conducting a full model specification search. The analysis and reporting would account for 5 days. UK-only variant c) therefore has a total cost of around £15K (15 person-days). UK-only variant d) would conduct a full specification search on the updated dataset. This would add 10 person-days for additional analysis, increasing the costs to around £25K. Extending the variants to mainland Europe adds 15 days to the data collection costs, increasing the costs of variants 1 and 2 to around £30K and £40K, respectively. All costs are at 2018 prices, excluding VAT, and using the same £1,000 day rate as previously. Specifications c) and d) thus cover the second and the fourth task listed in Section 3.2.3.

Finally, the third task, of validating the new evidence is on the one hand part of the estimation and reporting procedure in c) and d) as researchers should check for outliers and conduct sensitivity analyses for the influence of specific observations. Any additional contrasting between the outcomes of the updated meta-analysis and the outcomes of Options 1.1 and 1.2, i.e. to test if these are still fit-for-purpose will be in the order of magnitude of £3K (option e) and will only add marginally to the costs and is best conducted with the latest data.

4. Option 2.1.3 - Revealed Preference

4.1 Background

4.1.1 Introduction

Chapters 4 and 5 summarise engagement with the Market Research industry and further work by the Arup and ITS study team. We invited previous contractors with the Department for Transport, and contractors from existing industry panels[15] to complete an online survey as a structured means of obtaining group feedback on these complex issues. We then invited the participants, alongside commissioners of market research work, to a round table discussion that was hosted by Arup and ITS Leeds at Arup offices on 25th of January 2018. The discussion with a diverse range of suppliers ranged from data collection, the approach to procurement, to risks and managing modelling work, and helped to give a rich insight into many of the aspects of the brief. It also allowed a deeper dive into some of the issues than was possible in Phase 1, helped to establish a potential supply chain for servicing future DfT commissions and functioned as a market sounding exercise. Following the round table, a further round of the online survey, relating to costs, was undertaken.

The Phase 1 report acknowledged, in Option 2.1.3, the appeal of using emerging revealed preference (RP) data sources, such as smart-ticketing and GPS-based data, combined with new modelling approaches for estimating future VTTS values. Traditional RP methods are perceived as being cost intensive and its confidence intervals tend to exceed those for SP estimates. A classic weakness of SP work, however, is its basis on declared rather than actual travel behaviour (which explains the appeal of RP methods). In seeking to ameliorate these issues, the use of emerging methods, such as mobile phone data, and journey planning data, may offer a good means of collecting RP instead of (or alongside) SP. The Phase 1 report also identified that with advances in data collection methods, the wealth and volume of information included in such sources is promising, but that applicability for national VTTS studies remains challenging.

4.1.2 SP vs RP

The primary challenge of estimating the VTTS, either from SP or RP data, is to isolate the effect of journey costs and time on individual travel decisions. SP methods essentially construct a 'laboratory condition' whereby journey time and cost are varied independently across choice tasks making it easier, from a statistical perspective to isolate their impact on decisions. In RP this is much harder, since journey time and costs are highly correlated. Moreover, SP studies allow to study more travel decisions per traveller thereby increasing the number of observations in the dataset relative to RP. Both SP and RP make use of the same discrete choice modelling (estimation)

[15] Including Transport Focus's list of approved market research contractors

techniques. Models on SP are usually more complex due to the larger volume of data and the higher informational content of each specifically designed choice task relative to RP methods.

RP methods have had their limitations and these have been acknowledged since the work by Beesly (1970); these have not been overcome yet. Emerging big data sources can mitigate these somewhat through the volume of data collected (i.e. the average should outweigh all the peculiarities that are picked up at the individual level in SP methods), but do not fully overcome these either.

A downside of such emerging big data sources is that the large level of detail and context that has been implemented in current SP models will be lost. Namely, SP surveys that are used to estimate VTTS collect a wealth of information from each survey respondent, including:

- mode of travel.
- journey purpose.
- travel costs and travel time by mode (preferably what ticket each user is on).
- other travel conditions (e.g. crowding and congestion).
- socio-economic characteristics (e.g. income, gender, age).
- household characteristics (e.g. car ownership, household size).

On the one hand these characteristics are used to present the respondent with a set of specifically designed (i.e. relevant) within-mode choice tasks and on the other hand they are used as explanatory variables in the econometric modelling. Emerging RP big data sources, unless combined with relevant background data, are not able to provide that level of detail. For example, at the time of writing, not all of these aspects about their customers are recorded by mobile phone network carriers (although some could perhaps be inferred through other online sources, or directly requested from participants). This is explored in more detail in the following section.

4.1.3 ***Considerations in using RP***

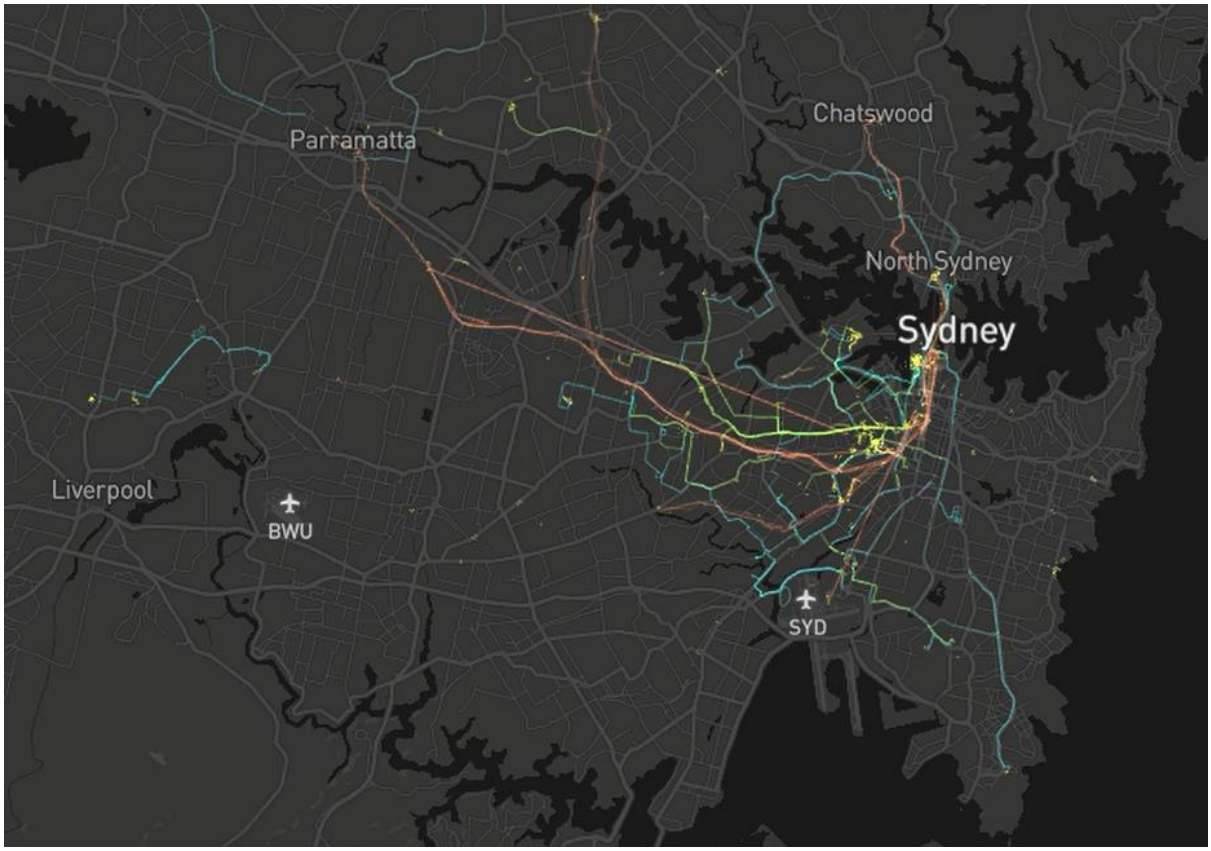
As mentioned in the Phase 1 report, there are three challenges in applying RP to value of travel time savings purposes, including national and representative coverage. The objective of sampling for a national VTTS survey should be to survey a truly random and representative set of travellers. Obtaining such a representative sample is typically done through setting quotas for specific segments of the population. Segmenting the population is typically done by means of income, gender and age data, but for VTTS studies geographical coverage, journey purpose, trip time, cost, duration and distance are other key variables. RP samples have potential to be more unrepresentative than SP samples (as the experimenter has less control). For example, the sample for smartphone based RP data is by definition restricted to smart phone users, which (although representing 70-80% of the population, and potentially a higher proportion of transport users) are likely to have a higher income level, and potentially, value of time, than the general population. As such, this would require “topping-up” with specifically targeted population segments, which has the potential to erode the cost savings from a move to RP.

Bespoke RP surveys

There are ways to control the sample by introducing RP techniques alongside traditional survey techniques. Inclusive RP surveys, where respondents have downloaded an app and responded to questions 'on the go', coupled with a special GPS device for non-smartphone users, could be used, and be paired with demographic information about the respondent (which overcomes some of the difficulties in the following section).

As an example of a GPS data source, an app called Mobility Mosaic (see Figure 1) tracks people's movements and trip chaining via their mobile phone. Other examples of similar apps are rMove and Future Mobility Sensing. Many mobile phone users have GPS enabled on their device and devices are increasingly transmitting passive information. Drop-out rates over the recording period and trip under-reporting (due to switching off phones or the GPS option), however, need to be monitored.

Figure 1 – Mobile phone (GPS) movements in Sydney from Mobility Mosaic App



Source – Mobility Mosaic, Arup

Essential respondent and trip characteristics, such as car ownership (or car access), can either be inferred by observing whether users take car trips from the home location over a period of time, such as one month; or such information can be observed when survey questions are included as an integral part of the app. Similarly, user origins and destinations can be inferred based on the time spent at the location and user characteristics can be obtained by matching to datasets where users live.

Some challenges associated with mobile phone data and other areas of interest are noted below:

1. **Passive data.**
    a. Celluar data accuracy varies depending on the generation (e.g. 2G / 3G / 4G). While earlier generations provide little “passive data” (periodic updates and crossing location area boundaries), newer generations provide much more. Therefore, as users gradually replace older devices, the quantity of mobile data will improve. This in turn will help to infer location and time of origin/destination, as well as routes, more accurately.
    b. GPS data offers significant spatial and temporal accuracy improvements over cellular data. This results in a more robust data source for origin/destination and route inference. However, current commercially available GPS data sources are generally single mode (e.g. car SatNavs) or heavily sample biased.
2. **Mode choice.** There are two issues related to this point. First, from passive mobile phone data it remains problematic to identify the mode of transport, or trip chain made. GPS sources combined with accelerometers can improve the necessary inference, but it remains a challenging task to separate cycling and walking trips from bus trips and car trips in congested urban conditions, and taxi from car trips in a wider range of conditions. Second, the 2014/15 (and preceding) SP VTTS studies have primarily focussed on estimating within-mode route choice models. RP studies used for VTTS purposes focus more on mode choice thereby requiring a shift in the types of choices models estimated.
3. **Vehicle type.** Vehicle type split is also difficult, such as commercial LGVs vs cars, however HGV are easier due to location identification.
4. **Journey purpose.** Purpose split can be done between Home-Based-Work and Home-Based-Other and Non-Home-Based, once home location and work location are inferred. Separating business trips is, however, more challenging. GPS route planning Apps now feature functionality which encourages a user to save trips under a label, such as “to work” and “to home”, enabling for limited journey purpose labelling.
5. **Individual attributes.** Mobile phone data has few user attributes (income, gender, age, car ownership etc.), but some of this might be available to the data owners via their customer databases, or through other online data sources. There are, of course, data privacy and confidentiality issues around this. Alternatively, it could be requested from the participants.
6. **Geographical coverage.** Passive mobile phone data may have sufficient geographical coverage for obtaining a representative national VTTS estimate. There is a risk that when moving towards a more app-based approach, the application becomes more of a case-study setting with limited geographical (and mode-purpose) coverage which can no longer be used for national VTTS purposes.

Ideal RP data:

If in the future these challenges could be overcome, technology may eventually produce enough data to enable VTTS to be estimated using established choice modelling techniques. This could, hypothetically, take the form of:

1. Obtain a 'one day snapshot' or more of all trips in the UK using mobile phone data – i.e. have all origins and destinations of movements in UK.
2. In parallel, derive a database of all travel choices available that day including actual journey times and fares / costs / fuel prices / crowding.
3. Estimate VTTS using (1) and (2) supplemented with the relevant background characteristics.

An example of this is that navigation apps are beginning to tailor their service offering based on their estimates of individual values of time. They may be able to infer people's value of time based on the choices made. However, there is the risk of introducing endogeneity when choice sets are generated when many more different options are available. Moreover, isolating the decision from noise (such as modal preference, interchanges, crowding levels, etc.), and knowledge about the subject remains challenging. In addition, the users of the apps currently may also be representative of a more affluent, urban section of the population, which may affect the inferred values of time.

It is important to make a distinction here between cellular and GPS data. Both have commercial value and are generally not available for use. Cellular data are held by mobile phone companies and require more manipulation inference to obtain the relevant origins and destinations and travel purposes. Mobile phone based GPS data are held by the likes of Apple and Google and contain a lot more detailed information.

A challenge with RP is obtaining robust data on non-chosen alternatives and the motivations for why a particular choice was made (e.g. prices, crowding and reliability levels). In the collection of RP data, it may be necessary to apply an intelligent screening process to ensure respondents understand the options available to them. On rail, technology change may be making the collection of RP data more robust – there is potential to expand on this through surveys while people are on their journeys when they connect to on-board Wi-Fi to interrogate further the reasons for the choice, aspects such as journey purpose, and other information, although this would in practice be limited to a few participants.

Existing RP data: individual sources

Nevertheless, there are potentially rich RP sources available that could be better exploited, potentially as benchmarks, for VTTS. There are a few corridors that have less noise (e.g. M6 Toll), and separately, some journey-planning apps (e.g. Citymapper, viaMichelin, Google Maps) offers choices about which mode and route to take, and give information about price, duration, mode and number of changes. Some also include information on congestion.

Through econometric-style techniques, and use of personal data (many apps have a sign-in process whereby home and work locations can be set), single variables could be isolated and journey purpose can be ascertained (or at least, guessed at e.g. home to work = commute). Potentially in future these could be expanded to cover more users, and more data about the subject – perhaps through sign in or other processes – could be obtained. In addition, other working-level, but important, issues might have to be overcome for the data to be valid (e.g. online tools may include certain journey options such as "fastest" at the top of the page), and there is a degree of inertia and habit in travel choices.

Finally, 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.

4.2 Timescales and budget

In summary, whilst there is an obvious attraction to the replacement of SP by RP methods, the case has not fully been made for this in the short term. In order to take steps to introduce RP as a method that takes more of a role in future valuation of travel time surveys, we recommend the following:

- To work alongside journey planning applications and other online providers to determine how suitable their data is to determine how people choose journeys based on travel time and cost. It may be that although the 'standard' version of the app cannot be used for VTTS, a variant version may be designed at small cost, or that future versions of the app may allow more data to be gathered about the user and their actual choices.
- In the meantime, there continues to be a role for RP as a means of validating SP using smaller scale case studies using GPS based apps or alternative mobile phone data.
- Similarly, macro-economic-style research into the impacts of Crossrail and HS2 (the latter in particular) could provide good RP case studies in the near future.

Prior to undertaking a large-scale data collection and analysis exercise, we suggest that the DfT should undertake the following tasks to ensure that the method is mature and accurate enough to be rolled-out fully:

- a (ongoing) scan of industry progress in the field of tracking mobile phones including GPS, cellular, but also WiFi and Bluetooth (could be done internally, but if done externally could cost some £5-10k)
- a review of the potential to link datasets to determine trip and user characteristics (around £10-£20k).
- a review of RP estimation methods if sufficient big data becomes available (around £5k).

- a pilot of the a GPS tracking app in the UK (potentially, one-third to one-half of the costs below).
- a pilot of how data scraping/harvesting and an RP method could work in practice (around £10k).

For RP app development, the app development itself could cost around £20-60k (2018 prices, excluding VAT), or potentially higher or lower depending on sophistication. Modification of an existing app (such as an existing journey planning platform) may be around the same costs, but would depend upon participation from a partner. The set-up costs are thus relatively small. Recruiting respondents and analysing the data would incur additional costs. Following this we suggest that the unit costs per survey collected would broadly be the same price as the online survey costs in the option 2.1.4 SP costs section below.

This would suggest itemised costs for new RP survey work of (2018 prices, exc VAT):

- App development £20-60K
- Data collection: £50-75K (1,000-1,500 respondents). Noting that incentive costs increase when apps need to record travel data /diary over a longer period.[16]
- Additional data is required to obtain the data of alternative travel options (i.e. the level-of-service data) if the design is not connected to another app.
- Data analysis and reporting may come to some £30-40K.

5 Options 2.1.4 and 2.2: Stated Preference

5.1 Background

We identified in the Phase 1 report that we believe that future large-scale VTTS SP studies should "move with the times" and that maintaining a given survey approach or modelling approach is not justified only on the basis of ensuring consistency. This introduces a discrepancy between Options 2.1.4 and 2.2 from the Phase 1 report. The former option aims to ensure that any future large-scale data collection exercise using SP methods does so using the latest survey design and estimation methods. The latter option, as part of the monitoring scheme, requires exact replication of the 2014/15 survey providing the cleanest comparison of the VTTS over time. That replicability is required to identify whether a new large-scale study is required.

The 2014/15 study was based largely on SP questionnaires, administered mainly through intercept surveys, and complemented by some telephone surveys. There are arguments in favour of both traditional and non-traditional methods of SP data collection: the advantage of the traditional intercept survey is that interviewers can be stationed where there is a good throughflow of travellers (e.g. at bus and rail stations),

[16] For example, a recent Arup-led study offered an incentive of £50 for three weeks RP app-based monitoring, compared with £10 for a traditional survey, which also reflective of a desire to recruit as many participants as possible.

with national representation gained through surveying at a range of different locations (and adjusting ex post for any residual bias using sample enumeration). In the 2014/15 study, telephone surveys were used to fill any gaps left by the intercept survey (and especially segments where trip rates were expected to be lower, and thus less likely to be captured by the intercepts).

We suggest that the key issues for taking forward emerging SP methods and the monitoring programme, are sampling, timing of survey, and medium of survey:

For **sampling**, VTTS faces the challenge that those with a high value of time are much less likely to take part in the survey[17]. Even in the face of a non-representative sample, post sampling adjustments (e.g. post-stratification weights) to correct for bias in terms of sample composition are helpful, but unable to correct for the self-selectivity issues of online panels or the bias incurred due to busier, high VTTS, individuals being less willing to participate. Without intercepts, potentially this risk is increased, and so ensuring quality data while still achieving a good sample size could inherently carry additional costs. This can be mitigated through survey improvements (easily controlled for by the market research team) and the sample for the survey (less easily controlled).

On the **timing of survey**, in order to help ensure robust results, it is necessary to try and minimise the cognitive gap between the survey and the reference trip, when moving away from direct intercept surveys; to self-completion on pen and paper; or online. The quality of responses can reduce due to the time elapsed between when respondents answer the survey and when the reference trip was taken.

For **medium of survey**, online surveys, which have been established for many years now, are a more cost-effective way to reach a wider sample, but are prone to some of the sampling issues described above (see the Dutch national value of time study as a case in point). Self-selecting online samples are particularly prone to bias; however, this can be partly mitigated by using randomly generated online panels. The latter have been used effectively by Nielsen, Kantar Public, Ipsos Mori and others. A recent study by RAND Europe, surprisingly found minor differences in the VTTS between respondents recruited through an online panel and intercepts.[18]

5.2 Timescales and Budget

As part of the follow-up to the round table discussions with a number of companies in the market research industry a follow-up online survey was undertaken relating the costs of alternative data collection methods. Below we provide insights into the costs of collecting SP data using alternative survey mediums.

[17] This also applies to many RP survey methods.

[18] https://www.rand.org/pubs/research_reports/RR2405.html

5.2.1 ***Intercept surveys***

Intercept surveys include interception of travellers during a transport trip, collection of data to check scope, and a further check that it aligns with sample quotas and collection of data for follow up.

Based on 1,000 rail intercept surveys on an annual basis, similar in scope to the last VTTS research and hence take about 25 minutes on average to complete, our market sounding with market research suppliers suggested cost of £30-£70k per year, including a £10 incentive for each participant (costs are at 2018 prices), depending on the presentation of results and other factors. This would be at the lower end of the range if reporting was simplified, or not needed at all. The costs may also differ according to the journey purpose with employers' business more costly, particularly for car (and not really available for bus), because of the ease of intercept and the response rate.

Indeed, a proportion of the variation in costs is driven by the level of screening at the intercept, as well as by the response rate (as not all the intercepts will complete the survey – and the survey length could be a deterrent to those with a high value of time).

Carrying out 300 surveys per quarter, to avoid seasonality, rather than 1,000 per annum could to push the costs up by some £10-15k.

5.2.2 ***Online panels***

The estimate for an online panel of 1,000 participants with a 30 minute questionnaire covering car, bus and tram for commuting or leisure would be about £20k (including an incentive worth £10 per complete, which could potentially be reduced for online participants). There might be significant project management costs associated with the presentation of results, and any tweaking to the survey, which could add a further £10k, to make the total of £30k. This could be avoided if the survey is developed as part of a wider piece of work, and if presentations are toned down or not needed. All costs are at 2018 prices and exclude VAT.

There would be no significant difference in costs for sampling three or four times a year.

5.2.3 ***App-based surveys***

SP app development is likely to be marginally cheaper than RP app development in terms of set-up costs (as the app has existed in some shape or form for several years). For SP app development, the app itself would cost around £10-30k (2018 prices, excluding VAT). Following this we suggest that the unit costs per survey would broadly be the same price as the online panel costs in the section above (but may have the same quality limitations due to selectivity).

5.3 **Large scale and rolling survey**

Since the 2014/15 study the SP playing field has not changed significantly. Hence, at this moment in time it is expected that, when the monitoring scheme indicates a new large-scale VTTS study is required, the costs for a new large-scale VTTS study (Option

2.1.4) would be in the same order of magnitude as the 2014/15 study. It is not foreseen that required sample sizes and levels of segmentation have changed in the current timeframe and similar resourcing would be required to develop the surveys and analyse the data.

Improvements to the framework can be made at small incremental costs by combining traditional sampling methods with non-traditional SP data gathering (through intercepts, then an app-based survey), or by collecting data on a seasonal basis. This helps shorten the cognitive gap between the journey and the survey, and provides the medium in a cost-effective means. Opportunities for cost savings are presented by moving to online-panels, but quality concerns may not support such a move.

A smaller scale annual rolling survey as part of the monitoring scheme (Option 2.2) would, however, come in at much lower annual costs. Relative to a large-scale VTTS survey, cost savings would be incurred in terms of i) survey development since an exact repeat of the 2014/15 survey is proposed, albeit on a smaller scale; ii) smaller samples would be collected, albeit on an annual basis and iii) costs would be saved on data analysis due to using the original 2014/15 models and estimation syntax, although some degree of cleaning and testing is still required.

What would be proposed under Option 2.2 would be to sample a different mode each year (e.g. year 1 car; year 2 rail; year 3 other PT and bus) on all journey purposes. Hence, after 3 years all mode purpose combinations would be surveyed. Original sample sizes were in the order of magnitude of 3,000 respondents for car, rail, bus and other PT, respectively. Taking at least 20% of those sample sizes, e.g. 250 respondents per journey purpose, would cost below £50K per year. In terms of analysis and reporting, approximately £15K would be needed on an annual basis since on average four models would need updating every year. Option 2.2 would thus be expected to come in at approximately £70K per year. All costs are at 2018 prices and exclude VAT.

6. Summary and recommendations

This Phase 2 report has provided more detail on several options that could be taken forward by the Department for Transport as part of a potential VTTS maintenance programme. This short list of options was identified at a project board meeting finalising Phase 1 of the project in November 2017. Before summarising the key findings of these specific options, we would like to re-iterate that it is important for the Department, and other European countries, to develop such a maintenance programme, particularly in a time of such rapid technological change in the transport and connectivity sectors. This would build a stronger case for conducting new large-scale VTTS studies when they are most appropriate instead of the infrequent and ad-hoc timing that is the state of current practice.

The Phase 1 report identified a potential structure for the maintenance programme, which has the overall aim of keeping the VTTS fit-for-purpose. Four streams were identified building on i) regular updates of the VTTS; ii) monitoring of changes in the VTTS; iii) improving the robustness of base and future VTTS values and iv) commissioning work on outstanding issues. The Phase 2 report primarily concerns the first two streams as together these form the core of the maintenance programme and place confidence in the derived 2014/15 behavioural models.

Table 4 presents the long-list of options of the Phase 1 report and indicates which options were feasible and which were taken forward in this Phase 2 report. For each of the options taken forward in Phase 2, we have described the relevant background, work requirements and relevant budget estimates. These budget indications are included in the final column of Table 4.

Table 4: Summary of options

Option	Name	Feasible	Phase 2	Cost*
1.1	Uprating the VTTS in line with GDP per capita growth	Yes	Yes	£55K
1.2	Uprating the VTTS using GDP per capita growth and NTS data	Yes (technically)	Yes	£29K
1.3	Adjusting future VTTS values using forecasts of NTS data	Too much uncertainty in forecasts	No	NA
2.1.1	Socio technological change	Yes	No	Low-medium
2.1.2	Meta-analysis	Yes	Yes	£144K
2.1.3	New RP data collection methods	Not yet	Yes	£100K-£175K
2.1.4	New SP data collection and analysis methods	Yes	Yes	Low-medium
2.2	Monitoring scheme	Yes	Yes	£70K
3.1	Understanding the 2014/15 confidence intervals	Yes	No	Low
3.2	Increasing robustness without introducing bias	Yes	No	Low
3.3	Uncertainty in future VTTS values	Yes	No	Low
4	Stream 4: Outstanding and emerging issues	Yes	No	Various
* Only options taken forward in Phase 2 have been given a monetary cost indication. Low <£50K; Medium <£200K				

6.1 Priorities

In terms of priorities, we recommend the Department to focus on Options 1.1; 1.2; 2.1.2 and 2.2. That is, Options 1.1 and 1.2 review the Department's current practice to adjusting the VTTS over time without the need for collecting new survey data. There is significant scope for setting new standards across Europe. Validation of the updated base year VTTS values from Option 1.1 and particularly Option 1.2 is, however, a key necessity to test the improved approach, but also to test whether these values are still fit-for-purpose 3-5 years (or perhaps 6-10 years) from now.

There are two options for validating the values from Option 1.1 and 1.2. The first one is meta-analysis (Option 2.1.2) and the second one is collecting small SP samples on an annual basis using exact repeats of the 2014/15 study (Option 2.2). Meta-analysis is perhaps the more cost-effective option of the two with the potential of obtaining buy-in from other European nations to keep the database up-to-date, although some upfront investment is needed to ensure the database is updated over the 2011-2018 period. The annual monitoring scheme using new SP data provides, however, a

clearer source of comparison for which the costs are primarily based on data collection. Namely, the existing 2014/15 model specifications and estimation code (on which ITS holds the IPR) can be used. There is again some degree of urgency since the data already dates from 2014 so already 4-5 years will have passed before a first attempt can be made. The longer an annual data collection scheme is postponed, the less likely it is that the additional 20% of SP data is comparable to the 2014/15 study and a larger initial sample may be required to improve the validity of the comparisons.

Options 2.1.3 and 2.1.4 are less urgent and are mainly concerned with ensuring that any future large-scale VTTS study will make use of the most appropriate data collection method. Since the early work in the 1970's, RP methods (Option 2.1.3) have been the most attractive source of data; but practicalities limit their application. There is an inherent attraction of grounding values in actual behaviour through RP. Our analysis has shown that these practicalities still exist despite promising developments in the context of big and emerging data sources. Based on discussions with Market Research companies in a roundtable discussion and responses to an online survey, we recommend the Department to review the ongoing evolution of RP methods and its fitness for purpose in terms of updating VTTS. Additionally, we do acknowledge that small scale SP validation exercises can be undertaken in the form of the use of mobility apps or real-world examples such as Crossrail and HS2.

Option 2.1.4 therefore focuses more on the development of SP data collection methods. Our analysis suggests with regards to major updates, that budget is unlikely to be significantly different from previous exercise. This result is primarily driven by the limited quality of online panels. Some improvements can, however, be implemented in the form of seasonal data collection and (or) the inclusion of app-based response formats. The general idea of an annual rolling survey (Option 2.2) was also well-received by the Market Research Companies making this monitoring scheme a viable option.

This study has been carried out alongside the introduction of the General Data Protection Regulations (GDPR) from 25 May 2018[19]. The most relevant of the reforms include the greater liability placed on data gatherers and controllers, an enhancement of individual rights – including a right to be forgotten – and data portability, as well as an obligation to promote these rights to individuals. Whilst we do not anticipate that the impact of these regulations would substantially affect our recommendations, we have not carried out a full assessment of this as part of our scope, and we recommend that DfT undertakes a separate piece of work in this regard. We suggest that because of the extra responsibilities shouldered by market research companies and other data holders, that they may, at the margin, increase the costs and risks of future VTTS studies.

[19] The UK Data Protection Bill will bring the European Union's General Data Protection Regulation (GDPR) into UK law

6.2 Costs, required expertise and procurement risks

Options 1.1 and 1.2 are very closely related and we would recommend bundling these together in a single tender costing around £84K. The referred costs are primarily associated with the development and testing of a framework. The nature of the work is therefore primarily academic. Once established the empirical work (e.g. sourcing NTS data) and running the Implementation Tool is suited for the regular procurement framework at significantly lower costs.

Option 2.1.2 is slightly more expensive at £144K (most extensive options) and this is primarily driven by the required 2011-2018 update. Costs could even turn out higher since there are IPR issues associated with the original UK and European dataset. This also limits the number of parties able to bid for this work. The required expertise also reflects this specialised work (see Chapter 3 for more detail).

Option 2.2 would be associated with an annual cost of approximately £70K. Working with the data would, however, require estimating the 2014/15 behavioural models which requires specialised software (Ox was used by the Arup and ITS team). Depending on the IPR of the original models, the work could be open to a larger group of bidders. Nevertheless, these still require a significant degree of econometric experience.

Options 2.1.3 and 2.1.4 are less clearly defined in terms of the actual work required and can take various directions. These options would however be accessible to a wider group of bidders.

The long-term nature of the maintenance programme in the form of annual data collection may cause some concern for the Department in terms of gaining a value for money price (for example, the letting of a contract over 10 years, say), although this could be overcome by splitting the time duration down into blocks of 3-4 years.

Despite the inevitable challenges we believe that the development of a VTTS maintenance programme is a necessary and promising avenue taken by the Department.

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

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

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

Daly, A. and Fox, J. (2012) Forecasting mode and destination choice responses to income change, *Paper presented at IATBR*, Toronto.

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

Jara-Diaz, S. and Rosales-Sarias, J. (2017) Beyond transport time: A review of time use modelling, *Transportation Research Part A: Policy and Practice*, 97, 209-230.

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.

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

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.

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.

Swärdh, J. (2008) Is the intertemporal income elasticity of the value of travel time unity? Working Paper 2008:3, Swedish National Road and Transport Institute.

Wardman M. (1998) The value of travel time: a review of British evidence, *Journal of transport economics and policy*, 32(3), 285-316.

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. and Shires, J.D. (2003) Review of fares elasticities in Great Britain. Working Paper. Institute of Transport Studies, University of Leeds, Leeds, UK.

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.

Zamparini, L. and Reggiani, A. (2007) Meta-Analysis and the Value of Travel Time Savings: A Transatlantic Perspective in Passenger Transport, *Network and spatial economics*, 7:377.

List of abbreviations

AHCG	Accent Marketing and Research and Hague Consulting Group
DfT	Department for Transport
EIB	European Investment Bank
EU	European Union
IPR	Intellectual Property Rights
ITS	Institute for Transport Studies
ITT	Invitation To Tender
GDP	Gross Domestic Product
NTS	National Travel Survey
OBR	Office for Budget Responsibility
SP	Stated Preference
VTTS	Value of Travel Time Savings
WTP	Willingness to Pay