



UNIVERSITY OF LEEDS

# Congested Values of Travel Time (CVTT)

## CVTT Forward Look

Richard Batley & Peter Mackie

29<sup>th</sup> September 2020

Report to the Department for Transport

ITS

# CONGESTED VALUES OF TRAVEL TIME (CVTT) CVTT FORWARD LOOK

# 1 INTRODUCTION

A team from the Institute for Transport Studies (ITS) at the University of Leeds was engaged by the Department for Transport (DfT) to undertake a 'Forward Look' in relation to the concept of Congested Values of Travel Time (CVTT).

In 2014/5, DfT commissioned Arup, ITS Leeds and Accent to undertake a major study to re-survey Values of Travel Time (VTT), and following a period of assimilation and consultation, TAG was updated in 2017 on the basis of the re-surveyed values. In the course of this study, Arup et al. employed Stated Preference (SP) survey methods to estimate not only a baseline VTT, but also a range of multipliers, which potentially allow the VTT to be adjusted up or down to reflect various aspects of comfort and quality, relative to the baseline<sup>1</sup>. Among these potential adjustments were a set of multipliers relating to traffic congestion – henceforth referred to as CVTT (Table 1).

These CVTT multipliers were not adopted in the 2017 update to TAG, due to concerns about their robustness (in particular, the multipliers appeared very high for some purposes), as well as the practicability of their implementation in modelling and appraisal. The high multipliers implied that car travellers would be willing to travel significantly longer distances to avoid heavy traffic, and the Department was concerned that this may not be consistent with actual travel behaviour. Instead the Department implemented values which were taken to represent 'Average' traffic conditions.

*Table 1: CVTT multipliers from 2014/5 Arup et al. study – expressed relative to Average traffic conditions*

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

Having specified the baseline VTT in terms of Average traffic conditions, the 2014/5 study reported CVTT as a multiple of the baseline VTT, for three different traffic conditions – namely Free-Flow (FF), Light and Heavy Congestion (Table 1).

Whereas the 2014/5 study was concerned with the estimation of VTT and CVTT, the current report is concerned more with the implementation of CVTT in modelling and appraisal and, as will become apparent in due course, it is more natural in this context to express CVTT as a multiple of the VTT for FF (as opposed to Average) traffic conditions (Table 2). Since the terminology has a degree of precedent, we will refer to multiples of the VTT for FF as 'M' multipliers.

---

<sup>1</sup> More specifically, in the case of car, SP1 considered time vs. cost, SP2 considered time vs. cost vs. reliability, and SP3 considered time vs. cost vs. congestion.

Table 2: CVTT multipliers from 2014/5 Arup et al. study – expressed relative to FF traffic conditions (aka ‘M’ multipliers)

Traffic conditions	Commute	EB	OtherNW
Free-flow	1.00	1.00	1.00
Light congestion	1.40	1.61	1.76
Heavy congestion	2.66	2.99	3.98

The Department is now reviewing the case for incorporating congestion multipliers in highways and other transport models – hence this Forward Look. DfT has so far commissioned a report from WSP, RAND Europe and Mott Macdonald on ‘Congestion Dependent Values of Time in Transport Modelling’, dated March 2018 and in the public domain. DfT has subsequently commissioned a follow-on report from the same team, which is dated May 2019 but not yet in the public domain. Together with the 2014/5 report, these two studies form the essential backdrop to the Forward Look.

Whilst the 2019 study will be considered in some detail, it is worth summarising the scope and outcomes of the 2018 study. WSP et al. were commissioned by DfT to conduct a feasibility study into the incorporation of congestion multipliers in highways and other transport models. The report considers various insights on this topic, namely:

- a technical review of the CVTT estimates emanating from the 2014/5 study by Arup et al.;
- a literature review of CVTT estimates from the UK and elsewhere in Europe;
- empirical testing of the route choice implications of CVTT using Trafficmaster data for a small sample of OD pairs.

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

Among these further works was a recommendation to investigate the practicability of implementing CVTT within both an existing highway assignment model and TUBA – and this in essence forms the basis of the 2019 follow-on study.

Against this backdrop, the Forward Look reported here is scoped around three overarching questions, underpinned by a series of motivating sub-questions, as follows:

**Q1. Has the 2019 CVTT study met its objectives?**

- Is the methodology proportionate and robust?
- Has it provided a fully defined proof of concept for applying CVTT?

- Has it included enough information about the potential impact of including CVTT?
- Has it highlighted sufficient barriers of implementation and are there any missing?

**Q2. How would the Department implement CVTT in appraisal and what are the barriers to this?**

- Does the concept of using CVTT fit within road scheme appraisal?
- If we did adopt benefits associated with CVTT in TAG, should they enter the initial BCR, adjusted BCR or the 'indicatively monetised' category of impacts?
- Would it be possible to develop an approach whereby CVTT are adopted in appraisal only, such as an 'uplift factor' to existing time benefits? What would be the risks associated with such an approach?
- Do we need to ensure there is functionality to apply CVTT in a wide range of commercial software packages before adopting them into TAG?

**Q3. What are the future research needs for CVTT?**

- Do you have any recommended research questions that need to be addressed after reading the 2019 CVTT study?
- Is it feasible to undertake a valuation study in order to estimate a delay time multiplier suitable for modelling and appraisal as set out in the report? How can we best avoid confounding congestion and reliability impacts in such work?
- What key things need to be considered within modelling and appraisal?
- Would CVTT for highway modelling cause a bias towards highways versus public transport and active travel and how would this be overcome?
- What is the importance of understanding the impacts the adoption of CVTT has on demand model calibration and matrix estimation?
- What further external expertise will be needed to progress this work?

In what follows, a chapter will be devoted to each of the overarching questions, and a closing chapter will summarise and propose recommendations.

## **2 HAS THE 2019 CVTT STUDY MET ITS OBJECTIVES?**

As part of our review of the 2019 CVTT study, we interviewed three key members of the study team (Andrew Gordon, James Fox and John Collins) using the script given in the annex to this report. We would like to put on record our thanks for their assistance and the valuable insights that they shared.

### **2.1 What were the objectives of the 2019 study?**

The objectives were threefold:

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

### **2.2 Methodology behind the work**

The first objective, which arguably represented the principal objective of the study, consisted of three tasks:

- i) Test CVTT in the base year highway assignment, to investigate how it affects the model validation.
- ii) Test CVTT in a future year forecast, with CVTT represented in the highway assignment model and the demand model. These tests included Do-Minimum (DM) and Do-Something (DS) scenarios, with the latter focussed upon a hypothetical major road scheme.
- iii) Use the results of ii) in DfT's TUBA economic appraisal software, to understand how the use of CVTT in modelling and/or appraisal affects the estimation of user benefits.

### **2.3 What was the modelling set-up for the 2019 study?**

#### ***2.3.1 Introduction to PRISM***

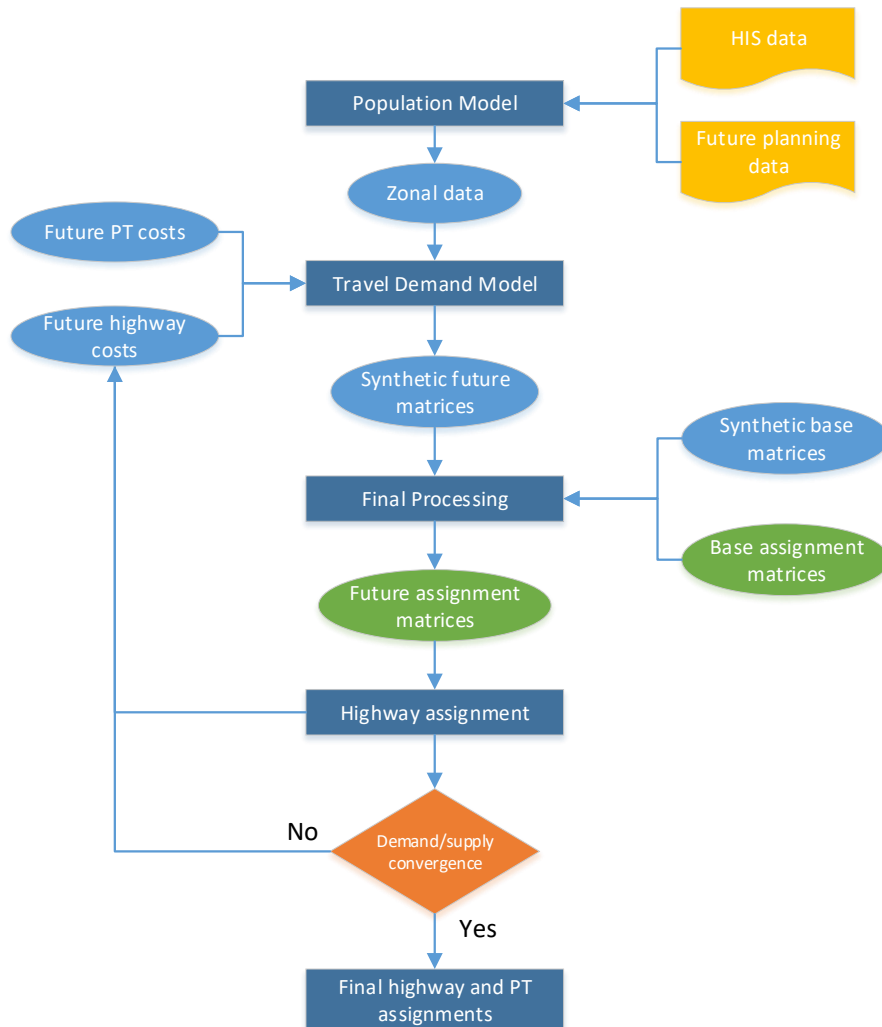
After consideration of various off-the-shelf models, the 2019 study opted to use the Policy Responsive Integrated Strategy Model (PRISM) – which is a multi-modal disaggregate demand model of the West Midlands Metropolitan Area. The model comprises separate Highway and Public Transport (PT) Assignment Models linked together with a Variable Demand Model (VDM).

Here we focus on the Highway Assignment Model (HAM), which represents an average weekday for three time periods (AM Peak from 0700 to 0930, Inter-Peak from 0930 to 1530, and PM Peak from 1530 to 1900). Five user-classes are modelled, namely Car Business, Car Work, Car Other, LGV and HGV.

The HAM employs an equilibrium assignment procedure incorporating detailed junction modelling and blocking back within the 'Area of Detailed Modelling'. This links to the VDM which comprises three main components, namely a Population Model, Travel Demand Model and Final Processing Model.

The VDM predicts Base and Future Year synthetic trip matrices for a given set of demographic data and travel costs, calculates the differences between these matrices, and applies the differences to the Base Year validated matrices. The outcome of this process is a set of revised trips matrices for assignment in the HAM. The structure of PRISM is summarised in Figure 1.

Figure 1: Structure of PRISM



### 2.3.2 Adjustments to 'standard' PRISM

From our reading of the 2019 report, we understand that the study team made various adjustments to the 'standard' PRISM model in order to introduce CVTT and set the model up for the subsequent tests. Among these adjustments, we believe that the following were of particular significance.

### a) Base Year matrix

The 'prior' Base Year (BY) matrix was built from various data sources and assigned to the network, wherein routing depended on various factors including a Generalised Cost (GC) function – within which VTT was a key variable. Ordinarily, routing information from the assignment would be used to update the BY matrix – with the objective of achieving better fit, and thereby producing a 'post-estimation' BY matrix. However, in the 2019 CVTT study, the latter stage was not conducted because of the time/resource that would have been involved in re-estimating the matrix. In effect, the 2019 work purposely employed the prior or 'pre-estimation' BY matrix.

### b) Demand/supply interaction

Again because of time/resource constraints, the VDM was not re-estimated on the prior BY matrix – and instead the pre-existing VDM estimated on the post-estimation BY matrix (not accounting for CVTT) was used. The opinion of the 2019 team was that the difference in the BY matrix would not have a substantive impact on the VDM<sup>2</sup>. Another way of putting this would be to say that, in the 2019 work, the VDM previously estimated on GC data<sup>3</sup> derived from the post-estimation BY matrix, was applied to forecasting using GC data derived from the prior BY matrix.

### c) Adjustment to generalised cost function in assignment

In order to introduce congestion to PRISM, the GC function was amended such that the time-based component of cost was separated into Free-Flow (FF) and Delay, essentially as follows:

$$VTT_{FreeFlow} \cdot T_{FreeFlow} + VTT_{Delay} \cdot T_{Delay} \quad (\text{Eq.1})$$

where:

$T_{FreeFlow}$  is the Free-Flow time (minutes)

$T_{Delay}$  is the Delay time (minutes)

$VTT_{FreeFlow}$  is the value of Free-Flow time (p/min)

$VTT_{Delay}$  is the value of Delay (p/min)

(Eq.1) was then operationalised by means of two sets of multipliers of VTT, as follows.

#### **The M multiplier**

First, time units were 'normalised' by expressing VTT entirely in terms of FF, and then representing the incremental disutility of Delay via M multipliers, thus:

---

<sup>2</sup> That is to say, the pre-existing VDM was originally estimated using household survey data from the West Midlands in combination with travel times and costs skimmed from the HAM – where the latter was based on the post-estimation BY matrix. If the matrix loaded into the HAM were changed – as was the case in the 2019 study – then strictly speaking the VDM should be re-estimated.

<sup>3</sup> In practice, different transformations were applied to time and cost variables within GC – as will become apparent when deriving VTT (see Table 4 below).



$$VTT_{FreeFlow} \cdot (T_{FreeFlow} + (M \cdot T_{Delay})) \quad (Eq.2)$$

where  $M = VTT_{Delay} / VTT_{FreeFlow}$ .

As we will see in due course, three assumed values for CVTT were applied to the BY tests (namely 'Low' (M=1), 'Medium' (M=3) and 'High' (M=5)), whereas only the latter two values were applied to the FY tests. It should be noted that these assumed values covered the range of estimates from the 2014/5 study (Table 2).

### **The r multiplier**

Second, since the TAG VTT is defined in terms of average traffic conditions, but (Eq.2) is defined in terms of FF, a second multiplier (r) was applied in order derive  $VTT_{FreeFlow}$  from TAG, i.e.

$$r = VTT_{FreeFlow} / VTT_{Average}$$

Different r multipliers were used for different journey purposes, specifically: Commute = 0.6968; EB = 0.5718; OtherNW = 0.5008. The 2019 report asserts that these were elicited from the relativities between the SP1 ( $VTT_{Average}$ ) and SP3 ( $VTT_{FreeFlow}$ ) values reported in the 2014/5 Arup et al. study. Checking their provenance more carefully, it would seem that these multipliers were drawn from Table 4.11 (p138) of the 2014/5 report, an excerpt of which is reproduced below.

*Table 3: Excerpt of estimates from behavioural model in 2014/5 VTT study*

<i>game specific <math>\theta_0</math> multipliers</i>	est.	rob t-rat. (1)	est.	rob t-rat. (1)	est.	rob t-rat. (1)
SP1 travel time	1	-	1	-	1	-
SP2 travel time	1.5988	4.05	1.1396	0.82	2.1875	5.52
SP2 std dev of travel time	0.5803	-4.75	0.8765	-1.04	0.8118	-1.48
SP3 free-flow	0.6968	-2.26	0.5718	-4.54	0.5008	-4.43
SP3 light traffic	0.9770	-0.14	0.9206	-0.74	0.8801	-0.90
SP3 heavy traffic	1.8557	2.98	1.7076	4.23	1.9955	4.05

Unfortunately, the 2019 study appears to have overlooked guidance given prior to the table (p134) in the 2014/5 report, thus: “As discussed above (cf. equation (4.46)), the game specific multipliers cannot directly be interpreted as the differences in the valuations across games as these differences are a function of  $\Delta t$  and the impact of this differs across the valuations in the different games. As such, analysis of the differences across games will happen at the implementation stage in Chapter 7, on the basis of specific assumptions about  $\Delta t$ ”.

The practical implementation of the above guidance is that the 2019 study should have based the r multipliers on the first row of Table 1, i.e. based on the nationally representative estimates of VTT and CVTT emanating from the sample enumeration process across the population of NTS travellers. Moreover, the correct multipliers (and the ratio to those actually used in brackets) are: Commute = 0.51 (0.73); EB = 0.42 (0.73); OtherNW = 0.47 (0.94). In short, the 2019 study inadvertently inflated the r multipliers across all journey purpose, and particularly so for Commute and EB.

### **d) Adjustment to utility function in VDM and implied VTT**

Corresponding to the above adjustments to the assignment models, adjustments were also made to the utility functions of the VDM, which in turn had implications for the implied VTT which we derive

here in Table 4. In particular, the table gives distinct VTTs for Free-Flow and Delay, expressed as a function of  $VTT_{Average}$  and the M and r multipliers.

Taking a closer look at the formulation of the VDM via public domain reports (Mott MacDonald, 2019), it might be noted that the VDM embodies a nested logit model with several different choice dimensions nested at different levels in the tree structure. Whilst practical operation is simplified through various working assumptions, the model in principle captures a wide range of choices, as follows:

- main mode choice
- car driver time period choice
- train & metro access mode choice
- destination choice
- train & metro station choice

The tree structure gives rise to a natural hierarchy of the choice dimensions, with main mode and time period (in that order) typically found to be the most important choices (i.e. in terms of the relative magnitude of their respective nesting – usually denoted ‘lambda’ – parameters).

In deploying the pre-existing VDM, the 2019 study effectively assumed that the sensitivity of travellers to the above choice dimensions was unaffected by the introduction of congestion into the model via GC. It is difficult to speculate upon the significance of this assumption – but intuition suggests that trip re-timing would be a primary mitigation for congestion, and it is therefore conceivable that sensitivity to time period choice would be affected by the introduction of congestion to the model via GC<sup>4</sup>.

---

<sup>4</sup> This point follows from footnote 2. In more technical terms, we are speculating that, if the VDM were re-estimated on the prior BY matrix, then the nesting parameter relating to time period choice could change – at least in relation to other nesting parameters.

Table 4: Adjustments to the utility function and implied VTT

		Original formulation (i.e. Average congestion)	Adjusted formulation (i.e. FF and Delay)
Non-HBEB	U	$U = \beta_{\text{time}} * T$ $+ \beta_{\text{cost}} \left\{ \gamma \cdot C + (1 - \gamma) \cdot \log(C) \cdot \frac{E(C)}{E(\log(C))} \right\}$	$U = r \cdot \beta_{\text{time}} \cdot (T_{\text{FreeFlow}} + M \cdot T_{\text{Delay}})$ $+ \beta_{\text{cost}} \left\{ \gamma \cdot C + (1 - \gamma) \cdot \log(C) \cdot \frac{E(C)}{E(\log(C))} \right\}$
	VTT	$VTT_{\text{Average}} = \frac{\beta_{\text{time}}}{\beta_{\text{cost}} \cdot \left\{ \gamma + \frac{(1 - \gamma)}{C} \cdot \frac{E(C)}{E(\log(C))} \right\}}$	$VTT_{\text{FreeFlow}} = r \cdot VTT_{\text{Average}} = \frac{VTT_{\text{FreeFlow}}}{VTT_{\text{Average}}} \cdot VTT_{\text{Average}}$ $VTT_{\text{Delay}} = r \cdot M \cdot VTT_{\text{Average}} = \frac{VTT_{\text{FreeFlow}}}{VTT_{\text{Average}}} \cdot \frac{VTT_{\text{Delay}}}{VTT_{\text{FreeFlow}}} \cdot VTT_{\text{Average}}$
HBEB	U	$U = \beta_{\text{time}} \cdot \left\{ T + \frac{C}{VTT_{\text{Average}}(d)} \right\}$	$U = r \cdot \beta_{\text{time}} \cdot (T_{\text{FreeFlow}} + M \cdot T_{\text{Delay}}) + \beta_{\text{time}} \cdot \frac{C}{VTT_{\text{Average}}(d)}$
	VTT	$VTT_{\text{Average}}(d)$	$VTT_{\text{FreeFlow}}(d) = r \cdot VTT_{\text{Average}}(d) = \frac{VTT_{\text{FreeFlow}}(d)}{VTT_{\text{Average}}(d)} \cdot \frac{\beta_{\text{time}} \cdot VTT_{\text{Average}}(d)}{\beta_{\text{time}}}$ $VTT_{\text{Delay}}(d) = r \cdot M \cdot VTT_{\text{Average}}(d) = \frac{VTT_{\text{FreeFlow}}(d)}{VTT_{\text{Average}}(d)} \cdot \frac{VTT_{\text{Delay}}(d)}{VTT_{\text{FreeFlow}}(d)} \cdot \frac{\beta_{\text{time}} \cdot VTT_{\text{Average}}(d)}{\beta_{\text{time}}}$

Where:

T is the travel time (minutes)

C is the travel cost (£)

d is the travel distance (miles)

## 2.4 Rationale behind the tests

The 2019 study conducted three sets of tests, referred to as Base Year tests, Future Year tests and TUBA tests. Here we briefly set out our interpretation of the rationale for these tests – using the standard demand/supply diagram as an expositional device (Figure 2).

Whilst the demand function is shown as linear, the supply function is shown as non-linear – since congestion causes supply to quickly become inelastic as road capacity is reached. Although not the focus of the discussion here, it should be borne in mind that, in practice, the magnitude of user benefits will be highly sensitive to the linearity of the supply function in the region of capacity.

To define the additional notation introduced here, the 0 and 1 superscripts refer to the Do-Minimum (DM) and Do-Something (DS) scenarios respectively, and the m and a subscripts refer to Modelling and Appraisal respectively.

According to our reading, the essence of the 2019 tests is as follows.

### 2.4.1 *Base Year tests*

Forecasts of Base Year demand/supply were analysed according to various calibration and validation criteria, namely:

- Convergence
- Run times
- Flow differences
- Screenline flows
- Validation – link flows
- Validation – journey times

### 2.4.2 *Future Year tests*

Forecasts of future year demand/supply for both DM ( $Q^0$  in the figure) and DS ( $Q^1$ ) scenarios – where the latter was focussed upon a hypothetical scheme (see below for further description) in 2036 – were analysed according to various forecasting criteria, namely:

- Highway assignment total travel time and distance
- Implied average VTT
- Mode shares
- Average trip lengths
- Link flows
- OD routeing

### 2.4.3 TUBA tests

These tests are arguably the most important component of the testing regime, and involved comparing the user benefits deriving from the hypothetical scheme under essentially two approaches:

- i) The conventional approach is followed, whereby the Rule-of-a-Half (RoH) measure of user benefits is calculated such that the generalised cost and demand inputs to the RoH are consistent with those emanating from the model. This gives rise to the benefit trapezium  $GC_m^0, a, b, GC_m^1$  (which is drawn in blue).
- ii) An alternative approach is followed, whereby for a given level of modelled demand (which for expositional simplicity, we assume here to be that emanating from i)), the GC inputs to the RoH measure are uprated (or downrated) based on various assumptions concerning the congestion multiplier, but the demand inputs to the RoH are as per i). In effect, this implies a manipulation of the demand and supply functions used to derive user benefits, such that supply becomes perfectly inelastic ( $S^0(GC_a)$  and  $S^1(GC_a)$ ) at the modelled quantity ( $Q^0$  and  $Q^1$ , respectively), and demand shifts outwards accordingly (from  $D(GC_m)$  and  $D(GC_a)$ )<sup>5</sup>. At face value, this manipulation would seem contentious – but would seem most defensible where the network is operating at or beyond capacity, since in this case the supply function would be genuinely vertical, and the generalised cost of travel would be ill-defined. In this latter case, it might be argued that the modelled generalised costs represent a lower bound estimate. This gives rise to the benefit trapezium  $GC_a^0, c, d, GC_a^1$  (which is drawn in red).

Moreover, it is clear from the RoH formula that if – *for purposes of appraisal* – the generalised costs for states 0 and 1 are adjusted through the addition of a common constant (as would apply if implementing a common congestion multiplier across the full distribution of traffic), then the conventional trapezium  $GC_m^0, a, b, GC_m^1$  would be equivalent to the adjusted trapezium  $GC_a^0, c, d, GC_a^1$ , i.e.

$$RoH = 1/2 \cdot (GC^0 - GC^1) \cdot (Q^0 + Q^1) = 1/2 \cdot ((GC^0 + k) - (GC^1 + k)) \cdot (Q^0 + Q^1)$$

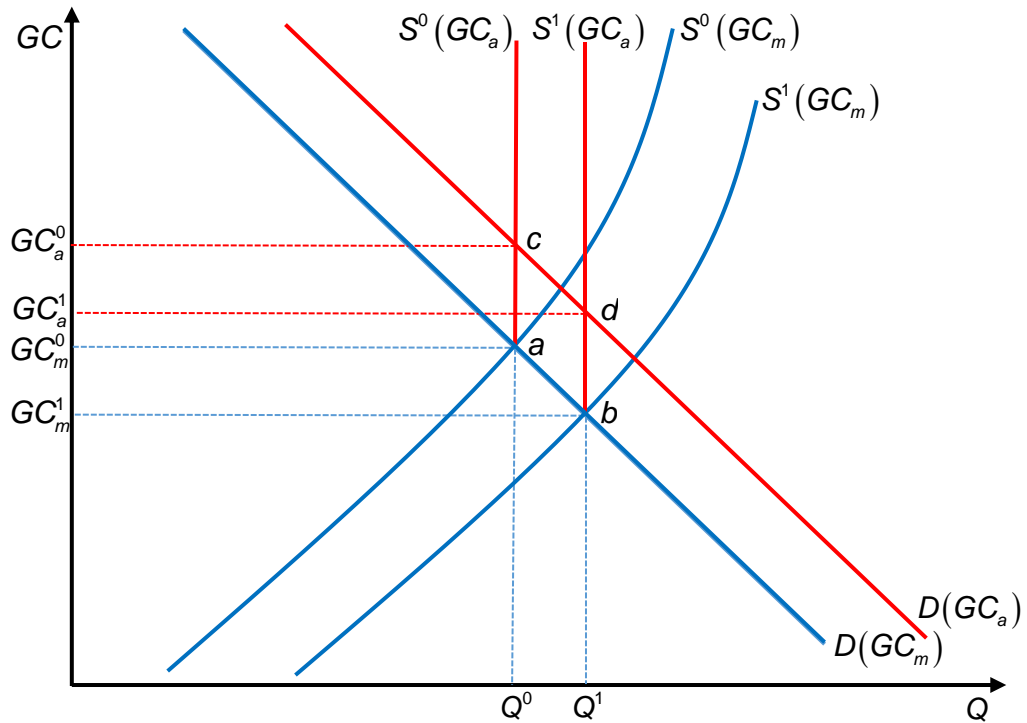
(Eq.3)

---

<sup>5</sup> As has been pointed out to us by both the Department and the WSP et al. (2019) team, approach ii) above makes no (explicit) assertions concerning the features of D/S equilibrium, and seeks only to uprate or downrate the inputs to the RoH measure. That said, when rationalised within the context of D/S equilibrium, it would seem inescapable that this implies a shift in the D function.

However, one would expect the two trapeziums to quickly diverge if different congestion multipliers are applied and/or the two states give rise to different mixes of traffic and thus different average VTTs across the distribution (indeed, one would in practice expect this to be the case, since the whole point of the scheme is to alleviate levels of congestion). The scale of this divergence in the RoH is the focus of the 2019 study.

Figure 2: Rationalising the TUBA tests in the context of the RoH diagram



## 2.5 The tests

Based on the rationale outlined above, the 2019 study reports the results from the BY, FY and TUBA tests for selective permutations of modelling and appraisal inputs as given in Table 5. The different M cases have been described previously. As regards the 'TAG' case in Table 5, we understand that this refers to the case where  $r=1$  and  $M=1$  – although the report could have been more definitive on this point. This is distinct from the Delay ( $M=1$ ) case in that the latter is associated with  $r<1$ .

Table 5: Schedule of tests

APPRAISAL					
MODEL		TAG (average traffic)	Delay (M=1)	Delay (M=3)	Delay (M=5)
	TAG (average traffic)	Table 7	Table 7	Table 7	Table 7
	Delay (M=1)	-	-	-	-
	Delay (M=3)	Table 8	-	Table 8	-
	Delay (M=5)	Table 8	-	-	Table 8

## 2.6 The hypothetical scheme

The hypothetical scheme is shown in Figure 3.

The key features of the scheme are as follows:

- A new D4M motorway to the west of the West Midlands conurbation, providing an alternative route to the existing M5;
- A new junction with the M6 at the northern end between M6 J11a and J12;
- A new junction with the M5 at the southern end between M5 J4 and J5a;
- Intermediate junctions with the A456, A454, and M54;
- All new junctions coded as free-flow (unlimited capacity) for all movements, to eliminate delays.

## 2.7 The results

### 2.7.1 Base Year tests

The 2019 study reports that: *“Overall, the level of model validation tends to deteriorate as the value of the congestion multiplier increases. The impact is largest for individual link flows. There is a smaller effect on screenline flows and no clear pattern for journey times”.*

However, it is qualified that: *“It would be unwise to draw any general conclusions from this. As noted above, the development of the prior matrices implicitly depends on WebTAG VTT. If different VTTs (i.e. with congestion multipliers) had been used to develop the prior matrices, the validation results would have been different”.*

They continue: *“Nevertheless, the results do show that use of the multipliers has the potential to affect the validation results, and could make the difference between achieving WebTAG targets and not achieving them”.*

Figure 3: Map showing hypothetical scheme – reproduced from WSP et al. (2019)





## 2.7.2 Future Year tests

The 2019 study finds that: “...total vehicle kms and average car trip lengths tend to increase with the multiplier, albeit by relatively modest amounts”.

However, they also report that: “Analysis of a small sample of OD pairs suggests that, in many cases, the impact on routeing in the model will be negligible and may be confined to those journeys where there is a realistic choice between a short, congested route, and a longer, less congested alternative”.

Among the various FY tests, the ‘outturn’ average VTTs (Table 4) would seem of particular interest given the focus of this Forward Look. The 2019 report does not document the basis of these calculations, but we presume that the consultants have applied equation 6 from their 2018 report, thus:

$$VTT_{\text{Weighted}} = \frac{VTT_{\text{FreeFlow}} \cdot T_{\text{FreeFlow}} + VTT_{\text{Delay}} \cdot T_{\text{Delay}}}{T_{\text{FreeFlow}} + T_{\text{Delay}}} \quad (\text{Eq.4})$$

Re-stating this in terms of  $VTT_{\text{Average}}$  and the r and M multipliers using Table 2, we have:

$$\begin{aligned} VTT_{\text{Weighted}} &= \frac{r \cdot VTT_{\text{Average}} \cdot T_{\text{FreeFlow}} + r \cdot M \cdot VTT_{\text{Average}} \cdot T_{\text{Delay}}}{T_{\text{FreeFlow}} + T_{\text{Delay}}} \\ &= \left( r \cdot VTT_{\text{Average}} \cdot \frac{T_{\text{FreeFlow}} + M \cdot T_{\text{Delay}}}{T_{\text{FreeFlow}} + T_{\text{Delay}}} \right) \end{aligned} \quad (\text{Eq.5})$$

Thus, the weighted average VTTs effectively pivot off  $VTT_{\text{Average}}$ , as a function of the assumed r and M multipliers, together with the quantities of FF and Delay time forecasted by the model. On this basis, if we were to consider whether the outturn weighted average VTTs differ significantly from  $VTT_{\text{Average}}$ , then the practical question is whether the following term is significantly different from one:

$$r \cdot \frac{T_{\text{FreeFlow}} + M \cdot T_{\text{Delay}}}{T_{\text{FreeFlow}} + T_{\text{Delay}}}$$

The substance of Table 6 is taken directly from the 2019 study report, but for comparison we also append the corresponding TAG guidance VTT (in the final ‘greyed out’ row of the table).

Table 6: Outturn weighted average VTT (£/hr per vehicle, 2010 prices)

Scenario	M	Commute	EB	Other
DM	TAG	16.82	25.25	11.96
DM	3	14.68	20.86	8.64
DM	5	17.11	26.38	10.91
DS	TAG	16.82	25.25	11.96
DS	3	14.59	20.72	8.59
DS	5	16.98	26.14	10.85
TAG guidance <sup>6</sup>		16.49	29.69	12.53

Focussing on the Commute journey purpose, which is where congestion will be most prevalent, it can be seen that the modelled VTT based on TAG is closely proximate to the TAG guidance value. This should come as no surprise bearing in mind that TAG was one of the benchmarks for the original VDM calibration.

The table shows a consistent pattern of the weighted average VTT for the medium multiplier (M=3) always being less than the corresponding TAG VTT. The weighted VTT for the high multiplier (M=5) is greater than TAG for Commute and EB, but not for Other. The VTTs for the DM and DS show little divergence.

The 2019 study reasons that: *“when the medium or high multiplier is used, the free-flow VTT is 30-50% lower than the TAG VTT. With the medium multiplier, delay is valued at 3 times the free-flow value, but it seems there is not enough delay in the model system to bring the average VTT up to TAG levels. With the high multiplier the level of delay is enough to raise the average VTT just above TAG for Commute and EB, but not for Other. This is probably a result of the free-flow VTT for Other being only 50% of TAG, significantly lower than for other purposes”*.

In other words, whereas the TAG model runs treat all travel time as representative of Average congestion, once congestion multipliers are introduced, the model effectively reallocates travel time between FF and Delay. Focussing again on the Commute journey purpose, it can be seen that a multiplier of M=3 undervalues VTT relative to TAG, whereas M=5 overvalues – albeit modestly so.

---

<sup>6</sup> These are ‘all week’ averages taken from the TAG databook tab A1.3.6.

Recalling that the original VDM calibration was based on the TAG VTT, the suggestion here is that a congestion multiplier intermediate between  $M=3$  and  $M=5$ , but closer to the latter than the former, would recover the TAG VTT. This result indicates that, despite not having been recalibrated, the VDM remains reasonably robust to the introduction of congestion multipliers in the upper range.

However, we would caution against using this result to draw more general interpretation concerning guidance values of  $M$  – since it relates to a particular model (PRISM), study area (West Midlands) and scheme. It remains to be seen whether the result would be transferable to other contexts.

Intuition suggests that, in general, the two primary mitigations available to car drivers in the event of congestion would be re-routing and re-scheduling. Whereas the former was the focus of the 2019 case study, the latter attracted limited attention – and this would therefore suggest that the congestion multipliers were largely picking up the (dis)benefits of re-routing, as well as re-distribution.

The VDM within PRISM does feature time period choice, but only in terms of a rather coarse segmentation into AM Peak, PM Peak and Inter-Peak, and only for car drivers and not passengers<sup>7</sup>. In order to better understand re-scheduling as a potential mitigation for congestion, a finer time resolution would be needed, especially around the shoulder of the peak. If such re-scheduling were implemented in PRISM, then re-routing behaviour might diminish, and by implication  $M$  might decrease – but this would then provoke a subsidiary question of whether the dis-benefits of re-scheduling are being omitted.

Moreover, behavioural mitigations on the part of car drivers (and indeed passengers) will generally manifest as demand changes – and it is thus crucial to understand how demand responds to congestion (mindful of course of the inherent endogeneity between demand and congestion). In our view, one substantive omission from the tests conducted in the 2019 study is inspection of the implied elasticities of demand with respect to travel time<sup>8</sup>. It would have been instructive to estimate such elasticities for the FY, with the change in demand between the DS and the DM formulated in terms of both trips and distance, and the corresponding change in travel time interpreted in terms of both the actual travel time and the implied travel time associated with the introduction of CVTT. More specifically:

---

<sup>7</sup> Footnote 14 of the 2019 study remarks: *“The model treats car driver and car passenger utilities slightly differently. The former mode has a time period choice model and utilities by time period are used in the demand model; the latter does not have a time period choice model and uses all-day average utilities”*.

<sup>8</sup> One of the recommendations of the 2019 report (paragraph 7.5.3) was to conduct realism testing – albeit in terms of the BY rather than the FY.

$$\varepsilon_{D,T} = \frac{(D^1 - D^0)/D^0}{(T^1 - T^0)/T^0} = \frac{\ln(D^1/D^0)}{\ln(T^1/T^0)} \quad (\text{Eq.6})$$

where:

$$\text{Actual travel time: } T = T_{\text{FreeFlow}} + T_{\text{Delay}}$$

$$\text{Implied travel time: } T = r \cdot (T_{\text{FreeFlow}} + M \cdot T_{\text{Delay}})$$

To give an example of the translation from actual to implied congested travel time using the relevant multipliers from the 2019 study, consider a Commute journey ( $r=0.51$ ) with an average travel time of 60 mins comprising 10 mins FF and 50 mins Delay. Assuming a medium congestion multiplier ( $M=3$ ), this would translate to:

$$T = 0.51 \cdot (10 + 3 \cdot 50) = 81.6 \quad (\text{Eq.7})$$

From other technical papers on PRISM we know that, in the course of the VDM calibration, the ‘overall’<sup>9</sup> elasticity of demand with respect to travel time<sup>10</sup> was -0.214 in terms of trips and -1.69 in terms of distance. Looking specifically at the Home-Work purpose, where intuition suggest that congestion would be especially prevalent, the corresponding elasticities from the VDM calibration were -0.153 and -0.760 respectively – meaning that the travel time elasticity was relatively inelastic in terms of trips, but approached unit elasticity in terms of distance.

Having now introduced CVTT to PRISM, it would be instructive to review the elasticity properties of the model, through both benchmarking against the calibrated elasticities, and assessment of their general plausibility.

### 2.7.3 TUBA tests

The 2019 reports that: *“TUBA runs using the VTT model show that changes in benefits in the tests with low, medium and high multipliers are proportionate to change in multiplier values. The tests suggest that TUBA runs with a CVTT multiplier would result in higher benefits compared to the TAG VTT when the multiplier values are higher than two”*.

They continue: *“TUBA runs with medium and high CVTT used in the model and the appraisal also show an increase in benefits compared to the TUBA runs with TAG VTTs”*.

However, they qualify that: *“...when TAG VTT is used in the appraisal the medium and high multiplier models produce lower benefits than the WebTAG VTT model. This is primarily due to*

---

<sup>9</sup> i.e. across all journey purposes.

<sup>10</sup> Recall that CVTT was not considered by the VDM calibration.

*vehicles choosing longer travel distances in response to the increase in CVTT multiplier value applied”.*

We see the core conclusions of the 2019 work embodied in Tables 7 and 8 below (Tables 29 and 30 in the original document). We have annotated the tables to make clear what permutation of modelling/appraisal inputs gives rise to each estimate of time saving benefit. In particular, we highlight (in blue, following Figure 2), those permutations where the modelling and appraisal inputs are internally consistent.

From Table 7 we can see that the absolute value of the scheme is highly sensitive to the value of the CVTT and that the current TAG value produces a scheme value intermediate between Low and Medium CVTT for this particular scheme. Then, comparing Tables 7 and 8, we observe a large difference in the results; that is, the scheme value is highly sensitive to whether the CVTT multiplier is used in appraisal only, or fed back into the behavioural choices reflected in the model (highlighted in blue). This pattern of results is not surprising in the context of the case study geography. We are satisfied that the 2019 case study has demonstrated the relevance and materiality to the scheme results of how congested values of travel time are handled in the modelling and appraisal. We think this core conclusion is unlikely to be disturbed by any of the comments we have made above.

Table 7: TUBA benefits with TAG VTT in modelling – reproduced from WSP et al. (2019)

Benefit	WebTAG VTT	with Multiplier VTT							
		Free-Flow	Low CVTT		Medium CVTT		High CVTT		
			Delay	Total	Delay	Total	Delay	Total	
Time Saving Benefit									
Commuting	9,966	1,308	5,635	6,944	16,906	18,215	28,177	29,486	
Other	7,034	481	3,041	3,523	9,124	9,605	15,206	15,687	
Business (Car)	10,991	1,385	4,899	6,285	14,698	16,083	24,497	25,882	
Business (Freight)	6,674	1,200	2,615	3,816	7,846	9,047	13,077	14,278	
<b>Total</b>	<b>34,664</b>	<b>4,376</b>	<b>16,191</b>	<b>20,567</b>	<b>48,574</b>	<b>52,950</b>	<b>80,957</b>	<b>85,333</b>	
Fuel VOC	-3,183	NA	NA	-3,183	NA	-3,183	NA	-3,183	
Non Fuel VOC	-2,414	NA	NA	-2,414	NA	-2,414	NA	-2,414	
User Charges	1,039	NA	NA	1,039	NA	1,039	NA	1,039	
Operator Revenues	-387	NA	NA	-387	NA	-387	NA	-387	
Greenhouse Gas	-1,437	NA	NA	-1,437	NA	-1,437	NA	-1,437	
Indirect Tax Revenues	4,325	NA	NA	4,325	NA	4,325	NA	4,325	
<b>Total</b>	<b>32,608</b>	<b>4,376</b>	<b>16,191</b>	<b>18,511</b>	<b>48,574</b>	<b>50,894</b>	<b>80,957</b>	<b>83,277</b>	
Change from WebTAG	1.00			0.59		1.53		2.46	
Change from Low CVTT			1.00		3.00		5.00		

TAG model,  
TAG appraisal

TAG model,  
M=1 appraisal

TAG model,  
M=3 appraisal

TAG model,  
M=5 appraisal

Table 8: TUBA Benefits with medium and high multiplier CVTT models – reproduced from WSP et al. (2019)

Benefit	CVTT with Medium Multiplier in Modelling				CVTT with High Multiplier in Modelling			
	WebTAG	Medium Multiplier CVTT			WebTAG	High Multiplier CVTT		
		Total	FreeFlow	Delay		Total	FreeFlow	Delay
Time Saving Benefit								
Commuting	9,262	16,510	1,425	15,085	8,256	21,853	1,729	20,124
Other	7,119	9,161	767	8,394	6,519	11,509	1,203	10,306
Business (Car)	10,147	13,472	1,967	11,505	10,750	17,440	3,324	14,117
Business (Freight)	5,696	7,483	1,143	6,340	4,827	9,332	1,117	8,216
<b>Total</b>	<b>32,224</b>	<b>46,626</b>	<b>5,302</b>	<b>41,324</b>	<b>30,352</b>	<b>60,134</b>	<b>7,371</b>	<b>52,763</b>
Fuel VOC	-3,696	-3,696	NA	NA	-3,498	-3,498	NA	NA
Non Fuel VOC	-3,279	-3,279	NA	NA	-4,983	-4,983	NA	NA
User Charges	1,854	1,854	NA	NA	2,379	2,379	NA	NA
Operator Revenues	-776	-776	NA	NA	-189	-189	NA	NA
Greenhouse Gas	-1,644	-1,644	NA	NA	-2,108	-2,108	NA	NA
Indirect Tax Revenues	4,974	4,974	NA	NA	6,417	6,417	NA	NA
<b>Total</b>	<b>29,657</b>	<b>44,059</b>	<b>5,302</b>	<b>41,324</b>	<b>28,370</b>	<b>58,152</b>	<b>7,371</b>	<b>52,763</b>

M=3 model,  
TAG appraisal

M=3 model,  
M=3 appraisal

M=5 model,  
TAG appraisal

M=5 model,  
M=5 appraisal



## 2.8 Synthesis

Our appreciation of the 2019 study is that what has been done is very good in its own terms, and taken together with the 2018 study, provides an excellent review of the status quo. Aside from the misinterpretation of the  $r$  parameters from 2014/5, we have not detected any obvious technical flaws within the 2019 work.

That said, it should be highlighted that – primarily because of budget constraints – the 2019 study deployed two substantive simplifications to their analysis, and it is therefore appropriate to comment on the likely implications of these simplifications. First, the prior or ‘pre-estimation’ BY matrix formed the basis of the PRISM tests – when ideally the BY matrix would have been re-estimated to take account of CVTT. Second, the VDM component of PRISM was not re-estimated on the prior BY matrix, and instead the pre-existing VDM estimated on the post-estimation BY matrix (not accounting for CVTT) was used – ideally the VDM would have been re-estimated on an updated BY matrix accounting for CVTT. Whilst both simplifications do not undermine the essence of the 2019 tests, intuition suggests that if both aspects of re-estimation had been conducted, then some degree of mitigation to congestion would already have been observed in the DM – primarily in terms of trip re-routing but also in terms of trip re-timing. On this basis, it is possible that the TUBA tests summarised in section 2.7.3 overstate the decongestion benefits of the scheme – by how much is impossible to say.

One way of building confidence in these simplifications would have been to examine the demand elasticities implied by the FY forecasts – in particular, comparing these against the elasticities estimated in the course of the VDM calibration.

The principal contribution of the 2019 study is to demonstrate that CVTT can be successfully implemented within PRISM. The following more specific findings emerge from the three forms of testing:

- BY tests: the introduction of CVTT may affect the validation results, and could make the difference between achieving TAG targets and not achieving them.
- FY tests: total vehicle kms and average car trip lengths tend to increase with the multiplier, albeit by relatively modest amounts. The impact on routeing appears negligible in many cases, but can have an effect when there is a choice between short distance/high congestion and long distance/low congestion.
- TUBA tests: reading across Tables 7 and 8, the user benefits of the case study scheme are highly sensitive to the choice of multiplier value (i.e. the size of the trapezium varies considerably with  $M$ ); then comparing Tables 7 and 8 for particular values of  $M$ , there are large differences between the appraisal-only user benefits and the modelling and appraisal user benefits.

So, to summarise, our appreciation of the position the Department has currently reached on this topic is as follows. Its consultants have demonstrated through the case study that it is technically feasible to represent CVTT in modelling and appraisal. There is no technical showstopper. The issues are (a) whether there is enough behavioural evidence to support a



value of M different from unity; (b) whether M multipliers should be deployed in appraisal only or throughout modelling and appraisal; and (c) what testing regime would be required to enable the Department to move forward from case study to implementation in guidance with confidence.

### **3 HOW WOULD THE DEPARTMENT IMPLEMENT CVTT INTO ITS APPRAISAL AND WHAT ARE THE BARRIERS TO THIS?**

Our response to this question can be broken down into a series of sub-questions.

#### **3.1 Is the evidence on the congestion multiplier sufficiently secure and robust to meet the standards for use in TAG?**

We think that WSP et al.'s (2018) review of this question is fair. There is strong evidence both from the 2014/5 VTT study and from meta-analysis evidence that people place a higher disutility on driving in congested conditions than in FF conditions. This is for a mixture of reasons: the higher cognitive load associated with the number of driving decisions per minute, higher annoyance associated with driver behaviour, pushing in etc., and increased uncertainty about arrival time at the destination, which is particularly a feature if the journey is not made regularly and delays are not fully anticipated. But we do not know the composition of that mixture, nor what is perceived and taken into account in advance and what is unperceived/unexpected.

We agree that some features of the 2014/5 study may have combined to produce a rather high value for the multiplier for 'Other' journey purpose. The pictorial representation was of conditions approaching gridlock and the mean journey length in the sample was high. While this was controlled for in deriving the recommended values of time, we agree with the conjecture that the congestion multiplier might be correlated with journey length. WSP et al. suggested a piece of analysis of variance work to understand better the pattern of relationships around the congestion multiplier in the data, and we support this proposal.

Having said the above, we do wonder whether, in the course of the WSP et al. 2018 and 2019 studies, the policy question has been scoped as broadly as it might have been. In the 2014/5 study, the Arup et al. team estimated VTT using three different SP games (SP1: time vs. cost; SP2: time vs. cost vs. reliability; SP3: time vs. cost vs. crowding/congestion), and then considered the respective merits of each game – from both conceptual and empirical perspectives – as the basis for the headline VTT (representative of Average traffic conditions) to be implemented in TAG. On balance, Arup et al. recommended SP1 for the headline VTT, but this should be qualified by noting recommendations 4 and 5 from the 2014/5 study:

*“R4: In the immediate term, we would recommend the values from SP1 as the basis for the 'headline' VTT, since these provide the closest comparator to the 2003 game, and most readily lend themselves to implementation in appraisal. It should be clarified that we interpret VTT from SP1 as referring to 'average' travel conditions, rather than free-flow or uncrowded conditions.*

*If however crowding/congestion data at an appropriate level of detail can be sourced, then there is a case for basing 'headline' VTT on appropriately weighted values from SP3 – instead of SP1.*

*R5: We recommend that the Department should undertake further work to examine the viability of using SP3, and its relative advantages/disadvantages against SP1”.*

In other words, if R4 and R5 are accepted, then the Department should consider the relative merits of two alternative approaches for implementing congestion multipliers in TAG, namely:

- Approach 1: derive multipliers from SP3, but implement these as multiples of the headline VTT from SP1, or;
- Approach 2: derive both the headline VTT and congestion multipliers from SP3<sup>11</sup>.

In effect, WSP et al. considered Approach 1, but not 2.

On pages 221-3 of the 2014/5 study report, there is discussion of the relative merits of SP1-3 as alternative bases for the headline VTT – a starting point should be to review this debate.

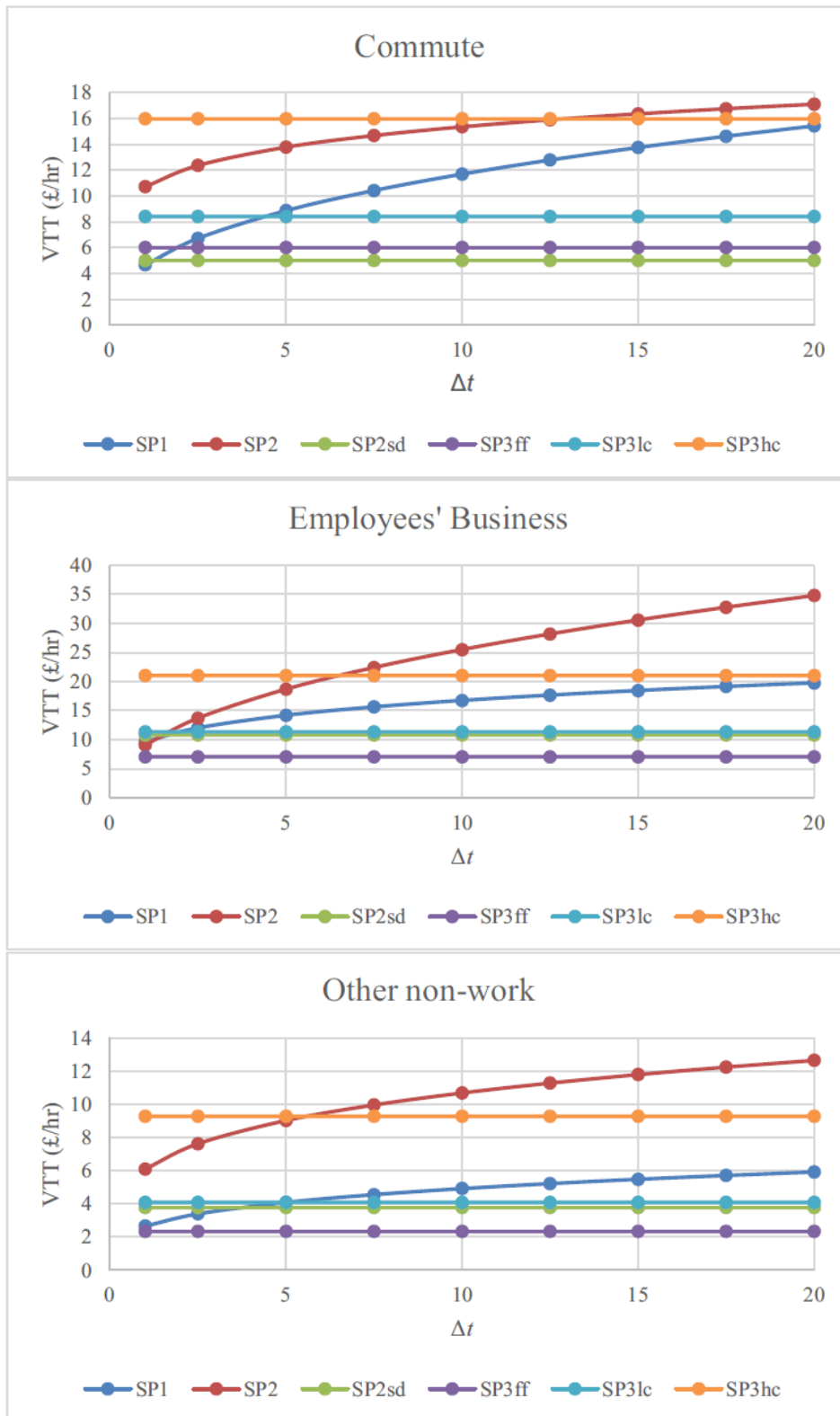
Subsequently, the modelling team from 2014/5 has undertaken some further analysis of SP1-3, and this is reported in the academic literature as Hess et al. (2017). This paper highlights the following key points:

- Whereas the VTTs emanating from SP1 and SP2 exhibit a significant ‘size effect’, this property is not discernible in SP3. The practical implication is that, when deriving VTT from SP1 or SP2, an assumption has to be made regarding the ‘size effect’, i.e. the change in travel time from the reference journey which conditions the VTT. This property is shown in the figure below from Hess et al. (2017), where VTT is plotted against the assumed ‘size effect’ – denoted  $\Delta t$ . It can be seen that, in the case of SP1 and SP2, the VTT increases, but at a decreasing rate, with  $\Delta t$ . By contrast, in the case of SP3, VTT is constant with  $\Delta t$ . Since  $\Delta t$  must be assumed and cannot be empirically identified, there is an argument that the VTT emanating from SP3 is more definitive than that emanating from SP1 or SP2.
- When following Approach 1, much is made of the relatively high congestion multiplier for OtherNW (especially in the case of Heavy Congestion), but one has to remember that the SP1 VTT for OtherNW is by accepted standards quite low. Indeed, in 2014/5, Commute values went up, but Other NW went down – relative to previous TAG values. Therefore, we need to be careful in dismissing the congestion multiplier for OtherNW as an outlier – since an alternative explanation could be that the SP1 VTT for OtherNW is itself too low.

---

<sup>11</sup> More specifically, in Approach 2, the headline VTT would be derived as a weighted average as per (Eq.4), and the  $r$  multiplier would be re-defined in relation to this weighted average from SP3 as opposed to the headline VTT from SP1. The definition of the  $M$  multiplier would be unaffected and common to Approaches 1 and 2.

Figure 4: Plot showing size effects on VTT by SP game and journey purpose – reproduced from Hess et al. (2017)



- As noted earlier in this report, congestion is likely to be especially prevalent in Commute journeys. An interesting feature of the Hess et al. figure for Commute is

that the VTTs emanating from SP2 and SP3hc are equivalent at  $\Delta t=12.5$ . Thereafter, SP2 is above SP3hc, but not dramatically so. Whilst we cannot be at all definitive, one possible explanation for this finding could be that SP2 (reflecting reliability) and SP3hc (reflecting heavy congestion, leading to delay) are capturing similar effects.

If we were starting with a blank sheet of paper and choosing a value of the congestion multiplier based on the evidence from the 2014/5 study and broader meta-analysis and international evidence, we would in principle favour differentiating between the value of congested time (i.e. conflating light/heavy congestion etc.) and the value of free flow time – such that  $M>1$ . We think the evidence, pretty strong although not conclusive, supports the proposition that congested time attracts a higher disutility than FF time. If we had to choose a single ‘best’ value of  $M$ , we would recommend an  $M$  value of 2 as the initial position for any potential update to TAG. This is on the basis that a  $M$  value of 2 is consistent with most of the cell values in Table 2 above and with the meta-analysis results range.

We observe from Figure 4 above that, whereas in WSP et al. (2019) the value of delay time is taken to be constant (e.g.  $M=3$ ), the behavioural evidence is that the value of travel time varies with the intensity of congestion itself, for example in the ratio 1 : 1.5 : 2 for Commuting and 1 : 2 : 4 for OtherNW in going from FF to Light Congestion to Heavy Congestion conditions. We find this quite plausible. So, if the Department is minded to move forward with CVTT, we would recommend working on the feasibility of adopting Approach 2 (i.e. based on 2014/5 SP3) rather than Approach 1. This would include revisiting the implications for modelling in the context of the value set in Approach 2, including the possibility of representing the value of delay time in modelling as a non-linear function of congested travel time.

### **3.2 Should congestion multipliers be implemented in modelling and appraisal or could they be implemented in appraisal only?**

Our answer is partly one of principle and partly one of practice. We see the congestion multiplier primarily as a reformulation of generalised cost which both drives demand and route choice in the model and is core to the user benefit calculations in TUBA. There are very strong arguments for consistency of the generalised costs in modelling and appraisal. Furthermore, perusal of Tables 7 and 8 above (reproduced from WSP et al. (2019)) shows that the differences between applying the multiplier in appraisal only and throughout are large and material to the BCR. This is because congestion multipliers cause heightened feedback both in routing and demand which moderate the first round benefit substantially in the conditions of the case study.

Therefore, we think it is all or nothing; if congestion multipliers are to be used, they should be used throughout the modelling and appraisal to provide the core user benefit estimates at Level 1. We appreciate that the DfT’s guidance and use of average values of time across congested and free flow conditions creates potential distortions in the relative values of different types of scheme. But, given the arguments of principle and materiality in the paragraph above, we do not see any viable temporary halfway house, for example introducing a decongestion benefit in appraisal only at Level 2.

As noted above, WSP et al. were faced with the problem of creating a continuous variable – essentially a composite of FF time and Delay time – which could readily be derived from standard model outputs for both links and junctions and to which a weight consistent with a given M value could be applied. Their solution of translating FF and Delay into generalised cost is simple but elegant. However, if congested values are thought to be better represented behaviourally as a continuous variable, further work would be required to obtain such values and then marry them up with FF time and Delay time outputs of the model.

The WSP et al. case study is inherently limited to the conditions of the scheme and context. This is a regional study area with a large scheme and many routeing and demand choices. The simulation work has been done for 2036, and could be done for 2026, but we understand that in some cases models are also run for 2051. The practical barriers therefore are: whether a broader range of testing (scheme types, area types, heavily congested urban conditions, more modelling and appraisal years) is required, whether the method works equally well for other models (especially SATURN/DIADEM), and whether the consequences for scheme ranking need to be established, since it is clear that scheme BCR relativities will change and the absolute BCR of ‘mainly FF’ type schemes may fall (given that the SP1 values implemented in TAG are intended to represent Average congestion).

### **3.3 Overlaps and biases**

Clearly, the main overlap is with reliability, since as well as delay, the uncertainty induced by congestion is probably one of the principal reasons why people do not like it. Therefore, there is a problem about including both congestion multipliers and (un)reliability in modelling and appraisal. One possible approach to ensuring there is no confounding between CVTT and reliability might be for the congestion multiplier to stand as a proxy for reliability for the range of the distribution classed as ‘normal conditions’, reducing the scope of reliability benefits essentially to the assessment of resilience measures affecting the right hand tail of the journey time distribution, perhaps roadworks, incident-related delays, closures due to extreme weather events etc. not captured in modelling ‘normal conditions’.

Would introducing congestion multipliers be biased? We think not between modes, on the basis that this is not dissimilar from the treatment of walking, waiting, crowding etc. in public transport appraisal. This is another step on the road to incorporating the quality of travel time, including its productive use, as well as the quantity of travel time in appraisal work.

Obviously taking this step would have broader implications, for example for the appraisal of schemes in the context of autonomous vehicles where an increase in the quality of in-vehicle time and consequently a drop in the unit value of travel time is to be anticipated. Issues relating to the assumptions surrounding the timing of such technological changes for the appraisal of current schemes in future years will become more transparent. The sensitivity of scheme results to policy assumptions/scenarios in this area will be enhanced. It would be biased to represent one source of the quality of travel time (congestion conditions) in appraisal but not others.

Although not a bias, we do expect that changes in the ranking order and in the Value for Money (VfM) category of individual road schemes will be politically sensitive in a broad sense, especially if the direction of change clashes with wider political imperatives.

### 3.4 Other issues

Our sense is that introducing congestion multipliers may make appraisal results even more sensitive to study areas and what routeing, departure time, and destination choices are allowed to vary in the model. This may have implications for TAG.

In principle, we think the arguments deployed for car users apply equally to light and heavy commercial vehicle drivers. As long as their time values remain based on the Cost Saving Approach (CSA), we do not think it is logical to introduce congestion multipliers for these user classes. If the basis of their time values were to change to a behavioural or willingness to pay approach, then that would change.

It strikes us that the challenges facing the Department in relation to the incorporation of congestion multipliers in the modelling and appraisal of road have direct parallels in relation to the incorporation of lateness multipliers for modelling and appraisal of rail. As far back as ten years ago, there was recognition in rail demand modelling (Batley, Dargay & Wardman, 2010), that the relatively high lateness (aka reliability) multipliers estimated in SP studies were inconsistent with the relatively modest lateness elasticities estimated in econometric studies. However, it is only very recently that official guidance and procedures (especially the Passenger Demand Forecasting Handbook (PDFH) and Schedules 4 and 8 of the regulatory regimes governing track access) have sought to decouple the two sources of evidence.

The reasons for this divergence between SP and econometric evidence is not fully understood, but Batley et al. (2010) reasoned that: *“To date, forecasts of the demand impacts of lateness and reliability have been derived largely from individual-level models taken at a snapshot in time. The contribution of this paper is to develop a dynamic model of rail demand at the market-level, yielding short and long-run elasticities with respect to lateness. Whereas individual-level models have suggested a high valuation of lateness and reliability, our market-level models indicate a relatively muted demand response. Reconciling these findings, we reason that, whilst rail travellers show considerable disdain for experiences of lateness, such experiences will not necessarily dissuade them from travelling by train”.*

It is not inconceivable that a similar phenomenon could apply to road congestion, and this therefore lends additional support to our recommendation that, at least in the short term, lower (i.e.  $M = 2$ ) rather than higher ( $M = 3$  to  $5$ ) bound estimates of CVTT should be the focus of potential implementation in TAG.

### 3.5 Conclusions

Overall, our answer to question 2 is as follows:

- Differential values of time according to congestion conditions are a real phenomenon.
- There is a strong conceptual case for implementing them in modelling and appraisal but there are challenges in doing so.

- There is no easy or quick win such as implementing a decongestion benefit in appraisal only at Level 2.

A key policy issue is whether to view the CVTT question as a standalone or in the broader context of bringing the treatment of the quality and productive use of travel time more centrally into the appraisal system.

In coming to these conclusions, we have been mindful of the Principles of Good Guidance provided to us by the Department. Section 4 sets out the steps which we think the Department would need to follow in order to satisfy these Principles.



## 4 WHAT ARE THE FUTURE RESEARCH NEEDS FOR CVTT?

We think that the first, immediate steps required are policy decisions into which this Forward Look could play. Weighing up the points we and WSP et al. have made against the Principles of Good Guidance, is there a sound basis for moving forward at all and if so, how? A key high level question is whether to consider a move to VTT varying according to congestion conditions for all schemes or whether to retain the current headline VTT (representative of Average traffic conditions) for most schemes but permit the use of CVTT for a subset of 'exceptional' schemes subject to specified criteria.

Then, following from the above discussion, we would distinguish between desk studies which could help to move the debate forward incrementally and major pieces of work.

In the former category, we recommend:

- A small piece of re-analysis of 2014/5 SP1 and SP3 to examine the relationships between CVTT values, journey purpose and journey length.
- Further work to compare the relative merits of Approaches 1 and 2 in deriving:
  - Some notion of a headline VTT representing Average traffic conditions<sup>12</sup>
  - CVTTs for both FF and Delay based on a single M value of 2
  - CVTT as a continuous function of congested travel time
- A small piece of re-analysis of WSP et al. (2019) to examine the implied elasticities arising from the FY forecasts.
- A review of work elsewhere on continuous values of CVTT together with consideration of how such values could be used to map on to model outputs such as FF time and congested time.

The above pieces of work are seen as the precursors to a decision on whether a new SP study would be required in order to develop a robust continuous function relating values of travel time to congestion intensity. Whether or not this is required, a major piece of work, probably crucial to implementation is:

- To prove the feasibility of the method on a typology of schemes, using a range of models and looking at scheme years further into the future such as 2051. We think this step, including discussion of the results with stakeholders, is likely to be the most critical one. A pre-requisite would be adaptation of the models so as to handle FF and

---

<sup>12</sup> With reference to (Eq.4), the weighted average VTT would need to be populated with data on travel times for FF and Delay conditions – and this provokes the question of whether such data should be scheme specific or national averages (with the latter perhaps segmented to reflect some notion of scheme type).

congested time to deliver the equivalent of what has been achieved with PRISM in the case study.

- Further consideration should be given to the treatment of CVTT for light and heavy commercial vehicles, which might imply revisiting the use of the Cost Saving Approach (CSA) for those vehicle categories.

Beyond this lies the larger question of a more holistic approach to the quality and productivity of time spent travelling and its representation in modelling and appraisal. If it is decided that CVTT needs to be placed in this broader context, then further behavioural studies might be required, perhaps within the next National Value of Travel Time study. Last but not least, it should be acknowledged that the whole approach of adopting congestion multipliers implies confidence in the baseline VTT from which multiples are derived – and arguably that baseline should be defined in terms of FF rather than Average traffic conditions. Therefore, the ITT for any future National Value of Travel Time study should be explicit regarding the traffic conditions (i.e. FF vs. Average) underpinning the baseline VTT.

## 5. CONCLUSIONS

*Robust Mandate – is there sufficient rationale for the change?*

Our answer is a qualified yes. The evidence base is reasonably strong and the WSP et al. studies have demonstrated feasibility at an outline level. We think the case for implementation as a core (Level One) user benefit throughout modelling and appraisal rather than in appraisal only is pretty clear. The rationale would be stronger if CVTT were viewed as a step on the road to a more holistic treatment of the quality and productive use of travel time in modelling and appraisal.

*Proportionate changes in modelling and appraisal costs*

This would need to be explored further, but we take WSP et al.'s work to be saying there are no insuperable obstacles here and the likely costs will be upfront changes to code. Issues relating to model validation have been flagged, however.

*Appropriate consultation on the change – has the impact on the Transport Business Case been calculated for a range of projects and have relevant stakeholders been consulted?*

Clearly not yet, and a big decision for the Department is whether to proceed to this step which involves application to a typology of schemes, a range of models, values over scheme lives, and engagement building with a range of internal and external stakeholders.

*Improving analytical assurance—presentation, complexity, innovation, risk of error/uncertainty*

Relative to many changes in TAG, this has the potential to be large and to require public explanation and defence. It will be necessary to work through how this plays out in the context of processes such as audit and Public Inquiry where concepts such as weighted minutes will come under scrutiny.

*Two concluding thoughts*

Viewed as a standalone issue, we think there is a programme of work which could lead to an implementation plan for CVTT in the medium term. There is no immediate term fix.

But really, we see CVTT as part of a bigger question which has policy ramifications. In particular, once the Department accepts that the 'quality' of time as well as the 'quantity' is important for modelling and appraisal, this does open the box of how to predict changes in the 'quality' of time over the life of current schemes associated with technical change.

## REFERENCES

Arup, ITS Leeds and Accent (2015) Provision of market research for value of travel time savings and reliability: phase 2 report. Prepared for the Department for Transport.

Batley, R., Dargay, J. and Wardman, M. (2010) The impact of lateness and reliability on passenger rail demand. *Transportation Research Part E: Logistics and Transportation Review*, 47 (1), pp61-72.

WSP, Rand Europe and Mott Macdonald (2018) Congestion dependent values of time in transport modelling. Report to the Department for Transport.

WSP, Rand Europe and Mott Macdonald (2019) Congestion dependent values of time in transport modelling. Report to the Department for Transport.

Hess, S., Daly, A., Dekker, T., Ojeda Cabral, M. & Batley, R. (2017) A framework for capturing heterogeneity, heteroskedasticity, non-linearity, reference dependence and design artefacts in value of time research. *Transportation Research Part B: Methodological*, 96, pp126-149.

Mott Macdonald (2019) PRISM 5.2, Model validation report, September 2019.