Article:
Watling, D.P. (2006) User equilibrium traffic network assignment with stochastic travel times and late arrival penalty. European Journal of Operational Research, 175 (3). pp. 1539-1556. ISSN: 0377-2217

https://doi.org/10.1016/j.ejor.2005.02.039

USER EQUILIBRIUM TRAFFIC NETWORK ASSIGNMENT WITH STOCHASTIC TRAVEL TIMES AND LATE ARRIVAL PENALTY

David Watling, Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, UK
email: dwatling@its.leeds.ac.uk

Abstract – *The classical Wardrop user equilibrium (UE) assignment model assumes traveller choices are based on fixed, known travel times, yet these times are known to be rather variable between trips, both within and between days; typically, then, only mean travel times are represented. Classical stochastic user equilibrium (SUE) methods allow the mean travel times to be differentially perceived across the population, yet in a conventional application neither the UE or SUE approach recognises the travel times to be inherently variable. That is to say, there is no recognition that drivers risk arriving late at their destinations, and that this risk may vary across different paths of the network and according to the arrival time flexibility of the traveller. Recent work on incorporating risky elements into the choice process is seen either to neglect the link to the arrival constraints of the traveller, or to apply only to restricted problems with parallel alternatives and inflexible travel time distributions. In the paper, an alternative approach is described based on the 'schedule delay' paradigm, penalising late arrival under fixed departure times. The approach allows flexible travel time densities, which can be fitted to actual surveillance data, to be incorporated. A generalised formulation of UE is proposed, termed a Late Arrival Penalised UE (LAPUE). Conditions for the existence and uniqueness of LAPUE solutions are considered, as well as methods for their computation. Two specific travel time models are then considered, one based on multivariate Normal arc travel times, and an extended model to represent arc incidents, based on mixture distributions of multivariate Normals. Several illustrative examples are used to examine the sensitivity of LAPUE solutions to various input parameters, and in particular its comparison with UE predictions. Finally, paths for further research are discussed, including the extension of the model to include elements such as distributed arrival time constraints and penalties.*

1. Introduction and Review

Equilibrium traffic assignment models are able to reflect drivers' path choice behaviour through a range of devices. The basic static, deterministic Wardrop User Equilibrium (UE) model explains the congestion-feedback effect, with the premise that drivers are able perfectly to predict the generalized travel times they will experience. The dynamic UE model (Ran & Boyce, 1996) takes this level of perfect predictive ability a stage further, assuming that drivers are able to perceive actual arc travel times differentiated by entry time to the arc. The Stochastic User Equilibrium (SUE) model (Daganzo & Sheffi, 1977) and Dual-Criteria UE model (Leurent, 1998) provide counterparts to the basic static/dynamic UE models, which include randomly-distributed elements in the drivers' perceptions of generalized travel times. This is intended to reflect the fact that as modellers, observing the transport system, we cannot hope to explain or observe all the factors that motivate path choice (typically only using readily-observed proxies for the explained elements, such as travel time, distance and tolls), and that these factors are likely to vary in importance across the travelling population.

A comment that may be made regarding all such tools is that they are based very much on how the transport planner views the whole transport system. In particular, they allow a reflection of (at least some elements of) the great deal of observational uncertainty faced by modellers of the transport system, yet provide no hypothesis as to how the drivers in the network might consider uncertainty. This latter point is the focus of the present paper. In the context of traffic assignment, where drivers are selecting a path from a fixed origin to destination, the primary source of uncertainty for the individual driver is surely the fact that for any particular trip, the driver will not know in advance the precise traffic conditions he/she would encounter on the alternative paths available. While this will partly be attributable to a lack of spatial knowledge (an individual driver will only have experience of a limited number of alternatives), there will be a good deal of uncertainty even for those paths of which the driver has considerable experience. This uncertainty derives from the ambient trip-to-trip variation in travel times, both within and between days, due to factors such as incidents, breakdowns, weather and the variability in activity patterns. These factors lead to variations in, among other things, flows and capacities, which in turn impact on travel times.

At one level, empirical work is clearly needed to address how drivers respond to such variability in particular case-study situations (see, for example, Abdel-Aty *et al*, 1997; De Palma & Rochat, 1999). However, in order to have practical value, such empirical work needs appropriate hypothetical models with adjustable parameters, in order that the behavioural information may be exploited, as well as a link from the behavioural data to applications in a potentially large-scale network assignment model with heterogeneous origin-destination movements. The purpose of the present paper is the deduction of such hypothetical models.

Research towards the development of such techniques has particular relevance for a number of reasons. Firstly, in the emerging topic of network reliability (Asakura, 1996; Du & Nicholson, 1997; Chen *et al*, 1999; Bell & Cassir, 2000), where one can view the traffic system as unreliable from an objective perspective, it seems inconsistent to assume that the drivers themselves have no view on such unreliability. Therefore, the techniques developed in this paper may be viewed as a step on the way towards a comprehensive model for network reliability assessment. Secondly, there has been a growing appreciation of the 'value' of unreliability (Hall, 1993; Porter *et al*, 1996; DETR, 2000), and the potential advantages of understanding, and thereby mitigating against, the impacts of travel time variability. Thirdly, the emergence of traveller information systems has provided the technology to make a significant impact on the level of drivers' network knowledge, yet the full potential of such systems can only be realised if their impact on behaviour can be understood and forecast (Mannering *et al*, 1994; Emmerink *et al*, 1995).

The purpose of the present paper is therefore to generalise existing equilibrium assignment approaches to accommodate travellers' reactions to variability in traffic conditions. As motivation, let us consider the simple network example defined below.

Example 1 A network serves a single origin-destination demand of q units through two parallel arcs/paths. The arc travel time functions on each arc a (a = 1,2) are functionally identical, i.e. the travel time function for arc a is $t_a(v_a) = \eta(v_a)$, where η is an arc-independent function assumed to be monotonically increasing, and where v_a denotes the flow on arc a. Drivers are assumed to perceive travel cost purely in terms of travel time. Suppose additionally that there are variations in the actual travel times on arc 1, but not on arc 2, such that the distribution of travel times at a given flow v_1 on arc 1 has mean $t_1(v_1)$ and variance σ_1^2 (for simplicity assumed flow-independent).

How might a conventional traffic assignment model deal with this example? If it is assumed, as is usual, that the population of drivers is a homogeneous group of risk-neutral, fully-informed travellers, then it would be natural to assume that drivers make their choice decision based on the *expected* travel costs (which equal expected travel times, if cost and time are perceived to be the same). Therefore, σ_1^2 would have no role, and in fact the split of demand between the two arcs/paths would be 50%/50% regardless of the value of σ_1^2, and regardless of whether a UE or SUE model is adopted.

This independence from σ_1^2 derives from a number of assumptions underlying conventional applications of traffic assignment models, be they UE or SUE. The first key assumption is the premise that all drivers are risk-neutral, whereas in reality there are likely to be a variety of drivers, risk-averse, risk-prone and risk-neutral (Yin & Ieda, 2001). A risk-prone driver is likely to gamble on being 'lucky', and would therefore see the variance σ_1^2 as a positive attribute of arc 1, providing him/her with the chance of getting a lower travel time than arc 2. The risk-averse driver would look at the variance σ_1^2 in a negative way[1], concerned that they might be one of the 'unlucky' ones who experiences a higher travel time than that on arc 2. If there are more risk-prone than risk-averse drivers in the population, and the travel time density on arc 1 is symmetric, then one would expect in aggregate arc 1 to be used more than arc 2, and *vice versa* if risk-averse drivers out-numbered risk-prone ones. The degree of asymmetry of the travel time density function would be a further confounding factor.

A second key assumption is that variability itself is not valued as an inconvenience (Noland *et al*, 1998). The issue here is that while individuals may have flexibility in re-arranging the arrival and departure times of their trips and associated activities, that all other things being equal they prefer not to incur the inconvenience of such re-scheduling. Therefore, they would tend to avoid the risk of having to do this wherever possible. For example, it may well be possible to bring forward or delay a meeting in response to travel conditions on the journey to work, but such re-arranging would have a nuisance value that might be avoided.

A third key assumption is that while being risk-neutral, drivers are also neutral to the time at which they arrive at their destination. In practice, it seem logical that, while any kind of perception of 'lateness' is likely to be valued negatively, the degree of negativity will

[1] The response of such a group may also be a reluctance to experiment with alternatives, a habit/inertia effect.

depend partly upon the flexibility of the activity at the end-point of the trip (Noland *et al*, 1998; Noland, 1999).

In the literature, a number of works can be found that address the issue of travel time variability in a traffic assignment context. The 'rational expectations' approach of Arnott *et al* (1991) was developed for modelling driver information systems, in which uninformed drivers perceive expected, day-averaged costs and informed drivers day-specific costs. Emmerink *et al* (1995) developed this method further to include some kind of proxy to risk aversion, by supposing that utility was a linear combination of the travel time mean and standard deviation. In a similar vein, Van Berkum & Van der Mede (1999) defined utility as a combination of the travel time mean and variance. Mirchandani & Soroush (1987) and Watling (2002) developed generalised characterisations of equilibrium for cases in which travel times vary stochastically, yet both assumed risk-neutral decision-making.

More recently, Yin & Ieda (2001) proposed a three user-class model, to reflect different attitudes to risk, in combination with a network assignment model with stochastically varying travel times. In order to reflect risk, they assumed disutility to be a user-specified, increasing function of the stochastic travel times, with drivers aiming to minimise expected disutility. Finally, Noland *et al* (1998) and Noland (1999) proposed a method that differs from Yin & Ieda's in that there is an explicit link to users' preferred arrival times, and the penalty that might be incurred by late arrival. For problems with parallel alternative paths, and either uniform or exponential travel time densities, it was shown that a simple analytic form for the expected disutility could be derived.

The approach developed in the present paper draws on aspects of the references in the latter paragraph. Like Yin & Ieda (2001), the aim is to develop a general network model based on arc travel time densities (which infer *path* travel time densities). However, following Noland *et al*, the utility function is developed from the user's arrival time at the destination relative to some preferred schedule. In addition, the aim is to develop an approach that accommodates a flexible range of travel time densities, that are more plausible than the uniform/exponential, and incorporate both distributional asymmetries and correlation between arcs.

The paper is structured as follows. Section 2 is concerned with the basic notation and model formulation, including a consideration of issues of existence, uniqueness and computation of the extended equilibrium model proposed. In section 3, specific examples are presented of the general travel time density considered in section 2, and the appropriate functional forms deduced. Section 4 consists of numerical results from a number of illustrative examples, with conclusions and areas for further research in section 5.

2. NOTATION AND MODEL FORMULATION

The transportation network is considered as a directed graph consisting of A arcs indexed $a = 1,2,...,A$. The arc travel times are represented as stochastic variables, to represent the actual trip-to-trip variability in travel times. Thus, the actual travel time on arc a is denoted T_a. The *mean* arc travel times are assumed themselves to be functions of the arc flows. Thus, if v_a denotes the flow on arc a $(a = 1,2,...,A)$, with $\mathbf{v}$ the A-vector of flows across all arcs, then the *mean* travel time on arc a is denoted $t_a(\mathbf{v})$. When all such functions are collected together across arc into an A-vector, the resulting vector function is denoted $\mathbf{t}(\mathbf{v})$. It is supposed then the joint density of arc travel times $\mathbf{T} = \{T_1, T_2, ..., T_A\}$ has a known distributional form which is parameterised by the arc flow vector $\mathbf{v}$, through $\mathbf{t}(\mathbf{v})$.[2]

To the arcs of the network are joined W origin-destination (O-D) movements indexed $k = 1,2,...,W$, each with a demand of q_k. It is assumed that the drivers that make up this demand make their path choice decision with a longest possible travel time in mind for their journey, which if exceeded would incur them some inconvenience. Since in the proposed model drivers have an already-fixed departure time, their perception of such a longest travel time implies a latest *acceptable arrival time* at their destination *given* their choice of departure time.

Therefore, while the interest here is in incorporating a penalty for late arrival, we may equivalently incorporate a penalty for the travel time being longer than some acceptable level. For each O-D movement k, there is assumed to be a single, constant longest acceptable travel time τ_k $(k = 1,2,...,W)$. It is noted that by division of the actual O-D movements into a larger number of virtual O-D movements representing 'user classes' by movement, this approach can be generalised to reflect (say), differing arrival time flexibilities by trip purpose. Since this adds nothing to the problem mathematically, and only serves to make the notation more complex, such user classes are not explicitly

[2]The joint density of $\mathbf{T}$ may be parameterised by $\mathbf{v}$ in a very general sense, rather than $\mathbf{v}$ just parameterising the mean. See, for example, Watling (2002).

represented here, in the understanding that the proposed approach is trivially adapted to accommodate them through such 'virtual' O-D movements.

Having introduced the arcs and O-D movements, we move on to the paths of the network. The acyclic paths of the network are indexed by $r = 1,2,...,N$, with the subset of path indices relating to O-D movement k denoted R_k $(k = 1,2,...,W)$. The flow on path r is denoted f_r and the N-vector of path flows across the whole network denoted **f**. In order to be feasible for the demand, we clearly must have $\mathbf{f} \in D$ where D is the convex set:

$$D = \left\{ \mathbf{f} : \sum_{r \in R_k} f_r = q_k \quad (k = 1,2,...,W) \text{ and } f_r \geq 0 \; (r = 1,2,...,N) \right\}. \tag{1}$$

The path flows are related to the arc flows by:

$$\mathbf{v} = \mathbf{\Delta f} \tag{2}$$

where $\mathbf{\Delta}$ is an $A \times N$ arc-path incidence matrix, with elements δ_{ar} the 0/1 indicators equal to 1 only if arc a is part of path r. The convex set of demand-feasible link flow vectors is similarly denoted:

$$\widetilde{D} = \{ \mathbf{v} : \mathbf{v} = \mathbf{\Delta f} \text{ and } \mathbf{f} \in D \}.$$

The path travel times are assumed to be random variables, with the travel time C_r on path r related to the arc travel time variables by the transformation:

$$C_r = \sum_{a=1}^{A} \delta_{ar} T_a \qquad (r = 1,2,...,N) \tag{3}$$

which (if **C** denotes the random vector of path travel times) may be written in vector form:

$$\mathbf{C} = \mathbf{\Delta}' \mathbf{T}$$

where $\mathbf{\Delta}'$ denotes the transpose of the matrix $\mathbf{\Delta}$. Through such logic, the marginal distributions of the C_r, which will be needed below, are parameterised by **v** (through the distribution of **T** and its mean **t**(**v**)).

The final, important element of the approach is the specification of the path selection model of drivers. The basis of this model is the formation of a composite disutility function which, for each path in the network, is able to reflect both:

- a driver's valuation of the path's expected attributes (distance, expected travel time, tolls, etc.); and

- the extent to which following that path is likely, in the light of travel time variability, to satisfy a traveller on that O-D movement in achieving an 'acceptable' arrival time at the destination.

The approach is based on the 'schedule delay' concept proposed by Vickrey (1969). This has enjoyed some considerable use in travel choice problems in the transportation field, primarily in a deterministic travel time setting (e.g., Alfa, 1986; Ben-Akiva *et al*, 1986; Arnott *et al*, 1990; Ran *et al*, 1996; Yang & Huang, 1997), with a handful of contributions in a stochastic setting (Sumi *et al*, 1990; Noland *et al*, 1998). In this way, it is supposed that an individual considering path $r \in R_k$, and perceiving a given path travel time of c_r, may be represented as forming a composite path disutility, incorporating both the standard 'generalized travel time' ($\theta_0 d_r + \theta_1 c_r$ in (4) below) and the travel time acceptability in the form of a lateness penalty:

$$u_r = \theta_0 \, d_r + \theta_1 \, c_r + \theta_2 \max(0, c_r - \tau_k) \qquad (r \in R_k ; k = 1,2,...,W) \ . \tag{4}$$

where d_r represents the composite of attributes (such as distance) that are independent of time/flow and θ_0 is the value placed on these attributes, θ_1 is the value-of-time, and θ_2 reflects the value of being later than acceptable, Clearly, since u_r is a *dis*utility, it would generally be expected to value both variable components negatively, i.e. $\theta_1 > 0$ and $\theta_2 > 0$.

Expression (4) then may be generalised by presuming that individuals now have a view of the likely variability in travel times. Indeed, they are assumed to possess such information that they perceive (or can be represented as perceiving) the full joint probability distribution for the actual path travel time random variables. Replacing c_r with the random variable C_r in (4), and taking expectations, leads to an expected composite path disutility:

$$u_r = \theta_0 \, d_r + \theta_1 \, \mathrm{E}[C_r] + \theta_2 \, \mathrm{E}[\max(0, C_r - \tau_k)] \qquad (r \in R_k ; k = 1,2,...,W) \ . \tag{5}$$

If ψ_r denotes the marginal density function of C_r then (4) may be written:

$$u_r = \theta_0 \, d_r + \theta_1 \, \mathrm{E}[C_r] + \theta_2 \int_{\tau_k}^{\infty} (c - \tau_k) \, \psi_r(c) \, \mathrm{d}c \ . \tag{6}$$

The expected disutility represents the average disutility this particular individual would attach to path *r*. This point is emphasized since the term 'expected utility/ disutility' is in common use in transport to reflect an expectation across a population of individuals. Above, it is something rather different, namely an expectation across the range of trip-to-trip travel times that could occur, relative to the desires of a particular individual.

Based on the components described above, a characterization of equilibrium may then be specified, extending the conventional concept of Wardrop User Equilibrium (Smith, 1979):

Definition Consider the definition of expected path disutility (6), and write this as $u_r(\mathbf{f})$ to reflect a sequence of causalities: namely, the fact that the path disutility u_r is computed from the marginal density ψ_r of C_r, which depends on the path flows **f** through the link flows **v** (from (2)), which impact on the mean travel times through **t**(**v**), which in turn parameterise the joint density of **T**, from which the marginal density ψ_r is imputed through (3). Then, a path flow vector $\mathbf{f}^* \in D$ is termed a Late Arrival Penalised User Equilibrium (LAPUE), if $\mathbf{f}^*$ is a Wardrop equilibrium based on path cost functions $u_r(\mathbf{f})$. That is to say, if $\mathbf{u}(\mathbf{f})$ denotes the *N*-vector with elements the functions $u_r(\mathbf{f})$ $(r = 1,2,...,N)$, then $\mathbf{f}^* \in D$ is a LAPUE if and only if

$$\mathbf{u}(\mathbf{f}^*)'(\mathbf{f} - \mathbf{f}^*) \geq 0 \qquad \forall\, \mathbf{f} \in D \ . \tag{7}$$

As in UE models, our interest is typically not in the path flows themselves, but in the induced link flows through the application of (2) to a LAPUE path flow solution.

Existence Consider the marginal path travel time density function for path *r*, and write it as $\psi_r(c_r;\mu_r)$ to denote its (partial) parameterisation by the mean path travel time μ_r. Suppose that the functions $F_r(\mu) = \int_{\tau_k}^{\infty} (c - \tau_k)\psi_r(c;\mu)\,dc \quad (r \in R_k;\, k = 1,2,...,W)$ exist and are continuous. Suppose further that **t**(**v**) is a continuous mapping. Then LAPUE solutions exist.

Proof The proof is a minor modification to the proof of Smith (1979) to apply to the path flow, as opposed to arc flow, domain. In outline, since D is a closed, convex subset, then for any $\mathbf{g} \in D$ there exists a unique point $\mathbf{p}(\mathbf{g}) \in D$ which is nearest to **g**. A mapping may then be introduced $\theta : D \to D$ such that $\theta(\mathbf{f}) = \mathbf{p}(\mathbf{f} + \mathbf{u}(\mathbf{f}))$. As established by Smith (1979), **f** is then a UE if and only if $\theta(\mathbf{f}) = \mathbf{f}$. Then, applying Brouwer's fixed point theorem, such a solution exists if θ is a continuous mapping. But θ is continuous if **u** is, and by hypothesis **u** is a continuous composition of continuous functions (the F_r, **t** and the expectation of (3)), and so the result is established.

Uniqueness Suppose that the conditions of the Existence result above hold, and that in (6), $\theta_1 > 0$ and $\theta_2 \geq 0$. Suppose further that the arc travel time functions are strictly monotone, in the sense that $(\mathbf{t}(\mathbf{v}) - \mathbf{t}(\mathbf{w}))'(\mathbf{v} - \mathbf{w}) > 0$ for any $\mathbf{v}, \mathbf{w} \in \widetilde{D}$ $(\mathbf{v} \neq \mathbf{w})$. Finally suppose that the functions $F_r(\mu_r)$ (as defined above) are non-decreasing. Then there is a unique induced LAPUE arc flow solution.

Proof Again the proof involves a modification of that given by Smith (1979). Since the conditions of the Existence result hold, LAPUE solutions exist. Consider any two such LAPUE (path flow) solutions **f** and **g**, with distinct induced arc flows, i.e. $\mathbf{\Delta f} \neq \mathbf{\Delta g}$. Consider the function $\boldsymbol{\rho}(\mathbf{f})$ relating path flows to mean path travel times, $\boldsymbol{\rho}(\mathbf{f}) = \mathbf{\Delta}' \mathbf{t}(\mathbf{\Delta f})$. Then for such an **f** and **g**:

$$\begin{aligned}(\boldsymbol{\rho}(\mathbf{f}) - \boldsymbol{\rho}(\mathbf{g}))'(\mathbf{f} - \mathbf{g}) &= (\mathbf{\Delta}'(\mathbf{t}(\mathbf{\Delta f}) - \mathbf{t}(\mathbf{\Delta g})))'(\mathbf{f} - \mathbf{g}) \\ &= (\mathbf{t}(\mathbf{\Delta f}) - \mathbf{t}(\mathbf{\Delta g}))'\mathbf{\Delta}(\mathbf{f} - \mathbf{g}) = (\mathbf{t}(\mathbf{\Delta f}) - \mathbf{t}(\mathbf{\Delta g}))'(\mathbf{\Delta f} - \mathbf{\Delta g}) > 0\end{aligned}$$

since **t**(.) strictly monotone and by hypothesis $\mathbf{\Delta f} \neq \mathbf{\Delta g}$. Since by the assumptions of the theorem, the functions $F_r(\mu_r)$ are non-decreasing, and $\theta_1 > 0$ and $\theta_2 \geq 0$, then the path disutility function $\mathbf{u}(\mathbf{f})$ also satisfies this strict monotonicity-like condition for path flow pairs inducing distinct arc flows, being a montonically increasing, linear combination of strictly and weakly monotone-like functions:

$$(\mathbf{u}(\mathbf{f}) - \mathbf{u}(\mathbf{g}))'(\mathbf{f} - \mathbf{g}) > 0 \quad \mathbf{f}, \mathbf{g} \in D \ \ (\mathbf{\Delta f} \neq \mathbf{\Delta g}). \tag{8}$$

On the other hand, consider and expand the following expression:

$$\mathbf{u}(\mathbf{f})'(\mathbf{f} - \mathbf{g}) = \mathbf{u}(\mathbf{g})'(\mathbf{f} - \mathbf{g}) + (\mathbf{u}(\mathbf{f}) - \mathbf{u}(\mathbf{g}))'(\mathbf{f} - \mathbf{g}) > 0$$

where the >0 condition holds because the first term in the expansion is non-negative (since **g** is a LAPUE by hypothesis, using the LAPUE definition) and the second term is positive (by (8)). But since **f** is also a LAPUE by hypothesis, then applying the LAPUE definition for **f** yields $\mathbf{u}(\mathbf{f})'(\mathbf{g} - \mathbf{f}) \geq 0$, i.e. $\mathbf{u}(\mathbf{f})'(\mathbf{f} - \mathbf{g}) \leq 0$, giving a contradiction. Hence the original hypothesis, of two LAPUE solutions with distinct arc flows, must be false.

Having formulated the LAPUE model, a natural next question is that of computation of LAPUE solutions. This is not entirely straightforward, since in general the expected path disutility (6) is not expressible as a sum of arc disutilities. This has the effect that standard shortest path methods are not readily applied, and that the usual computational strategy, of avoiding explicit storage of *path* flows (as in the Frank-Wolfe algorithm, for example), is not viable[3]. Recently, several authors have developed algorithms for non-additive equilibrium problems of this general kind (Gabriel & Bernstein, 1997; Scott & Bernstein, 1998; Lo & Chen, 2000), extending previous work that had used ‘gap’ functions and path flow variables for the standard additive equilibrium problem. The approach proposed by Lo & Chen (2000) is particularly attractive, whereby recent advances in the use of ‘merit functions’ in the non-linear complementarity field are applied to obtain a problem with a *smooth* (continuously differentiable) gap function under mild conditions. This allows standard algorithms for unconstrained optimisation to be applied.

In order to apply Lo & Chen’s algorithm, the LAPUE model is reformulated as a complementarity problem ($\mathbf{x} \geq \mathbf{0}$, $\mathbf{F}(\mathbf{x}) \geq \mathbf{0}$, $\mathbf{x}'\mathbf{F}(\mathbf{x}) = 0$) by introducing the $N+W$ vectors:

$$\mathbf{x} = \begin{pmatrix} \mathbf{f} \\ \mathbf{y} \end{pmatrix} \qquad \mathbf{F}(\mathbf{x}) = \begin{pmatrix} \mathbf{F}^{\mathbf{f}}(\mathbf{x}) \\ \mathbf{F}^{\mathbf{y}}(\mathbf{x}) \end{pmatrix} \quad (\text{with elements } x_i \text{ and } F_i(\mathbf{x}),\ i = 1,2,\ldots,N+W\,)$$

where

$\mathbf{y}$ = W-vector of minimum O-D disutilities, $y_k = \min\{u_r : r \in R_k\}$ $(k = 1,2,\ldots,W)$

$\mathbf{F}^{\mathbf{f}}(.)$ = N-vector function of elements $u_r(\mathbf{f}) - y_k$ $(r \in R_k ; k = 1,2,\ldots,W)$

$\mathbf{F}^{\mathbf{y}}(.)$ = W-vector function of elements $\sum_{r \in R_k} f_r - q_k$ $(k = 1,2,\ldots,W)$.

LAPUE solutions may then be determined by minimising the gap function:

$$G(\mathbf{x}) = \sum_{i=1}^{N+W} \varphi(x_i, F_i(\mathbf{x})) \quad \text{where} \quad \varphi(a,b) = \tfrac{1}{2}\left(\sqrt{a^2 + b^2} - (a+b)\right)^2 .$$

The minimisation may be carried out by any standard technique. The remaining complication is how to deal with the resulting non-additive shortest path problems. Lo & Chen (2000) suggest two alternative strategies, either using a pre-defined path set or building a path set as required using a kind of heuristic column generation method. In the case of the pre-defined path set, the paths may be generated by the application of a standard UE or SUE algorithm.

3. FORM OF THE ARC TRAVEL TIME DENSITY FUNCTIONS

The LAPUE approach introduced in section 2 is ‘generic’, in the sense that no particular form of travel time density function is assumed. It is natural to examine the impact of

[3]Gordon *et al* (2001) developed heuristics for problems of the kind considered in the present paper, whereby arc-additive Taylor-series approximations to the path disutilities are used in a shortest path method, but in numerical tests these had limited success.

alternative forms of density function, in order to understand better the implications of the modelling approach proposed. For the case of parallel paths, each consisting of a single arc, Noland & Small (1995) derived the appropriate expression resulting from (6) for the cases of exponential and uniform travel time distributions. In practice, it would seem unlikely that the uniform would be a reasonable fit to actual travel time variations. The exponential distribution is rather inflexible, having only a single parameter, and so again might be expected to be troublesome in fitting to real data.

In the general network case, one of the most straightforward and flexible assumptions is to assume that arc travel times follow a Multivariate Normal distribution, with a flow-dependent mean vector $\mathbf{t}(\mathbf{v})$ based on the arc performance functions, and a constant covariance matrix. This implies that the joint density of path travel times is (for fixed $\mathbf{v}$) also a Multivariate Normal, and so the marginal densities for individual paths (needed in (6)) will be univariate Normals. Such an approach has a number of advantages and disadvantages. The advantages are:

- The convolution of arc travel times gives marginal path travel time densities of a simple form.
- Correlations between arcs can be specified. Such correlations might be expected when the variation reflects, say, the impact of rain or snow, or a traffic incident with far-reaching consequences. Variations in demand, such as that due to a special event for example, are also likely to have a correlating effect. In such cases, on any single trip, one would tend to expect all travel times to be higher than average.
- While arc correlations naturally arise, actually estimating them is a different problem. If they can be assumed to be small in relation to the variances, then we might justify neglecting them. A special case of the distribution suggested is when the arc travel times follow independent Normal distributions (as assumed by Yin & Ieda, 2001).
- A quite different issue is that of *route* travel time correlations. It could be argued that, from the viewpoint of logical model consistency, these are considerably more important than the arc travel time correlations. Whether or not the arc travel times are independent or correlated, the path travel times will be correlated due to the fact that paths overlap, and so share some common arcs. Thus, a driver who perceives one arc to be 'risky' will perceive it as risky for all paths using that arc. The network structure means that these correlations must exist. The fact that the path densities are formed from component link densities ensures that such correlations are automatically taken account of.

Some disadvantages of the multivariate Normal assumption are:

- Negative link travel times are in principle permitted, though if the model fits well to observed data then the impact will be negligible (see the point below about fit).
- The implicit assumption of univariate Normal link density functions assumes symmetry. However, many of the sources of variability (in traffic flows and link capacities, for example) will likely have an asymmetric effect on travel times, yielding a positively skewed distribution.

We shall see how the disadvantages noted above may be overcome later in the section, by an extension of the basic Normality assumption. For the moment, however, we explore the implications of the multivariate Normal assumption for the expected utility function (6). In particular, this assumption implies that the marginal route travel times follow a (univariate) Normal distribution, say $C_r \sim \text{Normal}(\mu_r, \sigma_r^2)$ for each route r. It is worth remembering that although we require only the marginal route travel time densities for (6), we are not assuming routes to be statistically independent. Under this model, (6) then can be shown to be expressible as:

$$u_r = \theta_0 \, d_r + \theta_1 \, \mu_r + \theta_2 \, \sigma_r \, L\left(\frac{\tau_k - \mu_r}{\sigma_r}\right). \tag{9}$$

The function L, known as the unit normal linear loss integral, arises among other applications in inventory and stock control (e.g. Johnson & Montgomery, 1974; Dunkerley, 2000), and is defined as:

$$L(x) \;=\; \int_x^{\infty} (u - x)\, \phi(u)\, \mathrm{d}u \;=\; \phi(x) \;+\; x\, \Phi(x) \;-\; x \qquad (-\infty < x < \infty) \tag{10}$$

where $\phi(.)$ and $\Phi(.)$ are respectively the probability density function and cumulative distribution function of a Normal(0,1) variate. It is noted that L is a continuous, decreasing function, and thus (since the argument of L is monotonically decreasing in μ_r) the conditions of the Existence and Uniqueness results in section 2 are met.

A number of properties of (9) may be readily observed. In particular, as τ_k is increased, reflecting greater arrival time flexibility, then since L is a decreasing function, the expected late arrival time $\sigma_r \, L\left(\frac{\tau_k - \mu_r}{\sigma_r}\right)$ will decrease, and therefore have a lesser impact on the overall route disutility. This happens regardless of the size of σ_r; in the LAPUE model, it is not variability itself that is unattractive to drivers, only its potential impact on late arrival. As $\tau_k \to \infty$, $L\left(\frac{\tau_k - \mu_r}{\sigma_r}\right) \to 0$ and the LAPUE model approaches a standard UE model. Similarly, as $\sigma_r \to 0$ then $\sigma_r\, L\left(\frac{\tau_k - \mu_r}{\sigma_r}\right) \to 0$, and again a standard UE model is recovered.

It should be emphasised, however, that (9) does not provide support for a utility function that is linear in μ_r and σ_r, since the 'loss' terms are themselves functions of μ_r and σ_r. However, it is noted in passing that if we modified our approach so that drivers had a *path-specific* (rather than trip-specific) expectation of an acceptable travel time, then some considerable simplification is possible. In particular, if the latest path-specific acceptable arrival time $\tilde{\tau}_r$ for path *r* is supposed to be a path-independent function of the travel time standard deviation, $\tau_r = \mu_r + b\sigma_r$ for some constant *b*, then the coefficient of σ_r in (9) would collapse to the constant $\theta_2 L(b)$. However, the link to late arrivals for performing activities at the destination is then lost, for suppose $b = 2$ and consider two paths, with $\mu_1=\mu_2=10$, $\sigma_1=3$ and $\sigma_2=2$. Then a lateness penalty is incurred on path 1 if the travel time exceeds 16, whereas on path 2 a lateness penalty is incurred with a travel time of only more than 14. Since departure times are fixed, this implies that the penalty incurred depends not only on the arrival time but also on the particular path followed. Thus the 'lateness penalty' incurred by a trip arriving at the destination on path 2 may be greater than that of a *later* arrival on path 1. This may be said to occur as path 2 was not expected to give such an extreme travel time relative to its own distribution, even though the time may be a lesser one than that on path 1. This alternative formulation is mentioned only for the purposes of contrast, and will not be considered further in the paper.

As noted earlier, while the Normal distribution has a number of advantages, it also has some disadvantages, in particular its assumption of symmetry. Even if it could be argued that sources such as variation in the demand matrix might lead to an approximately symmetric arc *flow* probability distribution, this would not imply that the arc travel time distribution is also symmetric. This is due to the fact that the impact of high flow days on travel times would be expected to be more pronounced than the impact of low flow days. For similar reasons, capacity-reducing incidents also typically have an asymmetric effect on travel times. The problem is that, although it is entirely logical that variance increases with unreliability, this does not imply that variance is sufficiently fundamental to describe unreliability. In particular, with the Normal distribution, as we increase the variance, we are not only increasing the probability of particularly long journey times, but at the same rate are increasing the probability of particularly short journey times.

A more logical travel time distribution, it seems therefore, is an asymmetric distribution, and in particular some form of positively skewed distribution. While there are clearly a number of standard candidate forms, such as the Lognormal or gamma, for fitting to data it would seem particularly useful to hypothesise a distribution with a good deal of control (in terms of parameters) over the right-hand tail. Such a distribution could be a *mixture* distribution, as described below.

Suppose that part of the variation of the continuous random variable X is explained by knowledge of the value of some categorical/discrete variable Y, which can take one of a finite number of values $y_1, y_2, ..., y_m$. Suppose that the probability distribution of Y is denoted $p_i = \Pr(Y = y_i)$ $(i = 1,2,...,m)$, where $\sum_{i=1}^{m} p_i = 1$, and that the conditional probability densities of $X | Y$ are denoted $g(x | Y = y_i) = g_i(x)$ $(i = 1,2,...,m)$. Then the unconditional density $g(x)$ of the variable X is given by $g(x) = \sum_{i=1}^{m} p_i \, g_i(x)$, and is known as a *mixture distribution*.

As a simple example, suppose that an arc has two possible distinct states, which we might term "incident" and "non-incident". The probability of an incident occurring is denoted p. Suppose that the travel time probability density in the non-incident state is $g_1(t)$ and in the incident state is $g_2(t)$. The resultant distribution is a mixture distribution with density function $g(t) = (1 - p) g_1(t) + p g_2(t)$.

For example, suppose that in the two-state example above, g_1 is a Normal$(10,2^2)$ density, and g_2 a Normal$(15,2^2)$ density. The resultant density functions, for four choices of p, are illustrated in Figure 1, with summary statistics given in Table 1, illustrating the control possible over the right-hand tail.

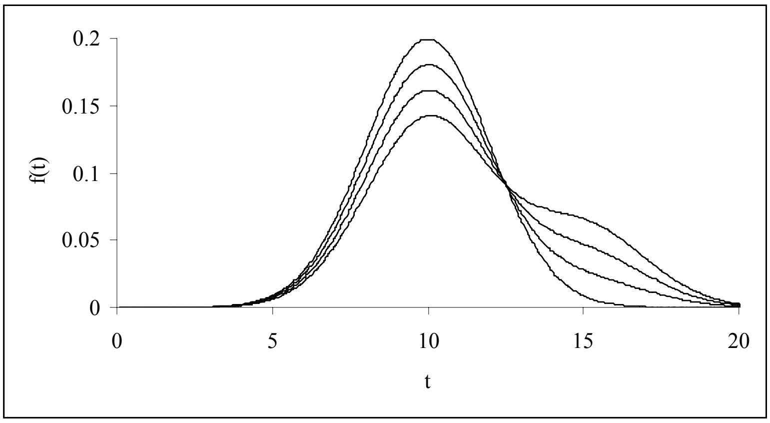

Figure 1: Density functions corresponding to mixture of Nor($10,2^2$) and Nor($15,2^2$)

p	Mean	St dev	Mode
0	10.0	2.0	10.0
0.1	10.5	2.5	10.0
0.2	11.0	2.8	10.1
0.3	11.5	3.0	10.1

Table 1: Summary statistics for mixture of Nor($10,2^2$) and Nor($15,2^2$)

If we suppose, for example, that the arc travel times are statistically independent, each following Normal mixture distributions, then the joint *path* travel times density will be a mixture of multivariate Normals, with the marginal path travel time densities univariate Normal mixtures. If the resulting travel time density for path r is then a mixture of m Normal(μ_{ir},σ_{ir}^2) densities ($i = 1,2,\ldots,m$) with mixing parameters p_{ir} ($i = 1,2,\ldots,m$), then the generalisation of (9) is:

$$u_r = \theta_0\, d_r + \theta_1 \sum_{i=1}^{m} p_{ir}\mu_{ir} + \theta_2 \sum_{i=1}^{m} p_{ir}\sigma_{ir} L\left(\frac{\tau_k - \mu_{ir}}{\sigma_{ir}}\right). \qquad (11)$$

Correlations between arcs can be introduced by grouping arcs and assuming multivariate Normal mixtures; again marginal univariate Normal mixtures arise for the path travel times.

4. ILLUSTRATIVE EXAMPLES

As a first step, we consider the two path network in Figure 2, and in Figure 3 the effect of changing the path 1 travel time variance on the LAPUE solution is examined. In the special case θ_2 = 0, no value is placed on late arrival, and LAPUE collapses to a standard UE. Similarly, as $\tau_1 \rightarrow \infty$, then even for non-zero θ_2, LAPUE approaches UE as drivers increase in flexibility. In the example, it can be seen how an increasing travel time standard deviation on path 1 leads, in the LAPUE model, to an increasingly deterrent effect on the flow using path 1, this flow substantially different from the UE solution. However, this result also depends on the late arrival flexibility τ_1; as τ_1 is increased the LAPUE graph will approach the UE line, as the perceived lateness penalty diminishes.

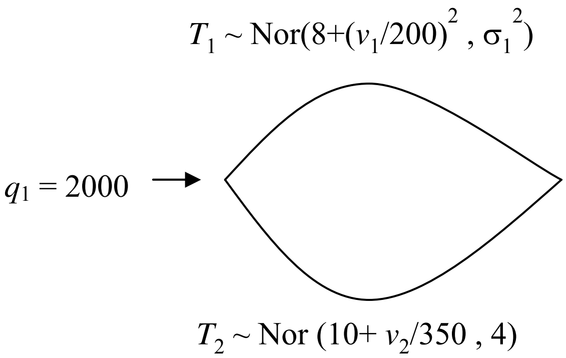

Figure 2: Test Network 1 ($\theta_0 = 0$, $\theta_1 = 1$, $\theta_2 = 2$, $\tau_1 = 15$).

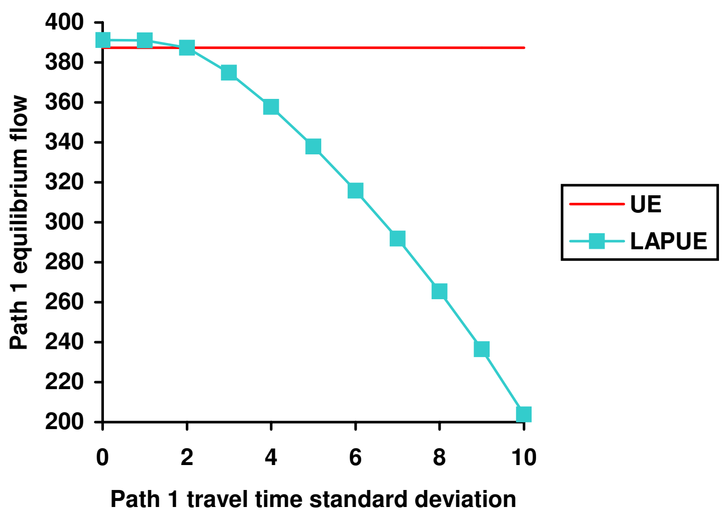

Figure 3: Impact of travel time variability on LAPUE solution, Network 1

An application of the normal mixture model specified in section 3 is now considered. With probability p_a an incident occurs on arc *a*. When an incident does *not* occur, the travel time density is Nor(μ_a,σ_a^2) (by an abuse of notation, we here allow μ and σ to denote properties of an arc, rather than a path). When an incident does occur, the density is Nor($k_a\mu_a$,σ_a^2), for some given $k_a \geq 1$. The overall mean travel time for each arc a is then $(1-p_a)\ \mu_a + p_a\ k_a\mu_a$, and it is these values that are equilibrated with the arc travel time functions in LAPUE.

Test network 2 is illustrated in Figure 4, with the three possible paths also labelled for reference. The UE path flows are (300, 100, 600) at path travel times of 22.0. For pure Normal arc travel times, LAPUE path flows are (338, 0, 662) at disutilities (22.0, 22.1, 22.0). Therefore the active paths are different in the UE and LAPUE cases. Also, in the LAPUE state, the path mean travel times are unequal (20.9, 20.7, 21.9), with the unused path actually having the lowest mean time (but unappealing as it uses two 'risky' arcs).

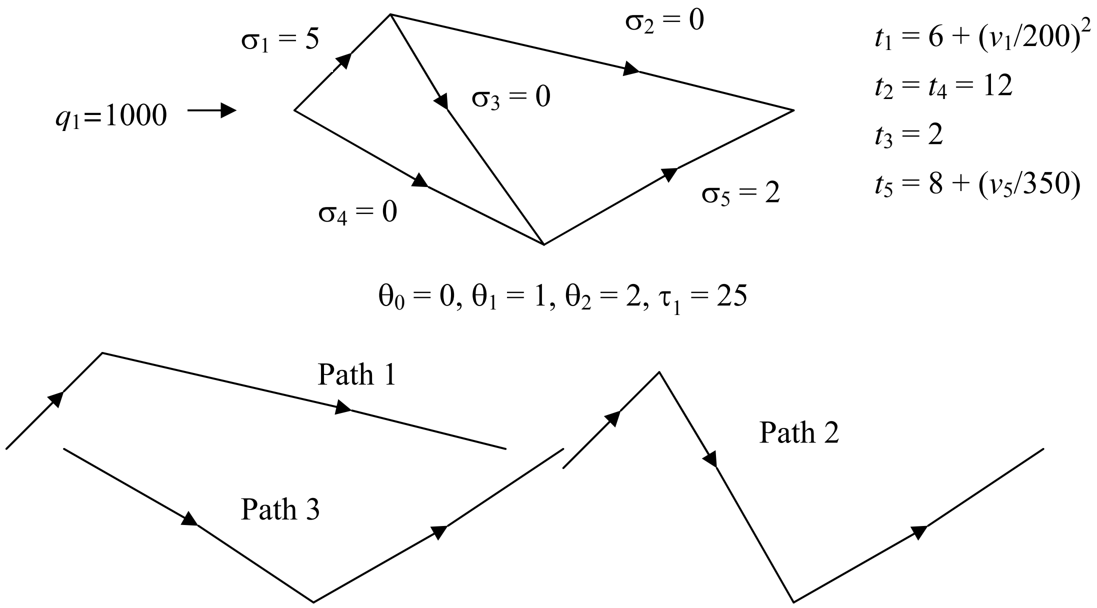

Figure 4: Test Network 2

For the same example, but assuming incidents now occur on arc 1 with probability $p_1 = 0.1$, the graphs in Figure 5, 6 and 7 illustrate respectively the LAPUE path flows, path disutilities and path mean travel times at varying values of k_1, the inflation factor for mean travel times on the 'incident' arc, arc 1. In all cases, path 2 remains unused. While an increase in k_1 would *increase* the path 1 mean travel time at *fixed* flows, it can be seen from Figure 5 that the LAPUE model predicts that drivers would mitigate this effect by diverting away from the risky route, to the extent that the *equilibrium* path 1 travel time *decreases* with k_1 on the used paths (see Figure 7). This is because, as illustrated in Figure 6, it is the path disutilities that are being equilibrated, not mean path travel times: therefore, in Figure 6 the disutilities for the used paths coincide, the disutility for the unused path (path 2) is higher than the disutility on the used path.

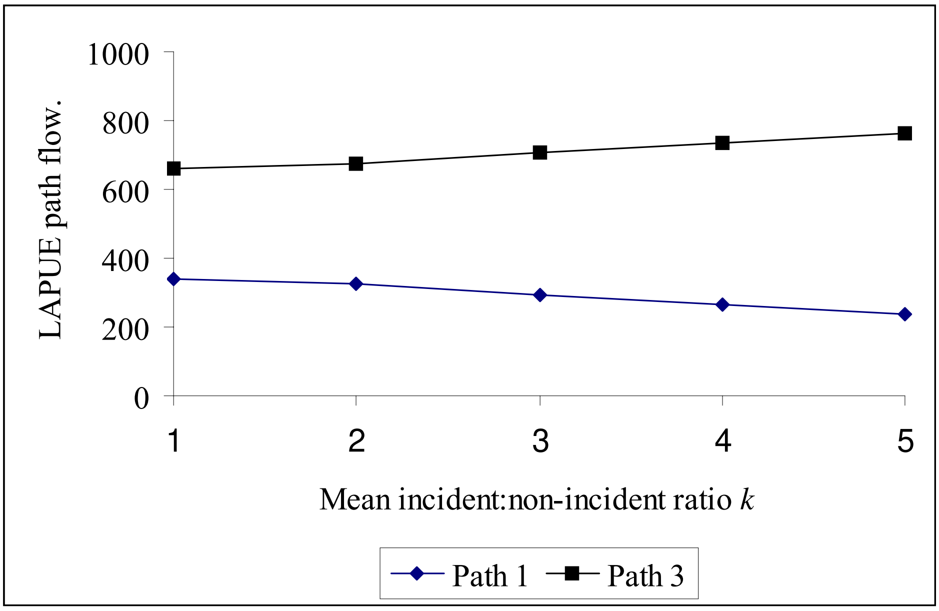

Figure 5: LAPUE used path flows for Network 2, under various values of k_1

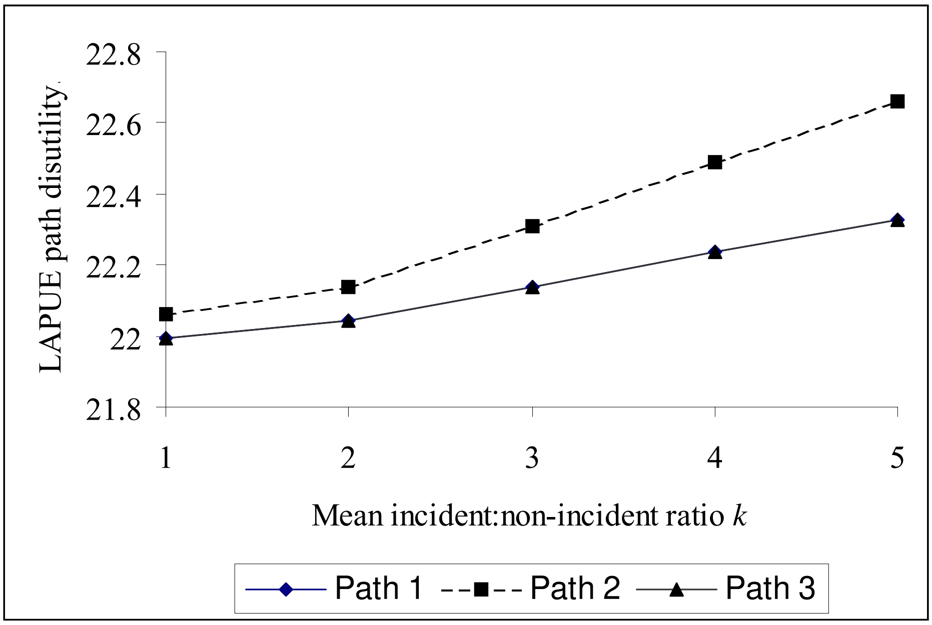

Figure 6: LAPUE path disutilities for Network 2, under various values of k_1

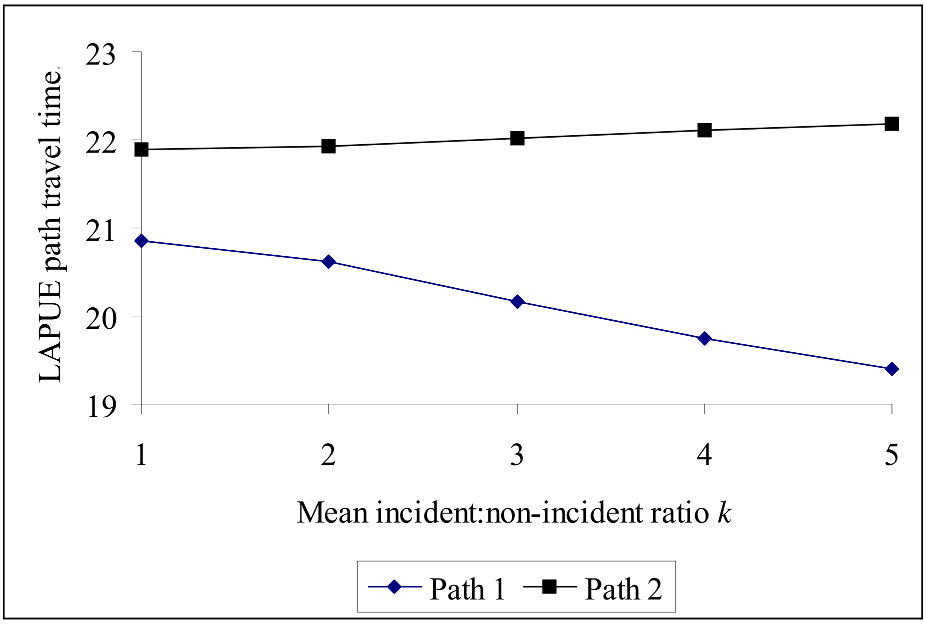

Figure 7: LAPUE used path mean travel times for Network 2, under various values of k_1

4. Stochastic User Equilibrium and the LAPUE Model

A generalisation of the conventional UE model—termed LAPUE—has been presented, to reflect driver responses to stochastically-varying travel times. It is then natural to ask how the concepts introduced relate to the established method for incorporating stochastic elements in an equilibrium framework known as Stochastic User Equilibrium (SUE), due to Daganzo & Sheffi (1977). The purpose of the present section is contrast the treatment of "stochasticity" in a LAPUE model with that in a conventional SUE model.

Definition **(Daganzo & Sheffi, 1977)** $\mathbf{f}^* \in D$ is a SUE if and only if

$$\mathbf{f}^* = \begin{bmatrix} \mathbf{f}^*_{[1]} \\ \mathbf{f}^*_{[2]} \\ . \\ . \\ \mathbf{f}^*_{[W]} \end{bmatrix} = \begin{bmatrix} q_1\, \mathbf{p}_{[1]}\left(\mathbf{c}_{[1]}(\mathbf{f}^*)\right) \\ q_2\, \mathbf{p}_{[2]}\left(\mathbf{c}_{[2]}(\mathbf{f}^*)\right) \\ . \\ . \\ q_W\, \mathbf{p}_{[W]}\left(\mathbf{c}_{[W]}(\mathbf{f}^*)\right) \end{bmatrix} \qquad (12)$$

where for each $k = 1,2,...,W$, the column vectors $\mathbf{f}^*_{[k]}$ and $\mathbf{c}_{[k]}(\mathbf{f}^*)$, of dimension $|R_k|$, respectively denote the flows and costs on paths serving movement k, and $\mathbf{p}_{[k]} : \Re^{|R_k|} \to [0,1]^{|R_k|}$ is a given mapping from the space of path cost vectors to the space of path choice fraction vectors for movement k.

This is a rather general definition, yet the specific forms proposed in the literature presume the path choice fractions to be given by a *random utility model*:

$$p_{r[k]}(\mathbf{c}) = \Pr\left(c_r + \varepsilon_r \le c_s + \varepsilon_s \quad \forall\, s \in R_k\right) \qquad (r \in R_k;\ k = 1,2,...,W) \tag{13}$$

where the $p_{r[k]}(.)$ $(r \in R_k)$ are elements of the vector function $\mathbf{p}_{[k]}(.)$, and where the ε_r terms $(r \in R_k;\ k = 1,2,...,W)$ are random variables, conventionally assumed to have zero mean, following a given joint probability distribution. The two commonest examples of such a model in the literature are the logit SUE (derived from an assumption of independent Weibull random disturbance terms) and the probit SUE (derived from multivariate normal random disturbance terms); see Sheffi (1985).

Now, in practice the generalised cost of any path *r* is typically assumed to be a linear composition of travel time t_r and other attributes that are independent of flow/travel-time; without loss of generality, we simply refer to these latter attributes in a composite form d_r (e.g. distance). Then the premise behind the application of random utility theory to SUE is that a randomly-distributed perceived cost is associated with each alternative path *r*, of the form:

$$\text{Perceived cost of path } r \ = \ \theta_0 d_r + \theta_1 t_r + \varepsilon_r \tag{14}$$

where θ_0 and θ_1 are constants, d_r and t_r are deterministic attributes of path *r*, and ε_r is a random 'error' term. Equivalently, since only path *differences* in cost are relevant in the random utility model (13), we may divide (14) throughout by θ_1 and instead consider:

$$\text{Perceived generalised travel time of path } r = \tilde{\theta}_0 d_r + t_r + \tilde{\varepsilon}_r \tag{15}$$

where $\tilde{\theta}_0 = {}^{\theta_0}\!/_{\theta_1}$ and $\tilde{\varepsilon}_r = {}^{\varepsilon_r}\!/_{\theta_1}$.

An obvious question to ask then is: what exactly do the random error terms in (14) or (15) represent? In practice, they are likely to represent a composition of a number of components, with plausible interpretations of the components including:

- *Representation of unexplained attributes*, i.e. factors that affect route choice which are not included in d_r or t_r.
- *Taste variation* in the sense of driver-to-driver differences in the weights attached to the explained attributes. While in structural terms this would more properly be reflected by postulating a multiplicative random error term such as variation in θ_1 (see, for example, Leurent, 1998; Nielsen *et al*, 2002), in models where θ_1 is assumed constant part of the additive variation may be attempting to explain this facet.

- *Lack of model fit* other than the two sources noted above, ranging from lack-of-fit of the random utility paradigm to actual decision processes, through to mis-specification, mis-recording and coding errors in the network and demand data.
- *Drivers' estimation errors*, in the sense of their inability to estimate correctly the actual mean travel times. For example, if the actual variance in travel times on a particular arc were σ^2, and if drivers' past trip experiences could be represented as a process of random sampling of realised actual travel times in a sample of size *n*, then the variance in their estimation of mean travel times would be σ^2/n.

Aside from such philosophical discussions, it may be tempting nevertheless to adopt a pragmatic approach whereby the term $t_r + \tilde{\varepsilon}_r$ in expression (15) is equated with the actual variation in travel times. Even if difficult to justify on behavioural/philosophical grounds, it may seem that this would provide a device for representing the impact of travel time variability in a similar spirit to the LAPUE model. However, consider Example 1 introduced in §1. Assuming that the actual travel time variation in arc 1 follows a Normal distribution, and equating the actual travel time distribution with the perceptual distribution of an SUE model, a probit SUE model arises:

$$v_1 = q \ \Pr\big(t_1(v_1) + \varepsilon_1 \le t_2(v_2)\big); \qquad v_2 = q - v_1 \ ; \qquad \varepsilon_1 \sim \mathrm{Nor}(0, \sigma_1{}^2)$$

which implies

$$v_1 = q \ \Phi\left(\frac{t_2(q - v_1) - t_1(v_1)}{\sigma_1}\right) \qquad (0 \le v_1 \le q)$$

where $\Phi(.)$ denotes the unit normal cumulative distribution function. By inspection, the unique SUE solution to this problem may be verified to be $v_1 = v_2 = q/2$, regardless of σ_1 (recalling that as the travel times are functionally identical, $t_1(\frac{q}{2}) = t_2(\frac{q}{2})$). Therefore, the existence of the variance $\sigma_1{}^2$ has no effect.

The example above represents an extreme case, but is chosen to illustrate that there is nothing intrinsic in the basic SUE model to deal with risk-aversion. An interpretation of the situation being modelled here is one in which each driver randomly samples a *single* travel time for each path from the distribution of actual travel times, and makes a decision on that basis. Yet this seems inconsistent with the hypothesis of equilibrium analysis, with the justification being that an equilibrium state may be assumed to prevail after some period of experimentation, adaptation and stabilisation of the system, in which drivers' experiences are implicitly more extensive than that of making a single previous journey.

This can be contrasted with a more logical interpretation of the SUE distributions, as representing differences across the population in their estimation of expected travel times; in that case, sampling from the distribution would now represent sampling different individuals across the population and their perception of expected travel time

In contrast to the SUE solution noted above, the LAPUE solution for this problem would be (neglecting the possibility that $v_1 = 0$ could be a LAPUE) the solution in v_1 to:

$$t_1(v_1) + \sigma_1 L\left(\frac{\tau - t_1(v_1)}{\sigma_1}\right) = t_2(q - v_1) \qquad (0 \le v_1 \le q) .$$

Since, as noted previously, the correction term on arc 1 is non-negative, then $t_1(v_1) \le t_2(q - v_1)$, and since the travel time functions are functionally identical and monotonically increasing, this implies $v_1 \le q - v_1$, i.e. $v_1 \le q/2$. Moreover, since the correction term is an increasing function of σ_1, then the arc 1 LAPUE flow will be a decreasing function of σ_1 (see also the example illustrated in Figure 3).

Therefore, it is appropriate to consider LAPUE and SUE as two distinct approaches. Effectively, the stochastic element in LAPUE deals with trip-to-trip variation in actual travel times and its effect on risk-taking strategies, whereas the stochastic element in SUE essentially deals with differences in perceptions of utility across the driver population. Indeed, the elements represented by the two modelling approaches are sufficiently distinct that (as explained in the conclusion section) it would be attractive to consider combining the LAPUE and SUE models into a single approach.

5. CONCLUSIONS AND FUTURE RESEARCH

A generalised form of the UE model, termed LAPUE, has been presented in this paper, which is able to reflect drivers aversion to the risk of late arrival, in the face of uncertain travel times. Existence and uniqueness of LAPUE solutions is guaranteed under mild extensions to the sufficient conditions for the UE case. While the non-additive path disutilities make the application of standard equilibrium algorithms problematic, it has been suggested that recent advances in smooth, gap-function based methods allow such problems to be solved, at least for small to moderate sized networks (though this has yet to be tested in practice). Two specific forms of the model, based on multivariate normal travel times and normal mixture distributions, have been investigated, and have been seen to produce plausible effects, which nevertheless differ substantively from those predicted by a UE model. Whether the LAPUE or UE model is the most realistic is an open question, but the preliminary results here indicate that this is a question that is worthwhile addressing through empirical studies, given the potential differences in model forecasts.

Aside from empirical work to understand individuals' attitudes to risk/variability, and numerical tests of the algorithm proposed on larger scale networks, there are many further areas for methodological research with this approach. In the basic model described it has been assumed that travel time variability is an exogenously-defined feature of the transport system, whereas in practice a primary source of travel time variation is the variation in activities. This variation manifests itself in variable link flows, which could in turn impact on the distribution of travel times; as shown in Watling (2002), such endogenous sources of variation may in principle be capture using generalised notions of equilibrium. The potential is attractive for combining such a model, or an approximation thereto, with the late arrival/travel time variability considerations proposed in the present paper

On the behavioural side, the realism of the LAPUE model would be greatly improved by reflecting the differences in individuals and trips, perhaps through randomly distributed model elements. Such an approach could be used to reflect: differing valuations of late arrival by journey purpose, for example, given a purpose-specific demand matrix; perceptual differences/errors in actual travel times between individuals; and, perhaps most significantly, a profile of preferred arrival times (through randomly-distributed τ_k values). In the spirit of current developments in mixed logit models, these elements might be combined to represent a single randomly utility model, with equilibrium then corresponding to an SUE counterpart to the LAPUE model (LAPSUE?).

Finally, the representation of the user's approach to network reliability can be viewed as one element of an overall framework for reliability assessment. It seems that a number of tools developed for the present application will also be useful in providing network performance measures of reliability, which is a highly active research topic.

ACKNOWLEDGEMENTS

I am grateful to a number of colleagues for their comments at earlier stages of this work, in particular Andrew Gordon, Martin Hazelton, Richard Clegg, Fabian Leurent, and Agachai

Sumalee, as well as those of two anonymous referees, though the views expressed in the paper are the responsibility of the author alone.

REFERENCES

Abdel-Aty MA, Kitamura R & Jovanis PP (1997). Using stated preference data for studying the effect of advanced traffic information on drivers' route choice. *Transpn Research* **5C**, 39-50.

Alfa AS (1986). A review of models for the temporal distribution of peak traffic demand. *Transportation Research* **20B**(6),491-499.

Arnott R, De Palma A & Lindsey R (1990). Departure time and route choice for the morning commute. *Transportation Research* **24B**, 209-228.

Arnott R, De Palma A & Lindsey R (1991). Does providing information to drivers reduce traffic congestion? *Transportation Research* **25A**(5), 309-318.

Bell MGH & Cassir C (2000). *Reliability of transport networks*. Research Studies Press Ltd, Baldock, England.

Ben-Akiva M, De Palma A & Kanaroglou P (1986). Dynamic model of peak period traffic congestion with elastic arrival rates. *Transportation Science* **20**, 164-181.

Chen A, Yang H, Lo H & Tang WH (1999). A capacity related reliability for transportation networks. *Journal of Advanced Transportation* **33**(2), 183-200.

Daganzo CF & Sheffi Y (1977). On stochastic models of traffic assignment. *Transportation Science* **11**(3), 253-274.

De Palma A & Rochat D (1999). Understanding individual travel decisions: results from a commuters survey in Geneva. *Transportation* **26**, 263-281.

DETR (2000). The Reliability Sub-Objective. *Guidance on Methodology for Multi-Modal Studies*, http://www.detr.gov.uk/itwp/mms/ .

Du ZP & Nicholson A (1997). Degradable transportation systems: sensitivity and reliability analysis. *Transportation Research* **31B**, 225-237.

Dunkerley CE (2000). The Impact of Road Transport Disruptions on Shippers and Transport Operators: A Case Study from Columbia. *PhD thesis, Institute for Transport Studies, Leeds.*

Emmerink RHM, Axhausen KW, Nijkamp P & Rietveld P (1995). The potential of information provision in a simulated road network with non-recurrent congestion. *Transportation Research* **3C**, 293-309.

Gabriel S & Bernstein D (1997). The traffic equilibrium problem with non-additive path costs. *Transportation Science* **31**(4), 337-348.

Gordon A., Van Vuren T., Watling D.P., Polak J., Noland R.B., Porter S. & Taylor N. (2001). Incorporating variable travel time effects into route choice models. *Proceedings PTRC European Transport Conference* (Methodological Innovations), PTRC Education & Research Services Ltd., London.

Hall RW (1993). Non-recurrent congestion: How big is the problem? Are traveler information systems the solution? *Transportation Research* **1C**(1), 89-103.

Johnson LA & Montgomery DC (1974). *Operations Research in Production, Scheduling and Inventory Control.* John Wiley & Sons, New York.

Leurent F (1998). Sensitivity and Error Analysis of the Dual Criteria Traffic Assignment Model. *Transportation Research* **32B**(3), 189-204.

Lo HK & Chen A (2000). Traffic equilibrium problem with route-specific costs: formulation and algorithms. *Transportation Research* **34B**(6), 493-514.

Mannering F, Kim S-G, Barfield W & Ng L (1994). Statistical analysis of commuters' route, mode and departure time flexibility. *Transportation Research* **2C**, 35-47.

Mirchandani P & Soroush H (1987). Generalized Traffic Equilibrium with Probabilistic Travel Times and Perceptions. *Transportation Science* **21**(3), 133-152.

Nielsen OA, Daly AJ & Frederiksen RD (2002). A Stochastic Route Choice Model for Car Travellers in the Copenhagen Region. *Networks and Spatial Economics* **2**(4), 327-346.

Noland RB & Small KA (1995). Travel time uncertainty, departure time choice, and the cost of morning commutes. *Transportation Research Record* **1493**, 150-158.

Noland RB, Small KA, Koskenoja PM & Chu X (1998). Simulating travel reliability. *Regional Science & Urban Economics* **28**, 535-564.

Noland RB (1999). Information in a two-route network with recurrent and non-recurrent congestion. In: *Behavioural and Network Impacts of Driver Information Systems* (eds. R Emmerink & P Nijkamp), Ashgate Publishing, Aldershot, 95-114.

Porter S, Wright I & Dale H (1996). Are there quantifiable benefits from reducing the variability of travel time*? Proceedings PTRC 24th European Transport Forum*, Seminars D & E – Part 2.

Ran B, Hall RW & Boyce DE (1996). A link-based variational inequality model for dynamic departure time/route choice. *Transportation Research* **30B**(1), 31-46.

Ran B & Boyce DE (1996). *Dynamic Urban Transportation Network Models*. Second Edition, Springer Verlag, Berlin.

Scott K & Bernstein D (1998). Solving a traffic equilibrium problem when paths are nonadditive. Paper presented at 78th annual Transportation Research Board meeting, Washington DC.

Sheffi Y (1985). *Urban transportation networks.* Prentice-Hall, New Jersey.

Sumi T, Matsumoto Y & Miyaki Y (1990). Departure time and route choice of commuters on mass transit systems. *Transportation Research* **24B**(4), 247-262.

Van Berkum E & Van der Mede P (1999). Driver information and the (de)formation of habit in route choice. In: *Behavioural and Network Impacts of Driver Information Systems* (eds. R Emmerink & P Nijkamp), Ashgate Publishing, Aldershot, 155-179.

Vickrey W (1969). Congestion theory and transport investment. *American Economic Review* **56**, 251-260.

Vythoulkas PC (1990). A dynamic stochastic assignment model for the analysis of general networks. *Transportation Research* **24B**(6), 453-469.

Watling DP (2002). A second order stochastic network equilibrium model. I: Theoretical foundation. *Transportation Science* **36**(2), 149-166.

Yang H & Huang H-J (1997). Analysis of the time-varying pricing of a bottleneck with elastic demand using optimal control theory. *Transportation Research* **31B**(6), 425-440.

Yin Y & Ieda H (2001). Assessing performance reliability of road networks under non-recurrent congestion. Paper presented at TRB annual conference, January 2001.