Article:
Clark, S.D. and Watling, D.P. (2006) Applications of sensitivity analysis for probit stochastic network equilibrium. European Journal of Operational Research, 175 (2). pp. 894-911. ISSN: 0377-2217

https://doi.org/10.1016/j.ejor.2005.06.012

Published paper

Clark, S.D.; Watling, D.P. (2006) *Applications of Sensitivity Analysis for Probit Stochastic Network Equilibrium.* European Journal of Operational Research, (175), pp.894-911.

APPLICATIONS OF SENSITIVITY ANALYSIS FOR

PROBIT STOCHASTIC NETWORK EQUILIBRIUM

Stephen D. Clark[1] & David P. Watling*

Institute for Transport Studies, University of Leeds, UK

**Corresponding Author:*

David P. Watling

Institute for Transport Studies

University of Leeds,

Leeds LS2 9JT, U.K.

Tel: +44 113 343 6612

Fax: +44 113 343 5334

Email: dwatling@its.leeds.ac.uk

Date: May 25th 2005

[1] This work was conducted whilst Stephen Clark was employed by the University of Leeds. He is currently employed in the Development Department of Leeds City Council, Leonardo Building, 2 Rossington Street, Leeds LS2 8HD, UK.

Abstract - Network equilibrium models are widely used by traffic practitioners to aid them in making decisions concerning the operation and management of traffic networks. The common practice is to test a prescribed range of hypothetical changes or policy measures through adjustments to the input data, namely the trip demands, the arc performance (travel time) functions, and policy variables such as tolls or signal timings. Relatively little use is, however, made of the full implicit relationship between model inputs and outputs inherent in these models. By exploiting the representation of such models as an equivalent optimisation problem, classical results on the sensitivity analysis of non-linear programs may be applied, to produce linear relationships between input data perturbations and model outputs. We specifically focus on recent results relating to the probit Stochastic User Equilibrium (PSUE) model, which has the advantage of greater behavioural realism and flexibility relative to the conventional Wardrop user equilibrium and logit SUE models. The paper goes on to explore four applications of these sensitivity expressions in gaining insight into the operation of road traffic networks. These applications are namely: identification of sensitive, 'critical' parameters; computation of approximate, re-equilibrated solutions following a change (post-optimisation); robustness analysis of model forecasts to input data errors, in the form of confidence interval estimation; and the solution of problems of the bi-level, optimal network design variety. Finally, numerical experiments applying these methods are reported.

Keywords: Traffic, equilibrium network flows, robustness and sensitivity analysis, probit choice model.

1. INTRODUCTION

The Deterministic User Equilibrium (DUE) is a multi-commodity network flow problem commonly found in road transport which arises from a supposed 'game' between the road-users in the style of a Nash equilibrium (Nash, 1951). In the DUE problem, a vector of network arc flows must be determined which is consistent with individual drivers selecting paths so as to minimise their own generalised travel times (a weighted combination of factors, in travel time equivalent units), where the arc travel times depend on the arc flows (Beckmann et al, 1956). Such models are used widely in practice to plan transport facilities for urban areas, in particular policies that may affect the demand for travel (in terms of the given trip rates between origin and destination nodes), the supply of road network capacity (impacting on the arc relationships between travel time and flow), or the non-time attributes, such as tolls, that impact directly on drivers' generalised travel times. The reader is referred to Nguyen *et al* (1996), for example, for a survey of the literature on the theory and application of such models.

Despite the optimization being at the individual driver, non-cooperative level, rather than a system optimization, it has long been known that a modified system-level, convex optimization problem exists which recovers the DUE solution, at least for 'separable' problems where an arc's travel time is independent of flows on other arcs (Beckmann *et al*, 1956), allowing efficient solution methods for convex problems to be applied (e.g. Frank & Wolfe, 1956). Many variants on this basic DUE model exist, with specialist solution techniques, such as problems with link interactions (e.g. Nguyen & Dupuis, 1984; Patriksson & Rockafellar, 2003), non-additive generalised path travel times (e.g. Chen *et al*, 2001b), and DUE incorporating intra-period dynamic phenomena (e.g. Wie, 1995; Chen *et al*, 2001a)

In practice, determining all the factors impacting on generalised travel times is no trivial task, and in any case the perceptions of these vary across individuals and between trips. An attractive generalisation of DUE is therefore the Stochastic User Equilibrium (SUE) problem, in which unobserved, uncertain or heterogeneous elements of the generalized travel times are reflected through randomly distributed additive components. Drivers select paths based on perceived generalized travel times — the sum of the deterministic part and random disturbance — based on behavioural random utility theory (Sheffi, 1985). Within the SUE family many specific model forms may be defined, depending on the assumed joint probability distribution of the random disturbance terms for the paths of the network. By far the most straightforward case to handle, often favoured in the literature for that reason, is the logit SUE (LSUE) model, where the random terms are assumed to be independent between paths and Weibull distributed, and a convex optimization formulation again exists that is able to recover the LSUE solution (see Fisk, 1980). The LSUE model, however, suffers from a major problem of lack of plausibility. In particular, it neglects correlations in the perception of overlapping paths, whereas intuitively, paths that share many common arcs will be perceived in a very similar way. Effectively, LSUE neglects the network structure.

The implausibility of the LSUE model for network problems has led to many alternatives being proposed within the SUE family, with research in this area especially resurgent in recent years. Notable among these are SUE models based on choice fractions given by the probit (Daganzo & Sheffi, 1977), C-logit (Cascetta *et al*, 1996), nested logit (Gentile & Papola, 2001), cross-nested logit (Vovsha & Bekhor, 1998), paired combinatorial logit (Gliebe *et al*, 1999; Prashker & Bekhor, 1999), mixed logit (Nielsen *et al*, 2002) and gammit (Cantarella & Binetti, 2002). Effectively these models all correspond to alternative statistical assumptions on the correlation structure of the random disturbance terms in a random utility model (C-logit being a slight anomaly in that its link to random utility is through the choice from a implicitly-available set of alternatives, not a fixed choice set).

The SUE family of models will be the focus of the present paper. The approaches we shall explore in sections 3 and 4 do not depend on the assumptions made regarding the particular form of SUE model, and so could in principle be applied to any member of the SUE family described above. However, for illustration we shall choose to focus on one particular such model, namely the Probit SUE (PSUE). There are a number of reasons for choosing to analyse PSUE from the wider SUE family. Firstly, as a model with a long history its properties and limitations are well understood (e.g. Daganzo, 1979), whereas the research community is still on a learning curve in understanding the difficulties with implementing more recent developments, such as mixed logit. Secondly, the probit has a claim to maximal flexibility, in that a wide range of structural relationships for the parameters may be accommodated for estimation purposes. Thirdly, it can have a claim to be the most challenging to analyse through sensitivity analysis, as we shall do, since (i) it admits no known convex optimisation formulation, and (ii) it is conventionally estimated by stochastic approximation methods that introduce the difficulty of handling Monte Carlo error in the computation of derivatives.

Focussing henceforth on the PSUE model, we shall see that it is able to address the problem of correlated paths in an appealing, intuitive manner. The random elements of path generalized travel times are implied from random components for *arc* generalized travel times; in its most general form, the random components of arc generalized travel times follow a multivariate Normal distribution. A commonly used special case assumes the random elements to be independent between arcs, yet the implied *path* random elements are still correlated when paths overlap. While an equivalent optimisation exists for PSUE (Sheffi & Powell, 1982), evaluation of this function is problematic, involving a multivariate integral (expectation) of dimension equal to the number of paths. In practice, Monte Carlo based, stochastic approximation methods are in common use (see Powell & Sheffi, 1982, based on Blum, 1954), as will be adopted in the present paper.

Specifically, we shall consider linear sensitivity analysis of the PSUE solution. Having described recent research by the authors addressing this issue, the paper goes on to illustrate a number of direct applications of such a sensitivity analysis, relevant to practical urban traffic problems, which are derived through alternative interpretations of the sensitivity parameters.

2. SENSITIVITY ANALYSIS OF PROBIT SUE: EXISTENCE, DERIVATION, COMPUTATION

Sensitivity analysis is concerned with the implicit relationship between the solution to a problem (e.g. optimisation problem) and changes in the input parameters to that problem. Given a solution for particular values of the input parameters, the objective of a first order sensitivity analysis is to obtain an approximate linear relationship between changes ('perturbations') to those input parameter values and changes to the solution. In the context of traffic network equilibrium, a variety of problem formulations have been used for the purpose of sensitivity analysis (optimization, variational inequality, fixed point formulations), with techniques derived for DUE (Tobin & Friesz, 1988; Cho *et al*, 2000; Patriksson & Rockafellar, 2003), logit SUE (Bell & Iida, 1997; Ying & Miyagi, 2001) and nested logit SUE (Gentile & Papola, 2001). Here we shall focus on probit SUE (PSUE), the basic definition of which is as a fixed point problem:

(1) $\mathbf{v}^*$ is a PSUE $\Leftrightarrow$ $v_a^* = \sum_{w=1}^{W} q_w \sum_{r \in R_w} \delta_{ar} P_r\left(\mathbf{c}^{(w)}(\mathbf{v}^*) \,\middle|\, \mathbf{\Sigma}^{(w)}\right)$

$(a = 1,2,...,A)$

where

v_a = flow on arc a $(a = 1,2,...,A)$, with $\mathbf{v}$ the vector of flows across all arcs

v_a^* = flow on arc a $(a = 1,2,...,A)$ at a PSUE solution (i.e. a $\mathbf{v}^*$ satisfying (1))

q_w = flow on origin-destination movement w $(w = 1,2,...,W)$

R_w = index set of acyclic paths serving origin-destination movement w

δ_{ar} = indicator variable, equal to 1 if path r contains arc a, equal to 0 otherwise

$c_r(\mathbf{v})$ = generalized travel time on *path* r as a function of the arc flow vector $\mathbf{v}$

$\mathbf{c}^{(w)}(\cdot)$ = vector of functions $c_r(\cdot)$ $(r \in R_w)$, for each $w = 1,2,...,W$

$P_r(\boldsymbol{\mu}|\boldsymbol{\Sigma})$ = for a given movement, probit probability of using path r as a function of the path travel time disturbance mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$

(2) $\mathbf{P}^{(w)}(\cdot|\cdot)$ = vector of functions $P_r(\cdot|\cdot)$ $(r \in R_w)$, for each $w = 1,2,...,W$ (used later) .

In this formulation, the path travel time functions are assumed link-additive, inferred from arc travel time functions (which we shall henceforth assume are 'separable', i.e. depend only on the flow on the arc in question):

$$(3) \qquad c_r(\mathbf{v}) = \sum_{a=1}^{A} t_a(v_a)\,\delta_{ar} \qquad\qquad (r \in R_w; w = 1,2,...,W)$$

where

$t_a(v_a)$ = generalized travel time on arc a as a function of flow v_a $(a = 1,2,..., A)$

$\mathbf{t}(\mathbf{v})$ = vector of functions $t_a(v_a)$ $(a = 1,2,..., A)$ (used later) .

Under mild regularity conditions (for example, assumptions (A2)–(A4) in the subsequent discussion suffice), a unique PSUE link flow vector $\mathbf{v}^*$ is guaranteed to exist (Sheffi, 1985). Furthermore, unlike its better known DUE counterpart, the existence of such a unique $\mathbf{v}^*$ implies the existence of a unique PSUE *path* flow vector, $\mathbf{f}^*$, derived from $\mathbf{v}^*$ by:

$$f_r^* = q_w P_r\left(\mathbf{c}^{(w)}(\mathbf{v}^*)\,\middle|\,\Sigma^{(w)}\right) \qquad (r \in R_w;\, w = 1,2,...,W)\,.$$

One of the main advantages of the probit model is that the path travel time disturbance covariance matrices $\boldsymbol{\Sigma}^{(w)}$ may be inferred from an *arc* travel time disturbance covariance matrix $\boldsymbol{\Lambda}$, with their elements linked by:

$$(4) \qquad \boldsymbol{\Sigma}_{rs}^{(w)} = \sum_{a=1}^{A}\sum_{b=1}^{A}\delta_{ar}\delta_{bs}\boldsymbol{\Lambda}_{ab} \qquad\qquad (r,s \in R_w; w = 1,2,...,W)$$

Assuming $t_a(.)$ to be a continuous and increasing function of v_a, an equivalent unconstrained optimisation formulation of PSUE is (Sheffi & Powell, 1982):

$$(5) \qquad \mathbf{v}^* = \arg\min\left\{ z(\mathbf{v}) = -\sum_{w=1}^{W} q_w\, S_w\left(\mathbf{c}^{(w)}(\mathbf{v})\,\middle|\,\boldsymbol{\Sigma}^{(w)}\right) + \sum_{a=1}^{A} v_a t_a(v_a) - \sum_{a=1}^{A}\int_0^{v_a} t_a(x)\,dx \right\}$$

where $S_w(\boldsymbol{\mu}|\boldsymbol{\Sigma})$, known as the satisfaction function, denotes the expected minimum perceived travel time on movement w corresponding to a probit $\mathbf{P}^{(w)}(\boldsymbol{\mu}|\boldsymbol{\Sigma})$ choice model — the only property of this function needed for later analysis is that $\partial S_w / \partial \mu_r = P_r \ \ (r \in R_w)$.

Let us now suppose there is a real parameter vector $\boldsymbol{\varepsilon}$, which may parameterise either changes to the input demand flows ($q_w = q_w(\varepsilon),\ w = 1,2,...,W$) or changes to the arc generalised travel time functions ($t_a = t_a(v_a, \boldsymbol{\varepsilon}),\ a = 1,2,..., A$). For example, in the first case the perturbations may represent the impact of a new development in generating or attracting trips, and in the latter case might represent network changes to link capacities, signal timings or tolls. For a given $\boldsymbol{\varepsilon}$ we may solve (5), and denote the solution by $\mathbf{v}^*(\boldsymbol{\varepsilon})$. By now varying $\boldsymbol{\varepsilon}$, each time solving (5) for fixed $\boldsymbol{\varepsilon}$, then $\mathbf{v}^*(\boldsymbol{\varepsilon})$ defines an implicit relationship between the PSUE flows and the perturbation vector $\boldsymbol{\varepsilon}$. Applying classical sensitivity analysis for non-linear optimisation problems (Fiacco, 1983) to problem (5) then yields a linear approximation to this implicit relationship, in the neighbourhood of $\varepsilon = \mathbf{0}$, as (Clark & Watling, 2000, 2002):

(6) $$\mathbf{v}^*(\varepsilon) = \mathbf{v}^*(\mathbf{0}) + \mathbf{M}(\mathbf{0})^{-1}\mathbf{N}(\mathbf{0})\ \varepsilon + \mathrm{o}(\|\varepsilon\|)$$

where the matrix $\mathbf{M}$ is given by:

(7) $$\mathbf{M}(\varepsilon) = \sum_{w=1}^{W} q_w \left(\nabla_{\mathbf{v}} \mathbf{c}^{(w)} .(-\nabla_{\boldsymbol{\mu}^{(w)}} \mathbf{P}^{(w)}).\, (\nabla_{\mathbf{v}} \mathbf{c}^{(w)})^{\mathrm{T}} \right) + \nabla_{\mathbf{v}} \mathbf{t}$$

and where the form of the $\mathbf{N}$ matrix depends on the kind of parameterisation represented by ε:

(8) Origin-destination demand: $$\mathbf{N}(\varepsilon) = \left[\left(\sum_{w=1}^{W} \nabla_{\boldsymbol{\varepsilon}}\, q_w \,.\, \mathbf{P}^{(w)} .(\Delta^{(w)})^{\mathrm{T}} \right) . \nabla_{\mathbf{v}} \mathbf{t} \right]^{\mathrm{T}}$$

(9) Arc function parameters: $$\mathbf{N}(\varepsilon) = \left[\sum_{w=1}^{W} q_w \nabla_{\mathbf{v}} \mathbf{c}^{(w)} . \nabla_{\boldsymbol{\mu}^{(w)}} \mathbf{P}^{(w)} . \nabla_{\boldsymbol{\varepsilon}} \mathbf{c}^{(w)} \right]^{\mathrm{T}}$$

where $\Delta^{(w)}$ is the arc-path incidence matrix for movement *w*, with elements δ_{ar} $(a = 1,2,...,A;\ r \in R_w)$ as defined earlier.

—

For this analysis, a number of assumptions will be made:

Assumption (A1): The perturbation Jacobians $\nabla_{\boldsymbol{\varepsilon}} q_w\ (w = 1,2,...,W)$ and $\nabla_{\boldsymbol{\varepsilon}} \mathbf{c}^{(w)}\ (w = 1,2,...,W)$ exist.

Assumption (A2): For each $w = 1,2,...,W$, the probit choice model $\mathbf{P}^{(w)}(\boldsymbol{\mu}^{(w)} | \boldsymbol{\Sigma}^{(w)})$ has a covariance matrix $\boldsymbol{\Sigma}^{(w)}$ that is independent of the mean $\boldsymbol{\mu}^{(w)}$.

Assumption (A3): For each $w = 1,2,...,W$, the covariance matrix $\boldsymbol{\Sigma}^{(w)}$ is non-singular.

Assumption (A4): The link travel time functions $t_a(v_a)\ (a = 1,2,...,A)$ are differentiable and strictly increasing in their argument.

It is then possible to establish the following result:

Lemma 1 (Existence of Sensitivity Analysis) Under assumptions A1–A4, the linear model (6) is well-defined, in the sense that all required derivatives and inverses exist.

Proof Assumption (A3) ensures that $\mathbf{P}^{(w)}$ defines a *random utility model* over the full set of alternative routes (see Daganzo, 1979, for a definition), and then (A2) ensures additionally that $\mathbf{P}^{(w)}$ defines a *regular* random utility model (Daganzo, 1979; Cantarella & Cascetta, 1995). For any regular random utility model, it is known that the Jacobian $\nabla_{\boldsymbol{\mu}^{(w)}} \mathbf{P}^{(w)}$ exists, since $-\nabla_{\boldsymbol{\mu}^{(w)}} \mathbf{P}^{(w)}$ is equal to the Hessian of a convex function, known as the satisfaction function (Daganzo, 1979). Coupled with assumptions (A1) and (A4), we have thus established that all required derivatives for the matrices $\mathbf{M}(\mathbf{0})$ and $\mathbf{N}(\mathbf{0})$ in (6) exist. It remains to show that $\mathbf{M}(\mathbf{0})$ is invertible. Now (A4) implies that the Jacobian $\nabla_{\mathbf{v}}\mathbf{t}$ in (7) is diagonal with positive entries, and so is positive definite. Considering, then, the sum of matrices over *w* in (7), it has been noted above that under (A2) and (A3), each Jacobian $-\nabla_{\boldsymbol{\mu}^{(w)}} \mathbf{P}^{(w)}$ is equal to the Hessian of a convex function and so is positive semi-definite. Then $\nabla_{\mathbf{v}}\mathbf{c}^{(w)} . \left(-\nabla_{\boldsymbol{\mu}^{(w)}} \mathbf{P}^{(w)}\right) . (\nabla_{\mathbf{v}}\mathbf{c}^{(w)})^{\mathrm{T}}$ is a quadratic form of non-zero terms ($\nabla_{\mathbf{v}}\mathbf{c}^{(w)}$ non-zero under assumption (A4)) applied to a positive semi-definite matrix, and so this quadratic form is also positive semi-definite. According to (7), $\mathbf{M}(\boldsymbol{\varepsilon})$ is therefore the sum of positive semi-definite matrices and a positive definite matrix, and so is positive definite. Hence the inverse of $\mathbf{M}(\mathbf{0})$ exists, as required in (6), since all positive definite matrices are non-singular, and the proof is complete. □

Regarding these assumptions, (A1) trivially holds for all cases considered in the present paper, where linear perturbations are made to the origin-destination demand levels and to the parameters of the link travel time functions, these latter functions being continuously differentiable in their parameters. (A2) is ensured by defining the link variance components in section 3 to be a function of the free flow travel time, rather than (say) the SUE travel time. (A4) also trivially holds for all travel time functions of the power-law form adopted here (assuming appropriate parameter values to ensure strictly increasing functions). The only condition that turns out to be not guaranteed *a priori* is (A3), and so we discuss this in some detail below. The derivatives in (7), (8) and (9) are commonly straightforward to compute, aside from the probit path choice Jacobians $\nabla_{\boldsymbol{\mu}^{(w)}} \mathbf{P}^{(w)}$, for which an attractive result due to Daganzo (1979) is exploited, expressing the off-diagonal terms as:

$$\text{(10)} \qquad \frac{\partial P_r(\mu, \Sigma)}{\partial \mu_s} = \sqrt{\frac{|\Sigma(r,s)|}{2\pi|\Sigma|}} \exp\!\left(\tfrac{1}{2} k(r,s)\right) P_r\left(\mu(r,s), \Sigma(r,s)\right) \qquad (r, s \in R_w;\ \ r \neq s)$$

where the right hand probit probability relates to a problem with alternative path s deleted. The matrix $\left(\Sigma(r,s)\right)^{-1}$ is obtained from $\boldsymbol{\Sigma}^{-1}$ by adding row s to row r, adding column s of the resultant matrix to column r, and finally deleting row s and column s of the resultant matrix. The vector $\mu(r,s) = \mathbf{d}(r,s)\Sigma(r,s)$, where $\mathbf{d}(r,s)$ is obtained from $\mu\Sigma^{-1}$ by adding (then deleting) the s^{th} element to the r^{th}. Finally, $k(r,s) = \mathbf{d}(r,s)\boldsymbol{\Sigma}(r,s)\mathbf{d}(r,s)^{\mathrm{T}} - \mu\Sigma^{-1}\mu^{\mathrm{T}}$, a scalar.

The choice probability for the ‘reduced problem’ (i.e. the $P_r\left(\boldsymbol{\mu}(r,s), \boldsymbol{\Sigma}(r,s)\right)$ term in (10)) is estimated by Monte Carlo simulation. This method may be implemented with a path-based variant of the Method of Successive Averages (MSA) algorithm (Powell & Sheffi, 1982) to compute the ‘unperturbed’ ($\boldsymbol{\varepsilon} = \mathbf{0}$) PSUE solution; only paths that are ‘active’ (carry positive flow) in the unperturbed state are considered in the analysis. It should be noted that while, in theory, *all* paths should be active at the PSUE solution, at the termination of any finite number of MSA iterations a number of paths will be so improbable that they will not have been sampled, and so the restriction to a smaller active path set effectively represents an estimate of a most probable set of paths. It is noted in passing that this is somewhat different to the selection of an *arbitrary* path set as suggested in some sensitivity analyses of DUE (e.g. Tobin & Friesz, 1988).

This latter feature, namely that the *estimated* PSUE path flow solution generally uses somewhat less than the full path set turns out to be a key element of our proposed computational method. In particular, it ensures that—for most origin-destination movements in larger networks—condition (A3) will be met, which effectively requires that for each movement, each active path has a link that is unique to that path, not used by other paths on that movement. However, it will be typically be the case that at least for some movements, the active path set is so large that (A3) fails, meaning that the Jacobian elements (10) cannot be computed. Two questions arise, then: (i) what causes this ill-conditioning, and (ii) how may it be overcome?

Considering, firstly, the cause of the ill-conditioning. This arises due to our decomposition of the derivative of the PSUE model into the product of the path travel time *vs* link flow Jacobian, and the path choice *vs* path travel time Jacobians. Such a decomposition is not guaranteed to be well-defined. The ill-conditioning could be avoided by decomposing into the product of a link travel time *vs* link flow Jacobian, and *link choice vs* link flow Jacobians for each movement, the latter referring to the proportion of flow on a particular movement that, across all paths for that movement, would use each link. The disadvantage of this alternative decomposition is, however, that the appealing computational result (10) does not apparently generalise to allow efficient computation of probit link choice Jacobians. That is to say, the ill-conditioning problem noted is not fundamental to performing sensitivity analysis of PSUE, but arises from the particular way we propose to *compute* the sensitivity analysis by exploiting result (10).

Secondly, then, we address the question of how to overcome the ill-conditioning noted. The problem has structural similarity to a known problem in the maximum likelihood estimation of probit choice model parameters, where particular parameterisations may not be *estimable* (Daganzo, 1979, pp 93-105). Daganzo suggests, at least for some simple types of problem, how the parameterisation may be adapted to avoid this problem. Alternatively, one may drop the appealing manner of defining the PSUE covariance matrix purely in terms of link error components (4). For example, Yai *et al* (1997) include additional path-specific error terms in their specification of the probit model, which we could simply *require* to satisfy (A3). While this approach has theoretical appeal, in that all ill-conditioning problems disappear and still

(10) may be exploited, it generates a new practical problem of how to generate/estimate the path-specific terms (unless they are treated as arbitrary, 'small' disturbances). In any case, the appealing practical appeal of PSUE, in defining the error terms as link components and letting the network structure define the path covariances, is lost.

In practice, neither of the strategies suggested above was adopted, but instead a simple technique was employed to 'prune' the path set in order to satisfy (A3). The aim of the heuristic is, for any origin-destination movement for which (A3) fails, to choose a linearly-independent subset of the active paths that explains the greatest proportion of the demand for that movement, in the unperturbed state. A simple greedy heuristic, whereby the active paths are ranked by flow carried in the unperturbed state, has been seen to overcome this problem at no appreciable loss of information (Clark & Watling, 2002). For paths not selected, the flows are held constant during the sensitivity analysis.

3 . SPECIFICATION OF TEST NETWORKS

In order to illustrate applications of the results above, two test networks will be used. The first is an artificial five-arc network, which has previously been used in the literature on related problems (Suwansirikul *et al*, 1987; Cho & Lo, 1999). The network structure and travel time functions are given in Figure 1. The origin-destination demand for the single movement is q = 100. The arc travel time disturbance covariance matrix Λ—which infers a *path* covariance matrix (in general, non diagonal) through (4)—is assumed to be diagonal, with the square root of the diagonal elements (standard deviations) proportional to the free-flow arc travel times. The base case used in all tests reported in section 3 uses a proportionality of 0.3, yielding:

$$\Lambda = \begin{pmatrix} 1.2^2 & 0 & 0 & 0 & 0 \\ 0 & 1.8^2 & 0 & 0 & 0 \\ 0 & 0 & 0.6^2 & 0 & 0 \\ 0 & 0 & 0 & 1.5^2 & 0 \\ 0 & 0 & 0 & 0 & 0.9^2 \end{pmatrix} = \begin{pmatrix} 1.44 & 0 & 0 & 0 & 0 \\ 0 & 3.24 & 0 & 0 & 0 \\ 0 & 0 & 0.36 & 0 & 0 \\ 0 & 0 & 0 & 2.25 & 0 \\ 0 & 0 & 0 & 0 & 0.81 \end{pmatrix}.$$

In conducting the sensitivity analysis, an unperturbed PSUE $\mathbf{v}^*(\mathbf{0})$ is first required; this is estimated using 32,000 iterations of the MSA algorithm, the large number of iterations used in order to minimise the confounding effect of Monte Carlo error in our test comparisons.

The second network considered, referred to as the Headingley network, represents an area of some 6km×3km in the city of Leeds, UK, based around the main arterial A660 Otley Road to the north-east of the city centre. It contains some 123 arcs, 29 origin zones and 29 destination zones, with a peak-hour demand matrix totalling 6,260 vehicles/hour, and with link travel time functions of the BPR form (Bureau of Public Roads, 1964). The path travel time covariance matrix is specified in an analogous way to that for the five-arc network, based on a diagonal arc travel time covariance matrix with standard deviations equal to 0.3 of free flow travel time. To compute the base case PSUE solution, 1,000 iterations of the MSA algorithm were used. Without loss of generality, travel time is assumed to be the only component of generalised travel time; in practice, all of the analyses presented are easily adapted to include non-time attributes such as vehicle operating cost and tolls within the generalised travel times.

Clearly, a key element of these network specifications that is particular to the present paper concerns the assumptions regarding the probit path choice covariance matrix. Now, before proceeding, while it is not the main purpose of this paper to compare alternative equilibrium models, it might reasonably be asked what the PSUE model adds, compared to the already-known (and arguably, more straightforward to implement) results for DUE and logit SUE sensitivity analysis. In fact, since PSUE is able to approximate both such models, we may gain insight into the effect of such alternative model forms simply by varying the covariance structure of the probit model. As a preliminary step, then, to illustrate the alternative structural specifications that are possible with the probit model, we report some illustrative results here for the five-arc network.

As a first test, the impact is examined of including the additional random, unobserved terms; since, as the probit covariance matrix approaches the zero matrix, PSUE will approach DUE, this suggests examining the impact of increasing variance. Three cases, with arc covariance matrices of $\frac{1}{9}\Lambda$, $\frac{4}{9}\Lambda$, Λ (corresponding to proportionalities 0.1, 0.2, 0.3), are considered. For the sensitivity analysis, ε represents a scalar, namely a single additive perturbation to the travel time function on arc 3, so that $t_3(v_3,\varepsilon) = t_3(v_3) + \varepsilon$ and $t_a(v_a,\varepsilon) = t_a(v_a)$ $(a \neq 3)$.

The first three columns of Table 1 give the approximate linear sensitivity relationship between the PSUE arc flows and ε, for the three cases defined above. Evidently, as the random variation decreases in magnitude, so the arc flows become more sensitive to ε. This is to be expected since with lower variation, the deterministic component of travel time is more critical. This demonstrates that the random terms do not simply cancel; ultimately neglecting unobserved attributes (as in DUE) will lead to different sensitivities, and therefore different predicted impacts of a transport policy.

The comparison above establishes a case for including random terms, yet this still may be achieved (with much greater simplicity than for probit) by use of the logit SUE. Logit SUE, however, lacks the intuitive correlation in travel times between overlapping paths. To demonstrate how different sensitivities can arise if this correlation is neglected, a further comparison is made with a 'zero path correlation' case, obtained by re-setting all off-diagonal terms to zero in the implied *path* covariance matrix corresponding to arc covariance matrix Λ. Comparing the Λ and 'zero path correlation' columns in Table 1, there is evidence that less positive correlation implies less sensitivity; this is logical since in neglecting path overlaps, we suppose drivers are over-optimistic in believing there will be an attractive substitute path available when one path is perceived as unattractive. This demonstrates that the simplistic assumption of uncorrelated paths can give rise to misleading effects that neglect the network structure. In fact, logit is even more extreme than the probit 'zero path correlation' case considered here, since the former also requires the path disturbance terms to be identically distributed, whereas in the latter case we have still permitted differential variance terms across the diagonal.

Having made a case, therefore, for considering the computationally more complex probit SUE, the remainder of the paper will focus on this particular model.

4. APPLICATIONS OF PSUE SENSITIVITY EXPRESSIONS

4.1 Interpretations of sensitivity analysis

In section 2, techniques for *computing* a sensitivity analysis of the PSUE model were described; in section 4, we shall explore various *applications* of the results (6)–(10) through a series of numerical examples. The interpretations we shall give to sensitivity analysis in these applications are not, however, the only ones that are possible. Since an appreciation of our particular manner of interpretation is important in understanding the philosophy and motivation behind the numerical tests, we explicitly address this issue here.

In particular, the applications reported have a common theme of exploiting (6) as a linear model, relating the perturbation vector $\boldsymbol{\varepsilon}$ to the equilibrium flow vector $\mathbf{v}^*$. Two points can be made regarding this interpretation. **Firstly**, the underlying relationship between $\mathbf{v}^*$ and $\boldsymbol{\varepsilon}$ is certainly non-linear, and so (6) can only be regarded as an approximation to the true relationship. A strict interpretation of (6) is then that the approximation is only valid in a local neighbourhood of $\boldsymbol{\varepsilon} = \mathbf{0}$, yet our computational experience with the PSUE model has suggested that (at least for demand perturbations) it can hold as a reasonable approximation over a rather wider range: see Clark & Watling (2000, 2002), as well as the comparisons we shall implicitly make with the underlying non-linear relationship in sections 4.2 to 4.6.

It is our belief that there are two features of the PSUE model, not possessed by DUE, that particularly contribute to the strength of this approximation. These features are namely: (i) the probabilistic choice mechanism in PSUE has a smoothing effect relative to the 'discrete' nature of DUE, where effectively changes in the active path set lead to non-smooth behaviour at the DUE arc flow level, as parameter perturbations are made; and (ii) the large number of active paths, per origin-destination movement, in the PSUE sensitivity analysis effectively allows changes in the dominant path set (the paths carrying most of the demand flow) as $\boldsymbol{\varepsilon}$ is changed, effectively 'smoothly' mimicking the active path set changes in DUE. For example, a number of paths with a negligible PSUE flow at $\boldsymbol{\varepsilon} = \mathbf{0}$, but active in the sensitivity analysis, may attract a non-negligible PSUE flow as $\boldsymbol{\varepsilon}$ is moved far away from $\mathbf{0}$.

This latter feature can be empirically confirmed even when a sequence of PSUE solutions are estimated *by Monte Carlo methods*, for gradually changing values of $\boldsymbol{\varepsilon}$, specifically from the graph of some measure (path flow of significant paths, total travel time, …) evaluated at PSUE versus $\boldsymbol{\varepsilon}$—the reader is referred to the examples reported in Clark and Watling (2000, 2002). (Of course, this requires care in setting up the sequence of Monte Carlo runs with a common random number seed, initial condition and initial permissible/universal path set, and in allowing a large number of iterations to minimise convergence error). This is the case even when, given some slight change to $\boldsymbol{\varepsilon}$, the active PSUE path set arising from the Monte Carlo algorithm includes a previously-unused path or drops a previously-active path. The reason for this is that such paths were previously inactive in the Monte Carlo algorithm because they had an extremely low probability of being chosen, and even when active in the perturbed situation will continue to have an extremely low probability of being chosen, and hence an extremely small effect on the flows on the significant paths used in the network. Such behaviour would not be expected with a path flow analysis of models such as DUE, since in that case the active paths generated by an algorithms are entirely arbitrary (path flows non-unique); in contrast, in the DUE case, one cannot infer anything about the unattractiveness of the inactive paths, and so the argument used here for PSUE cannot be extended to the DUE model.

Returning, then, to the issues concerning the interpretation of the sensitivity analysis, the point we would make **secondly** is that at no stage shall we apply the interpretation of the sensitivity analysis as a *gradient* (or sub-gradient) of $\mathbf{v}^*(\boldsymbol{\varepsilon})$ at $\boldsymbol{\varepsilon} = \mathbf{0}$, as one may do, for example, in some applications to bi-level optimisation (Suh & Kim, 1992; Davis, 1994; Josefson & Patriksson, 2003; Patriksson, 2004). In particular, our algorithm in section 4.6 makes no explicit use of the sensitivity analysis as a gradient. Throughout the paper, our key requirement is that an approximate linear relationship exists between $\mathbf{v}^*$ and $\boldsymbol{\varepsilon}$; our concern will therefore be with the quality of this approximation relative to the underlying non-linear relationship, not with gradient properties. This is in truth a rather subtle point, since the case we prove for existence of the sensitivity analysis effectively can be used to establish it as a gradient, under some additional natural technical conditions. The point we are making, rather, is that the philosophy of the techniques we subsequently describe for utilising the sensitivity analysis do not exploit gradient properties; this is not intended as a technical point, so much as to explain our philosophy in using sensitivity analysis for wider issues than simply inferring gradient properties.

The specific applications described in the following sections 4.2–4.6 are derived from alternative interpretations of $\boldsymbol{\varepsilon}$, either as a given perturbation vector, a random vector with a given joint probability distribution, or a vector variable to be optimised. It is noted also that we shall use the different applications to illustrate the interpretation of $\boldsymbol{\varepsilon}$ as a perturbation to both the travel demand and the network characteristics. Numerical results are given for both the five-arc network with arc travel time disturbance matrix $\boldsymbol{\Lambda}$ as defined at the start of section 3, and for the Headingley network also defined in section 3.

4.2 Identification of sensitive parameters

Perhaps the most direct practical use of the sensitivity expressions is to gain an understanding of how responsive the arc flows within the network are to changes in the parameters. Where there are high sensitivities (in absolute terms) then there is a need to concentrate resources so as to ensure that such regions of the network are accurately described. We suppose here that ε is a vector representing an additive perturbation to the arc functions, so $t_a(v_a,\varepsilon) = t_a(v_a) + \varepsilon_\alpha \quad (a = 1,2,...,A)$. For example, for the test network of Figure 1, we then obtain:

(11) $$\begin{pmatrix} v_1^*(\varepsilon) \\ v_2^*(\varepsilon) \\ v_3^*(\varepsilon) \\ v_4^*(\varepsilon) \\ v_5^*(\varepsilon) \end{pmatrix} \approx \begin{pmatrix} 55.48 \\ 44.52 \\ 12.39 \\ 43.10 \\ 56.90 \end{pmatrix} + \begin{pmatrix} -2.67 & 2.67 & -2.17 & -0.51 & 0.51 \\ 2.67 & -2.67 & 2.17 & 0.51 & -0.51 \\ -2.17 & 2.17 & -4.92 & 2.76 & -2.76 \\ -0.51 & 0.51 & 2.76 & -3.26 & 3.26 \\ 0.51 & -0.51 & -2.76 & 3.26 & -3.26 \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{pmatrix} .$$

This matrix may be used to infer the importance of any errors that might be made by the model-user in estimating the arc travel time functions, in terms of the errors that would ensue in the resulting equilibrium arc flows. For example, suppose that in a particular policy context it was particularly important to accurately estimate the flow on arc 3, then it can be seen from the third row of the matrix in (11) that v_3^* is most sensitive to errors in ε_3, so that any additional resources which are available to estimate the travel time function parameters more accurately are wisely spent on those for arc 3.

4.3 Computation of approximate, re-equilibrated solutions ("post-optimisation")

There are many practical situations in which it is required to determine a number of traffic equilibrium solutions for problems with slightly modified input data; Nguyen *et al* (1996) cite examples ranging from on-line route guidance, to origin-destination matrix estimation from arc counts under equilibrium constraints on the flows. Determining the required equilibria sequentially, based on already-determined equilibria, has the potential for large computational savings.

Thus, one immediate application of (6) that may be envisaged is as a means of arriving at an approximate PSUE solution following some given change to the base input data. By way of illustration, for the five-arc network we write the origin-destination demand as $q = 100 + \varepsilon$ (here ε is a scalar), and give in Table 2 a comparison between the linear estimate and re-estimated PSUE solution for the case $\varepsilon = 10$. The agreement between the methods in terms of flows is good, and compares favourably with the computationally simpler logit-based results reported by Ying & Miyagi (2001). In our tests, while the computational times for both methods (linear estimation of PSUE and re-estimated PSUE) were small, the linear estimate took a fraction of the time needed for the re-estimated solution.

The main potential for computational savings arises in larger network applications. For illustration, in the Headingley network, a single origin-destination flow was increased by 10%, representing in practice (say) the impact of a small in-fill housing development. The solutions were compared from two methods: linear approximation from the unperturbed case, and re-estimation of the equilibrium. The network impact on arc flows is illustrated in Figure 2, with the width of the arc lines representing the absolute increase in flow relative to the base case. On a 450MHz Pentium II PC, the time to compute the sensitivity expressions was 14 minutes, but subsequently any number of linear-approximated equilibria resulting from changes to any demand or any travel time function could be computed in negligible time. On the other hand, re-computing an equilibrium took around 1 minute. Comparing the approximate and re-estimated solutions on an arc-by-arc level, the average percentage absolute difference in arc flows was 0.19%, with the largest error only 2.99%. The distribution of these differences is illustrated in Figure 3. On a network-wide level, the increase in demand led to a forecast increase in total network travel time of 1.124% by the linear approximation method, comparing favourably with the forecast 1.118% increase obtained by re-estimation of the equilibrium.

4.4 Estimating confidence intervals for equilibrium flows

The major items of input data to a traffic network model are the origin-destination trip matrix and arc travel time functions. In practice, such input data are prone to potentially large estimation errors, yet typically no effort is made to quantify the impact on the resulting prediction errors in the model outputs. In statistical terms, a sensible interpretation would seem to be that the conventional predictions of flows from an equilibrium model are *point estimates* of mean flows, but we may reasonably ask for standard errors and/or confidence intervals for these point predictions. For this purpose, we explore the use of (6) with both $\boldsymbol{\varepsilon}$ and its linear estimator (transformation) $\mathbf{v}^*(\boldsymbol{\varepsilon})$ assumed now to be random vectors, and with a given probability distribution representing the sampling error in $\boldsymbol{\varepsilon}$. For example, in the case of the origin-destination demand matrix, the sampling error represents the uncertainty in estimating the *mean* demands from some given survey data.

Three alternative techniques were tested for estimating output sampling distributions from given input sampling distributions for the trip matrix and link capacities:

- *Re-estimation:* Monte Carlo simulation with the full non-linear model: i.e. simulate from assumed input data sampling distributions, and solve a PSUE for each simulated scenario.
- *Linear simulation:* Monte Carlo simulation with linear approximation (6): i.e. simulate from the input data sampling distributions, then use sensitivity expressions (6)–(10).
- *Linear analytic*: By (6), $\mathrm{var}\left(\mathbf{v}^*(\boldsymbol{\varepsilon})\right) \approx \mathrm{var}\left(\mathbf{v}^*(0) + \mathbf{M}^{-1}\mathbf{N}\,\boldsymbol{\varepsilon}\right) = \mathbf{M}^{-1}\mathbf{N}\,\mathrm{var}(\boldsymbol{\varepsilon})\left(\mathbf{M}^{-1}\mathbf{N}\right)^{\mathrm{T}}$ is used to propagate standard errors; and by assuming Normality, confidence intervals are computed.

These approaches have alternative merits: re-estimation is the most computationally demanding, using the correct non-linear relationships but being subject to Monte Carlo error; linear simulation reduces the computational demand, but is subject to linearisation error in addition to Monte Carlo error; linear analytic is the most computationally attractive, but is subject to linearisation error, and must also assume a particular (e.g. Normal) output sampling distribution.

To illustrate each of these techniques, the five-arc network is considered with a (scalar) origin-destination demand of $q = 100 + \varepsilon$, where ε is a random variable representing the sampling error in the estimation of the mean demand level. In particular, we assume ε to be Normal with zero mean and variance 25 (such an assumption could be justified, for example, based on a Normal approximation to underlying Poisson demand variation, if the mean demand were estimated from $n = 4$ sampled observations). For both the re-estimation and linear simulation techniques, 400 realisations of q were sampled and the empirical 5% and 95% percentiles obtained from the resultant 400 sets of estimated PSUE link flow patterns corresponding to the sampled demand levels. With the linear analytic technique, the 5×5 arc-flow covariance matrix $\mathrm{var}(\mathbf{v}^{*}(\varepsilon))$ was calculated using the **M** and **N** matrices, with var(ε) = 25.

Table 3 shows how the percentiles compare across the three techniques. There is generally quite good agreement between all three methods. The re-estimation and linear simulation estimates show the greatest similarity, indicating that the linearisation error is relatively small; since they are based on the same Monte Carlo draws of the demand, the Monte Carlo effect can be neglected. Turning attention, then, to a comparison of the linear simulation and linear analytic techniques, while differences in the estimates produced by the two methods could in principle be attributable to either Monte Carlo error (in the linear simulation estimates) or violation of the Normal approximation (made in the linear analytic results), the latter source may be ruled out in the present case. This is due to the fact that demands have been sampled from a Normal distribution, and so link flows are linear combinations of Normal variables (under the linearised model) and so are themselves Normal. Thus, the difference in the results can be entirely attributed to Monte Carlo error in the linear simulation estimates. Taken with the earlier conclusions that linearisation error appears to be small, the overall implication is that one should place greatest confidence in the linear analytic results, the differences with the alternative two methods being attributable mainly to Monte Carlo error (this was also confirmed by re-applying the linear simulation method with an increased number of pseudo-random draws). While, clearly, the Monte Carlo error can always be reduced by increasing the number of random draws, there will always be the question of how many draws should be performed, an issue that does not arise with the linear analytic method.

Further tests were conducted on the three methods (re-estimation, linear simulation, linear analytic) in realistic networks under various assumptions on the demand levels, input variation levels, and input sampling distributions (Poisson, Normal, Lognormal). The detailed results are not given here, but overall a similar pattern in the results was evident to that reported above. In particular, the difference between the methods in estimating the limits of 95% link flow confidence intervals was seen to be less than 2% (based on 100 Monte Carlo replications), but both methods which used the linear sensitivity result gave an enormous saving in computational effort.

4.5 Estimating a confidence interval for total travel time

The approach described in section 4.4 may be further extended to quantify the uncertainty in what is a primary practical measure of network performance, total travel time. Suppose the arc functions are of a form consistent with the commonly used Bureau of Public Roads form (for other forms, a polynomial Taylor series approximation may be used):

(12) $$t_a(v_a) = \alpha_a + \beta_a v_a^{\,n_a} \qquad (n_a \geq 0 \text{ and integer; } \alpha_a \geq 0;\ \beta_a \geq 0;\ a = 1,2,\ldots,A)$$

then the mean and sampling variance in the total travel time may be written:

(13) $$\mathrm{E}\left[\sum_{a=1}^{A} v_a t_a(v_a)\right] = \sum_{a=1}^{A} \alpha_a \mathrm{E}[v_a] + \beta_a \mathrm{E}\left[v_a^{\,n_a+1}\right]$$

(14) $$\mathrm{var}\left(\sum_{a=1}^{A} v_a t_a(v_a)\right) = \sum_{a=1}^{A}\sum_{b=1}^{A} \alpha_a\alpha_b \operatorname{cov}(v_a, v_b) + \beta_a\alpha_b \operatorname{cov}(v_a^{\,n_a+1}, v_b) + \alpha_a\beta_b \operatorname{cov}(v_a, v_b^{\,n_b+1}) + \beta_a\beta_b \operatorname{cov}(v_a^{\,n_a+1}, v_b^{\,n_b+1}) \; .$$

Expressions (13) and (14) are used to exploit knowledge of the link flow covariance matrix from the linear analytic method of section 4.4. The flow variances and covariances on the right-hand side of (14) depend on flow moments and cross-moments, and supposing the joint link flow sampling distribution is approximately multivariate Normal, these required moments may be computed from standard expressions for multivariate normal moments (e.g. Isserlis, 1918). By a further Normality approximation for the total travel time sampling distribution, a confidence interval is thereby computed.

This technique was applied to the Headingley network, assuming there to be sampling error in the origin destination trip demands. Assuming, for each O-D movement *w*, Poisson variation in the underlying flows with mean q_w (approximated, for large q_w by a Normal distribution), then for a survey based on *n* independent observations of flows the sampling distribution for ε_w (sample mean flow for O-D movement *w*) would be approximately Normal with a variance of $\frac{q_w}{n}$. The tests reported here correspond to *n* = 1, a not unusual case in practice. In Figure 4, the base situation corresponds to a 'Demand Multiplier' of 1 on the horizontal axis. Other scenarios involve multiplying all O-D matrix elements in the base model by a constant factor (the Demand Multiplier), which by the Poisson assumption infers an increase by this factor in both the mean and variance in the underlying O-D flows. On the vertical axis, a corresponding 95% confidence interval in the total network travel time is illustrated. Clearly, the width of the interval varies with the demand, an illustration that the magnitude of the uncertainty in the model outputs is context-dependent.

In practice, such confidence intervals may be used in a before-and-after study of some hypothetical policy measure, in order to test whether the measure is forecast to lead to a statistically *significant* improvement in traffic conditions, in the light of the uncertainty in the forecasts that arises from the uncertainty in the input data. Thus statistical hypothesis testing may be integrated with the equilibrium analysis of networks, allowing conservative decisions to be made in the face of uncertainty in the model predictions. Furthermore, such techniques could also be applied in the context of survey design, such as the problem of determining the minimum sample size required to achieve a given level of precision. For example, in the application reported above, one question that could be addressed is: what sample size n is required when estimating the O-D demand levels, in order that total travel time may be estimated to within some given level of precision, at a given level of confidence.

On a technical level, it should also be noted that the simple normality assumption adopted above (for the total travel time distribution) is not critical to the analysis. It may reasonably be argued that total travel time is much more likely to follow some form of positively skew distribution, for example. In such a case, more general families of probability densities may be estimated using higher order moments, using elements of the techniques reported in Clark & Watling (2004).

4.6 Network Design

The ability to vary 'design parameters' of a network in order to optimise a network characteristic, while anticipating the response of drivers to the changed parameters, is commonly termed *network design*. Yang & Bell (1998) review the algorithms known and adopted by the transportation field for solving such problems. In the continuous network design problem, the design parameters may, for example, be road widths/capacities, traffic signal timings or road tolls. For illustration here, we re-visit an example considered by Suwansirikul *et al* (1987) for the DUE case, on the five-arc network of Figure 1. The travel time functions given in Figure 1 are re-written with constrained design variables $\boldsymbol{\varepsilon}$ (representing capacity changes) introduced in the flow divisor for each arc:

$$(15) \qquad t_a(v_a, \boldsymbol{\varepsilon}) = \alpha_a + \beta_a \left(\frac{v_a}{\kappa_a + \varepsilon_\alpha} \right)^4 \qquad (0 \le \varepsilon_a \le 4) \qquad (a = 1,2,\ldots, A)$$

so that for example, $\alpha_1 = 4, \beta_1 = 0.6, \kappa_1 = 40$. Here, the network design problem has as its upper level the criterion of total travel time plus a penalty term to reflect the cost of making changes to the design variables (with below, $d_1 = d_2 = d_4 = d_5 = 2$ and $d_3 = 1$):

$$(16) \qquad \text{Minimise } f(\boldsymbol{\varepsilon}) \equiv \sum_{a=1}^{5} \left\{ v_a t_a(v_a, \boldsymbol{\varepsilon}) + 1.5 d_a \varepsilon_\alpha{}^2 \right\} \text{ subject to } 0 \le \varepsilon_a \le 4 \ \ (a = 1,2,\ldots,5)$$

which is optimised in our case subject to a lower level PSUE (rather than DUE) relationship:

(17) **v** is a PSUE given $\boldsymbol{\varepsilon}$, based on arc functions (15).

As noted by Fisk (1984), this problem has the structure of a Stackelberg game (van Stackelberg, 1952), with the planner — who is responsible for network changes — acting as a 'leader', and the road users as a collection of 'followers' (albeit here with unobserved components represented as random variables in the road users' decision process). With (5) used to represent the lower level (17), the network design problem (16)/(17) has the form of a bi-level optimisation. The implicit, non-linear, lower level constraints (17) make this a demanding problem to solve, yet a number of authors have reported success in exploiting sensitivity analysis information in this context (Yang & Bell, 1998). One potential use of sensitivity analysis in this context is as a gradient function for the lower level (see the comments and references to such applications in section 4.1), but we shall continue with the interpretation of sensitivity analysis as a linear approximation model, particularly exploring whether a *single* sensitivity analysis could be sufficient to use throughout the course of the upper level optimisation. We test this hypothesis on the five-arc network by comparing the 'optimal' solution estimated by using a single linear approximation with that obtained by using the full, implicit PSUE relationship.

To solve the example problem, we used Powell's method (Powell, 1964) for the optimisation, a well-known gradient-free method, based on two alternative approaches for evaluating the objective function. In what we shall call the *re-estimation method*, for a single function evaluation $f(\varepsilon)$ at a given ε, we solve a full PSUE problem (17), and then substitute the resulting flows with the given ε into (16). (The random seed was re-set to the same value for each function evaluation, in order to minimise the potential for problems with Monte Carlo error in estimating the PSUE solution.) This method thus follows in the spirit (though not the specific approach) of the study of derivative-free techniques for the network design problem (Suwansirikul *et al*, 1987; Friesz *et al*, 1992). In contrast, with what we shall term the *linearised method*, to evaluate $f(\varepsilon)$ at a given ε we use the linear relationship (6) (as an approximation to (17)), and substitute the approximate flows with the given ε into (16); this follows in the spirit of the sensitivity-based techniques studied by Bell & Iida (1997). For this latter method, a *single* linear approximation ((6) evaluated at ε=0) is deduced before the optimisation commences. In both the re-estimation and linearised methods, the stopping criterion used was that the difference between objective function values over successive iterations should be less than 10^{-4}. An initial condition of ε=0 was used; other initial conditions were tested, but did not find other local optima.

Table 4 presents the results of these experiments. ε^{L} denotes the final solution obtained by the linearised method. This method crucially assumes the linear approximation about ε=0 to be valid across the whole feasible region; a check on whether this might be reasonable can be made by comparing the linear approximate and re-estimated PSUE flows at the final solution, ε^{L}. This comparison is given in the table, and a close correspondence is observed, indicating that the linearisation may be reasonable. The final solution from the re-estimation method, ε^{R}, took around 435 seconds to compute, compared with less than 1 second to compute ε^{L}. The objective function values produced by the various methods are also given;

and for information, the initial value of the objective function is $f(\mathbf{0}) = 1259.70$. The final values of the design variables from the two methods are extremely close, and the resulting value of the objective function almost identical. It is noted finally that the solutions obtained here were appreciably different, though the same order of magnitude, as the optimal values (1.35, 1.22, 0.00, 0.95, 1.08) obtained by Cho & Lo (1999) for the case of DUE.

5. CONCLUSION

The development of network equilibrium methods has been geared towards the efficient estimation of point solutions at given values of the input parameters. They are much less suitable for applications in which an explicit relationship is needed between variations in the input parameters and the model predictions. We have demonstrated that, by using a linear approximation to this explicit relationship, it is possible to address in a natural way problems of statistical inference and network optimisation. Direct applications of these techniques include, for example, the determination of optimal toll levels, hypothesis testing of before-and-after studies and survey design.

There are many possible directions for further developments of these methods. One such class of developments involves generalisation of the model, e.g. to account for elastic origin-destination demand (where the demands are a function of the generalized travel times), non-separable travel time functions, and multiple classes of traveller.

ACKNOWLEDGEMENTS We would like to thank two anonymous referees for their detailed and constructive comments on earlier drafts of this paper.

REFERENCES

Bell MGH & Iida Y (1997). *Transportation Network Analysis.* John Wiley &Sons, Chichester

Beckmann MJ, McGuire CB & Winsten CB (1956). *Studies in the Economics of Transportation.* Yale University Press, New Haven.

Blum JR (1954). Multidimensional Stochastic Approximation Methods. *Annals of Mathematical Statistics* **25**, 737-744.

Bureau of Public Roads (1964). *Traffic assignment manual.* US Department of Commerce.

Cantarella GE & Binetti MG (2002). Stochastic Assignment with Gammit Path Choice Models. In: *Transportation Planning: State of the Art* (ed. by M. Patriksson & M. Labbé), Kluwer, Dordrecht, Netherlands, pp. 53-68.

Cascetta E, Nuzzolo A, Russo F & Vitetta A (1996). A Modified Logit Route Choice Model: Overcoming Path Overlapping Problems. Specification and Some Calibration Results for Interurban Networks. *Transportation and Traffic Theory: Proceedings of the 13th International Symposium on Transportation and Traffic Theory,* ed J.B. Lesort, pp. 697-711, Elsevier Science, Oxford.

Chen A, Chang MS & Yang CY (2001a). Dynamic capacitated user-optimal departure time/ route choice problem with time-window. *European Journal of Operational Research* **132**, 603-618.

Chen A, Lo HK & Yang H (2001b). A self-adaptive projection and contraction algorithm for the traffic assignment problem with path-specific costs. *European Journal of Operational Research* **135**, 27-41.

Cho H-J & Lo SC (1999). Solving Bilevel Network Design Problem Using a Linear Reaction Function Without Nondegeneracy Assumption. *Transportation Research Record* **1667**, 96-106.

Cho H-J, Smith TE & Friesz TL (2000). A reduction method for local sensitivity analyses of equilibrium arc flows. *Transportation Research* **34B**, 31-51.

Clark SD & Watling DP (2000). Probit Based Sensitivity Analysis for General Traffic Networks. *Transportation Research Record* **1733**, 88-95.

Clark SD & Watling DP (2002). Sensitivity analysis of the probit-based stochastic user equilibrium assignment model. *Transportation Research* **36B**, 617-635.

Clark SD & Watling DP (2004). Modelling network travel time reliability under stochastic demand. *Transportation Research B*, in press.

Daganzo CF (1979). *Multinomial Probit: The Theory and its Application to Demand Forecasting.* Academic Press, New York.

Daganzo C & Sheffi Y (1977). On Stochastic Models of Traffic Assignment. *Transportation Science* **11**(2), 253-274.

Davis GA (1994). Exact local solution of the continuous network design problem via stochastic user equilibrium. *Transportation Research* **28B**, 61-75.

Fiacco AV (1983). *Introduction to Sensitivity and Stability Analysis in Non-linear Programming.* Academic Press, New York.

Fisk C (1980). Some Developments in Equilibrium Traffic Assignment. *Transportation Research* **14B**(3), 243-255.

Fisk CS (1984). Game Theory and Transportation Systems Modelling. *Transportation Research* **18B**(4/5), 301-313.

Frank M & Wolfe P (1956). An algorithm for quadratic programming. *Naval Research Logistics Quarterly* **3**(1/2), 95-110.

Friesz TL, Cho H-J, Metha NJ, Tobin RL & Anandalingam G (1992). A simulated annealing approach to the network design problem with variational inequality constraints. *Transportation Science* **26**, 18-26.

Gentile G & Papola N (2001). Network design through sensitivity analysis and singular value decomposition. Paper presented at TRISTAN IV, San Miguel, Azores, June 13th–19th 2001.

Gliebe JP, Koppelman FS & Ziliaskopoulos A (1999). Route choice using a paired combinatorial logit model. Paper presented at the 78th meeting of the Transportation Research Board, Washington D.C.

Isserlis L (1918). On a Formulation for the Product-Moment Co-efficient of any Order of a Normal Frequency Distribution in any Number of Variables. *Biometrica* **12**(1/2), 134-139.

Josefson M & Patriksson M (2003). Sensitivity analysis of separable traffic equilibria, with application to bilevel optimization in network design. Department of Mathematics, Chalmers University of Technology, Gothenburg, Sweden, submitted for publication.

Nash JF (1951). Noncooperative Games. *Annals of Mathematics* **54**, 286-298.

Nguyen S & Dupuis C (1984). An efficient method for computing traffic equilibria in networks with asymmetric transportation costs. *Transportation Science* **18**, 185-202.

Nguyen S, Pallottino S & Inaudi D (1996). Postoptimizing equilibrium flows on large scale networks. *European Journal of Operational Research* **91**, 507-516.

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

Patriksson M (2004). Sensitivity Analysis of Traffic Equilibria. *Transportation Science*, forthcoming.

Patriksson M & Rockafellar RT (2003). Sensitivity Analysis of Aggregated Variational Inequality Problems, with Application to Traffic Equilibria. *Transportation Science* **37**(1), 56-58.

Powell MJD (1964). An efficient method of finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal* **7**, 155-162.

Powell WB & Sheffi Y (1982). The Convergence of Equilibrium Algorithms with Predetermined Step Sizes. *Transportation Science* **16**(1), 45-55.

Prashker JN & Bekhor S (1999). Stochastic User-Equilibrium Formulations for Extended-Logit Assignment Models. *Transportation Research Record* **1676**, 145-152.

Sheffi Y (1985). *Urban Transportation Networks.* Prentice Hall, New York.

Sheffi Y & Powell WB (1982). An Algorithm for the Equilibrium Assignment Problem with Random Link Times. *Networks* **12**(2), 191-207.

Suh S & Kim TJ (1992). Solving nonlinear bilevel programming models of the equilibrium network design problem: A comparative review. In *Hierarchical Optimization* (ed. by G Anandalingam & TL Friesz), JC Baltzer, Basel, Switzerland.

Suwansirikul C, Friesz TL & Tobin RL (1987). Equilibrium Decomposed Optimization: A Heuristic for the Continuous Equilibrium Network Design Problem. *Transportation Science* **21**(4), 254-62

Tobin RL & Friesz TL (1988). Sensitivity Analysis for Equilibrium Network Flow. *Transportation Science* **22**(4), 242-250.

Van Stackelberg H (1952). *The Theory of Market Economy*. Oxford University Press, Oxford.

Vovsha P & Bekhor S (1998). Link-Nested Logit Model of Route Choice Overcoming Route Overlapping Problem. *Transportation Research Record* **1645**, 133-141.

Wie B-W (1995). A differential game approach to the dynamic mixed behavior traffic network equilibrium problem. *European Journal of Operational Research* **83**, 117-136.

Yai T, Iwakura S & Morichi S (1997). Multinomial probit with structured covariance for route choice behavior. *Transportation Research* **21B**(3), 195-207.

Yang H & Bell MGH (1998). Models and algorithms for road network design: a review and some new developments. *Transport Reviews* **18**, 257-278.

Ying JQ & Miyagi T (2001). Sensitivity Analysis for Stochastic User Equilibrium Network Flows – A Dual Approach. *Transportation Science* **35**(2), 124-133.

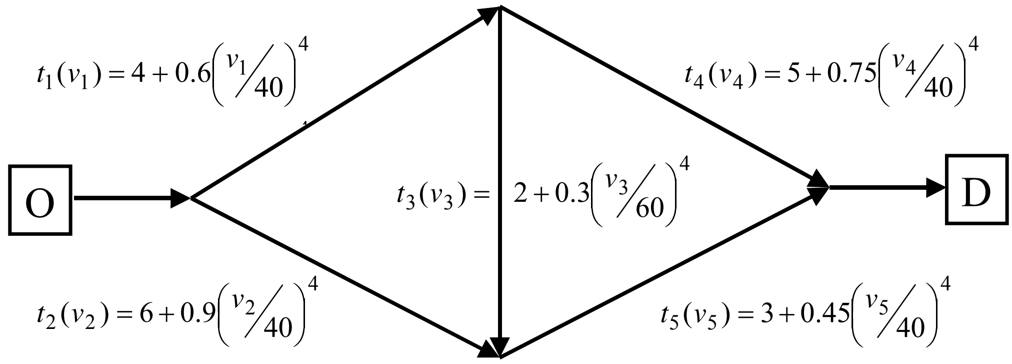

Figure 1: Five-arc example network (Suwansirikul *et al*, 1987), OD demand $q = 100$

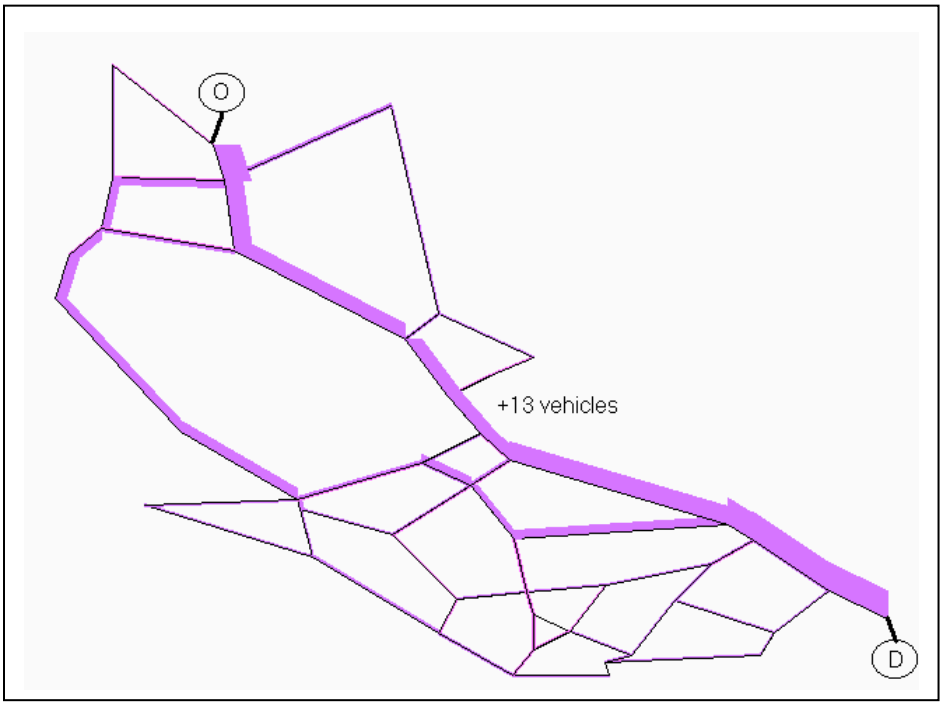

Figure 2: Diagrammatic representation of increase in arc flows following increase in an origin-destination flow (Headingley network)

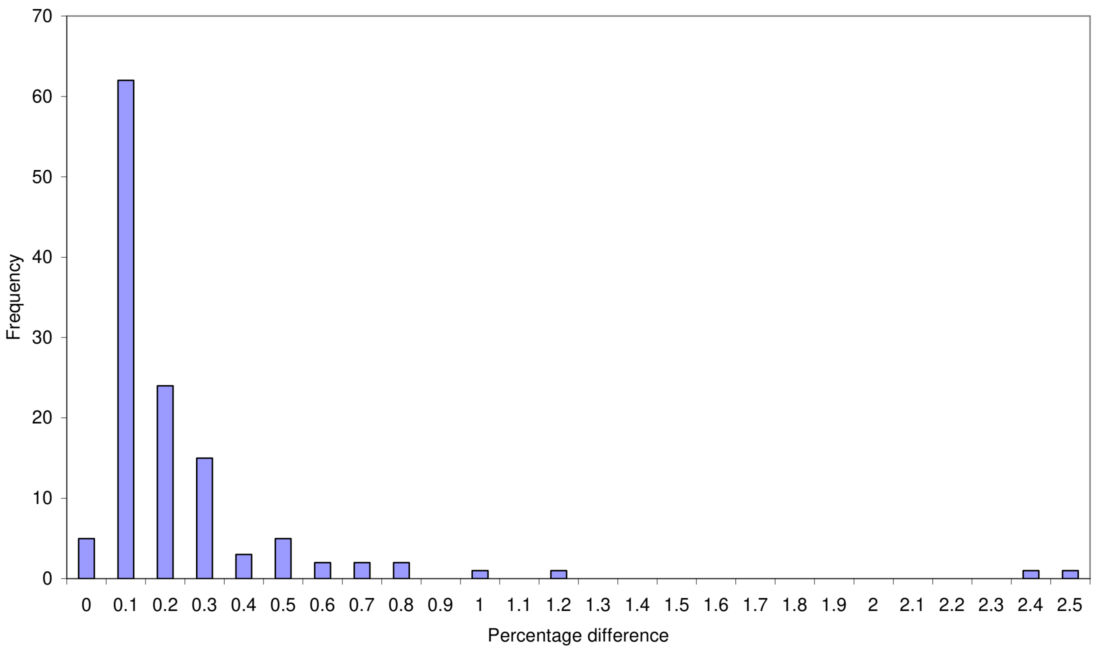

Figure 3: Arc-by-arc comparison, distribution of percentage differences in arc flows between linear appoximate and re-estimated solutions (Headingley)

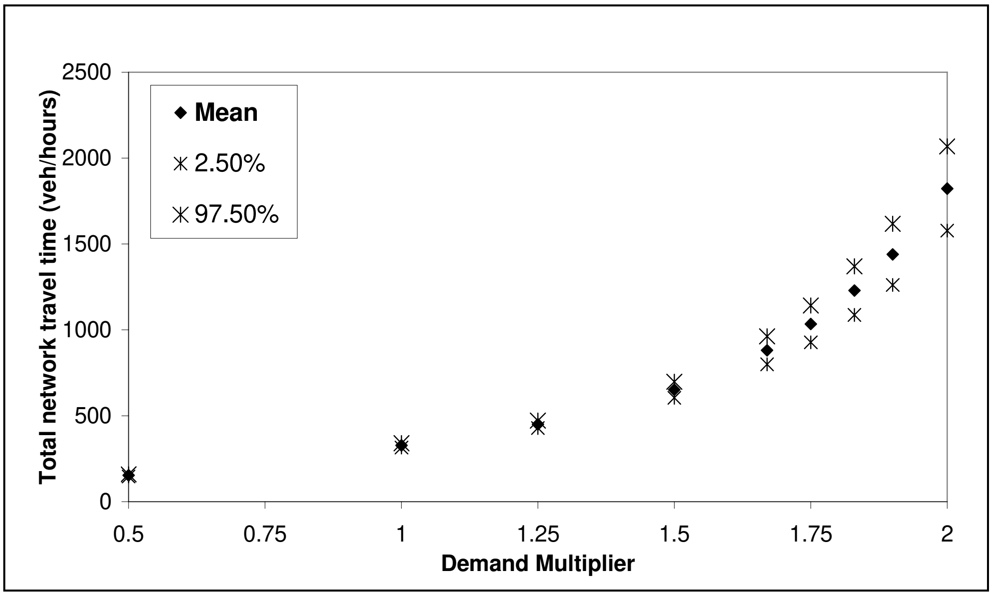

Figure 4: 95% confidence interval in total network travel time (Headingley network, O-D matrix uniformly factored by 'Demand Multiplier')

Arc a	Λ	$\frac{4}{9}\Lambda$	$\frac{1}{9}\Lambda$	Zero path correlation
1	$55.48 - 2.1650\,\varepsilon$	$54.79 - 2.4185\,\varepsilon$	$53.89 - 2.7823\,\varepsilon$	$56.62 - 2.0620\,\varepsilon$
2	$44.52 + 2.1650\,\varepsilon$	$45.21 + 2.4185\,\varepsilon$	$46.11 + 2.7823\,\varepsilon$	$43.38 + 2.0620\,\varepsilon$
3	$12.39 - 4.9211\,\varepsilon$	$10.82 - 5.4996\,\varepsilon$	$8.77 - 6.3242\,\varepsilon$	$14.54 - 4.6292\,\varepsilon$
4	$43.10 + 2.7558\,\varepsilon$	$43.97 + 3.0811\,\varepsilon$	$45.12 + 3.5419\,\varepsilon$	$42.08 + 2.5672\,\varepsilon$
5	$56.90 - 2.7558\,\varepsilon$	$56.03 - 3.0811\,\varepsilon$	$54.88 - 3.5419\,\varepsilon$	$57.92 - 2.5672\,\varepsilon$

Table 1: Linear sensitivity estimate of PSUE arc flows as a function of additive perturbation ε to the travel time function for arc 3, for alternative covariance assumptions (five-arc network)

Arc a	Base $v_a^*(0)$	Flow change per unit ε	Linear approx. to $v_a^*(10)$	Re-estimation of $v_a^*(10)$	Percentage difference
1	55.4843	0.4818	60.3023	60.3041	0.00%
2	44.5157	0.5182	49.6977	49.6959	0.00%
3	12.3875	-0.0408	11.9795	11.9900	-0.09%
4	43.0970	0.5226	48.3230	48.3140	0.02%
5	56.9030	0.4774	61.6770	61.6860	-0.01%

Table 2: Linear approximate and re-estimated solution for an $\varepsilon = 10$ unit increase in q (five-arc network)

Arc (a)	5% percentile			95% percentile		
	Re-estimated	Linear simulated	Linear analytic	Re-estimated	Linear simulated	Linear analytic
1	52.06	52.01	51.58	60.07	60.06	59.39
2	40.74	40.78	40.19	49.44	49.44	48.84
3	12.01	12.00	11.93	12.75	12.68	12.84
4	39.31	39.33	38.74	48.06	48.06	47.45
5	53.49	53.47	53.03	61.45	61.44	60.77

Table 3: 90% Confidence intervals for PSUE arc flows (five-arc network)

	Linearised method			Re-estimated method	
Arc (a)	Final solution $\boldsymbol{\varepsilon}^{\mathrm{L}}$	Approximated flows at $\boldsymbol{\varepsilon}^{\mathrm{L}}$	Re-estimated flows at $\boldsymbol{\varepsilon}^{\mathrm{L}}$	Final solution $\boldsymbol{\varepsilon}^{\mathrm{R}}$	Re-estimated flows at $\boldsymbol{\varepsilon}^{\mathrm{R}}$
1	1.4672	55.8933	55.8968	1.4469	55.8719
2	1.0538	44.1067	44.1032	1.0804	44.1281
3	0.0044	13.1704	13.1813	0.0002	13.1313
4	0.8086	42.7181	42.7157	0.8433	42.7406
5	1.2256	57.2819	57.2843	1.2017	57.2594
$f(\varepsilon)$		1214.39	1214.68		1214.39

Table 4: Network design solutions for re-estimation and linearised methods (five-arc network)