Article:

Clark, S.D. and Watling, D.P. (2002) Sensitivity analysis of the probit-based stochastic user equilibrium model. Transportation Research B : Methodological, 36 (7). pp. 617-635. ISSN 0109-2615

https://doi.org/10.1016/S0191-2615(01)00021-2

Published paper
Clark, S.D.; Watling, D.P. (2002) *Sensitivity Analysis of the Probit-Based Stochastic User Equilibrium Model*. Transportation Research Part B: Methodological 36(7) 617-635

SENSITIVITY ANALYSIS OF THE PROBIT-BASED STOCHASTIC USER EQUILIBRIUM ASSIGNMENT MODEL

Stephen D. Clark and David P. Watling,

Institute for Transport Studies,

University of Leeds, UK.

(email: sclark@its.leeds.ac.uk**,** dwatling@its.leeds.ac.uk**)**

Revised version: 30/4/01

ABSTRACT

The probit-based Stochastic User Equilibrium (SUE) model has the advantage of being able to represent perceptual differences in utility across the driver population, while taking proper account of the natural correlations in these utilities between overlapping routes within the network (which the simpler logit SUE is unable to do). Its main drawback is the potentially heavy computational demands, and this has previously been thought to preclude a consideration of the *sensitivity analysis* of probit-based SUE, whereby an approximation to changes in the equilibrium solution is deduced as its input parameters (specifically origin/destination flows and link cost-flow function parameters) are perturbed. In the present paper, an efficient computational method for performing such an analysis for general networks is described. This approach uses information on SUE path flows, but is not specific to any particular equilibrium solution algorithm. Problems inherent in the consideration of general network topologies are identified, and methods proposed for overcoming them. The paper concludes with an application of the method to a realistic network, and compares the approximate solutions with those obtained by direct estimation methods.

1. INTRODUCTION

The basic function of a traffic assignment model is to relate various input data (e.g. origin-destination demand matrix, network topology, capacities, signal timings) to various output measures (e.g. link flows, travel times). In practice, the approach typically adopted is: (i) to 'calibrate' the input data/parameters so as the model reproduces currently observed traffic conditions; and then (ii) to test a number of alternative hypothetical policies by adjusting the input data, and re-running the model for each policy. Since each policy option requires a new model run and output analysis, time limitations mean that it is rare for anything more than a small number of alternatives to be tested. The model therefore has a rather passive role, providing limited help to the planner in understanding the relationship between input data (including potential policies) and output measures.

In the present paper, we shall explore the technique of *sensitivity analysis*, which aims to extract explicit information on the relationship noted above, by postulating simple approximating functions to describe the impact of changes in the model inputs on changes in the model outputs. Given its long history in the transport research literature, it is surprising that this technique seems to have been little exploited in practice.

Specifically, the aim will be to deduce expressions for a first order sensitivity analysis of an equilibrium assignment model, yielding a linear relationship between the equilibrium flows and the input data. Although it may seem an unusual step, to approximate such an intrinsically non-linear system as a traffic assignment process by a linear model, the claim is only that such an approximation will be valid in the neighbourhood of some 'nominal' solution (this latter computed from a standard run of the equilibrium model). The advantage of such an approach is that it provides information on a range of hypothetical scenarios and their related equilibrium solutions. The potential applications include:

Identification of 'critical' variables, such as identifying the most sensitive links to demand or capacity changes (see, e.g., Yang, 1997).

Error analysis, whereby the propagation of given input sampling distributions to output error distributions may be traced through the sensitivity expressions (Leurent, 1996, 1998; Bell & Iida, 1997), thus quantifying the impact of measurement/sampling errors. Ultimately, in addition to point estimates such as the mean, the approach can be used to provide confidence intervals.

Optimal design. One particularly fruitful way of exploiting the sensitivity expressions is to embed them in a solution algorithm (particularly for bi-level problems), with the aim to determine optimal values of some design variables, subject to the constraint that the resulting flows are in equilibrium (Magnanti & Wong, 1984; Davis, 1994; Yang, 1997; Bell & Iida, 1997). Example applications of such an approach include optimal pricing (Yang & Bell, 1997) and traffic control (Yang *et al.*, 1994).

Estimation problems, a particular example being the problem of origin-destination matrix estimation from link counts (Yang *et al.*, 1992; Verlander & Heydecker, 1994). Again, as in the optimal design problem, sensitivity analysis may be used to approximate a user equilibrium constraint, as part of a solution algorithm.

The paper will begin with a review of previous work on sensitivity analysis of the traffic assignment problem. In section 3, a result is stated for a first order sensitivity approximation for a general non-linear program, and in section 4 this result is subsequently applied to the stochastic user equilibrium assignment problem. Section 5 provides a method for calculating the first derivatives for the probit choice probabilities, the most complex component of the sensitivity expressions. The problem of linear dependence between routes is explained in section 6, and methods for overcoming it investigated. Section 7 provides a case-study application of the methods to a realistic network.

2. REVIEW OF SENSITIVITY ANALYSIS FOR TRAFFIC ASSIGNMENT

Following early investigations by Hall (1978) and Dafermos & Nagurney (1984), which essentially established the "direction of change" following perturbations to inputs of a traffic assignment model, a major breakthrough was made with the work of Tobin & Friesz (1988). Exploiting recent results which formulated the deterministic user equilibrium (DUE) model as a variational inequality (Smith, 1979; Dafermos, 1980), Tobin & Friesz developed theoretical results and computational procedures for estimating the magnitude of the sensitivities for general networks. In particular, they examined the derivatives of the equilibrium link flows under perturbations to the elements of the origin-destination demand matrix and/or link cost function parameters. The primary advantage of the variational inequality approach is that it permits the travel cost on a given link to depend not only on the flow on that given link, but also (in a restricted way) on the flows on *all* links[1].

The major hurdle Tobin & Friesz faced was the well-known non-uniqueness of DUE path flows, even in cases where the DUE link flows are unique (e.g. strictly monotone cost-flow relationship, as in Smith, 1979). This difficulty arises even if there is no underlying interest in the perturbed path flows, since it is not possible to write the DUE feasibility constraints in terms of link flows only, without reference to path flows. They overcame this difficulty by

---

[1] As noted by Tobin & Friesz, a parallel derivation of their results is possible in the special cases where DUE may be formulated as an equivalent optimization problem, such as cases where the link travel cost/flow Jacobian is diagonal or symmetric.

demonstrating that a restricted formulation of the DUE problem did possess the required uniqueness property. The solution they adopt is as follows:

1. Choose an arbitrary DUE path flow solution, satisfying two conditions, namely that the solution is:
   (a) an extreme point of the path flow feasible region; and
   (b) non-degenerate, namely that the number of paths with positive flow must be equal to the rank of $[\Delta^{T}|\ \Lambda^{T}]$, where Δ^{T} is the transpose of the path-link incidence matrix, and Λ^{T} is the transpose of the origin/destination-path incidence matrix.
2. Compute the sensitivities of the path flows, assuming that in the perturbed state the only used paths are those that had positive flow in the unperturbed state selected in 1. (the "active paths").
3. Hence, by transformation, deduce the sensitivities in terms of link flows.

Tobin & Friesz go on to show that the link sensitivities resulting in 3 are independent of the arbitrary path flow solution selected in 1, but only for a 'restricted problem' where additional conditions hold. These conditions are namely that the perturbed variational inequality is strictly monotone, and that a strict complementarity condition holds (implying that only links with positive flow in the unperturbed solution need be considered). They do not, however, examine under what conditions it is reasonable to assume these additional assumptions to be valid. The technique was illustrated with a small example network, but the feasibility of the method for large realistic networks was not examined.

The approach of Tobin & Friesz has subsequently been extended to examine the sensitivity analysis of a number of generalisations of the DUE model, including a steady state queuing version of DUE (Yang, 1995), the elastic demand DUE model (Yang, 1997), and a DUE model with randomly distributed values of time (Leurent, 1998). In addition, Bell & Iida

(1997) considered the logit-based stochastic user equilibrium (SUE) assignment model, and showed how sensitivity expressions may be derived from an equivalent optimization problem, or by the Tobin & Friesz approach.

In conclusion, existing work in the literature on traffic assignment sensitivity analysis derives almost exclusively from applications and extensions of Tobin & Friesz's results. In any such DUE-style model, a major hurdle is the non-uniqueness of the equilibrium path flows, meaning that a restricted problem needs to be defined, and additional assumptions made, in order to derive sensitivity expressions. A major advantage of the SUE model, however, is that it is known to give rise to unique path flows under mild conditions (Sheffi, 1985). It also has an advantage over DUE of greater behavioural realism, in that DUE assumes the utilities of each available path to be known and identically perceived across the population. In reality, taste variation means that this premise is unlikely to be true, and a more appropriate assumption is that the utilities contain a random element. Moreover, a limiting case of SUE—as the perceptual dispersion in the population approaches zero—is the DUE model, and so the former model may be regarded as more general, with DUE a special case. Effectively, if our underlying interest were in DUE, we could even regard SUE with a small perceptual variance as an alternative means of defining a "restricted problem" with unique path flows.

There is therefore a good case for considering sensitivity analysis of the SUE problem. As demonstrated by Bell & Iida (1997), this is relatively straightforward for the case where choice fractions are assumed to follow a multinomial logit model, yet there are well-known deficiencies with using such a model in a network context. In particular, by assuming path utilities to be statistically independent, the model neglects what could be claimed to be the most important structural element of a network: namely that paths are formed from links. For example, two paths that overlap for virtually their whole length are likely to be perceived very similarly by an individual since they have many links in common, but this cannot be captured by a logit model. The probit SUE model (e.g. Sheffi, 1985) is able to overcome these drawbacks, by supposing path utilities (cost) are formed from a sum of link utilities (costs), with the error distribution specified for the latter. The disadvantage of the probit model is its computational complexity: there is no closed form expression for probit choice fractions, and evaluating them directly involves the evaluation of a multivariate normal integral, of dimension equal to the number of feasible paths (for a given origin-destination movement). This appears to have precluded a previous consideration of probit SUE sensitivities.

The purpose of the remainder of the paper will therefore be to deduce an efficient procedure for such a sensitivity analysis of the probit SUE model. A particular goal is that this procedure is feasible for large realistic networks, and so some attention is paid to selecting a computationally efficient method. It is notable that the previous sensitivity work reviewed in this section has largely been reported in applications to small illustrative examples, with little direct evidence of the computational efficiency of the methods for larger networks.

3. METHODS OF DERIVING SENSITIVITY EXPRESSIONS

For the DUE model, there are two main ways of deriving the sensitivity expressions: either through an equivalent optimization (EO) or a variational inequality (VI) problem. The main distinction is that the VI approach can be said to be more general in that it permits mild interactions between links in the cost-flow relationships. The benefit of this generalisation is, however, somewhat diminished by the fact that it is rather difficult to test for in practice (Heydecker, 1983), with the possibility of multiple link flow equilibria existing if it is just violated (Watling, 1996). Therefore, in most practical situations, regardless of whether the EO or VI approach is applied, one commonly has in mind the same class of problem, namely that with separable, increasing cost-flow functions.

For the SUE model, in addition to an EO and VI formulation, one also has the third possibility of a fixed point (FP) formulation. Again, there are no great computational differences between the three approaches, the FP formulation having the claim to admitting the greatest generality of problems, requiring only differentiability of the involved functions to provide sensitivity estimates. This could, however, be said to be a little misleading, since there are implicit smoothness properties assumed when applying sensitivity analysis to any kind of formulation. For this reason, in our subsequent study of SUE, we have chosen to use the EO formulation. Desirable properties of the objective function involved have been established by Sheffi (1985), such as convexity in the neighbourhood of an SUE solution, which we believe means that a Taylor series expansion may be used with greater confidence. In contrast, the VI approach proposed by Ran and Boyce (1993) has had relatively little attention paid to it. In any case, we believe it is superseded for general non-separable problems by the FP formulation. For completeness, in Appendix 1, a derivation of the sensitivity expressions for the FP formulation is presented. It is shown that in the special case where link cost-flow functions are separable and increasing (required for the EO formulation to be valid), then the EO and FP formulations provide equivalent sensitivity expressions.

Henceforth, we shall therefore restrict attention to problems expressible as an equivalent optimization. As a precursor to the application to traffic assignment, the present section therefore summarises a result for the first-order sensitivity approximation of a general non-linear optimization problem, applicable when either the constraints or parameters in the objective function are perturbed. In particular, for a given vector of perturbations $\boldsymbol{\varepsilon}$ (of length equal to the number of perturbations), consider the problem:

Minimize $f(\mathbf{x},\boldsymbol{\varepsilon})$ with respect to $\mathbf{x}$ subject to

$$g_i(\mathbf{x},\boldsymbol{\varepsilon}) \geq 0 \qquad (i=1,2,\ldots,m)$$

$$h_j(\mathbf{x},\varepsilon) = 0 \qquad (j=1,2,\ldots,p)$$

A first-order Taylor series approximation to the solution to this problem, as $\boldsymbol{\varepsilon}$ is varied, in the neighbourhood of the original solution is then (Fiacco, 1983, pp 76-77):

$$\begin{bmatrix} \mathbf{x}(\varepsilon) \\ \mathbf{u}(\varepsilon) \\ \mathbf{w}(\varepsilon) \end{bmatrix} = \begin{bmatrix} \mathbf{x}(\mathbf{0}) \\ \mathbf{u}(\mathbf{0}) \\ \mathbf{w}(\mathbf{0}) \end{bmatrix} + \left(\mathbf{M}(\mathbf{0})^{-1} \mathbf{N}(\mathbf{0}) \right) \varepsilon \quad + \mathrm{o}(\|\varepsilon\|)$$

where $\mathbf{x}(\boldsymbol{\varepsilon})$ is the solution vector (dimension n) at $\boldsymbol{\varepsilon}$;

$\mathbf{u}(\boldsymbol{\varepsilon})$ is the m-vector of Lagrangian non-negativity multipliers at $\boldsymbol{\varepsilon}$;

$\mathbf{w}(\boldsymbol{\varepsilon})$ is the p-vector of Lagrangian equality multipliers at $\boldsymbol{\varepsilon}$;

$\mathrm{o}(\|\boldsymbol{\varepsilon}\|)$ represents a real valued function, $\phi(\boldsymbol{\varepsilon})$, such that $\phi(\boldsymbol{\varepsilon})/\|\boldsymbol{\varepsilon}\| \to 0$ as $\boldsymbol{\varepsilon} \to \mathbf{0}$

and where the matrices **M** and **N** (as functions of $\boldsymbol{\varepsilon}$) are given by:

$$\mathbf{M}(\varepsilon) = \begin{bmatrix} \nabla^2 L & -\nabla g_1^T & \cdots & -\nabla g_m^T & \nabla h_1^T & \cdots & \nabla h_p^T \\ u_1 \nabla g_1 & g_1 & & 0 & & & \\ \vdots & & \ddots & & & 0 & \\ u_m \nabla g_m & 0 & & g_m & & & \\ \nabla h_1 & & & & & & \\ \vdots & & 0 & & & 0 & \\ \nabla h_p & & & & & & \end{bmatrix} \qquad \mathbf{N}(\varepsilon) = \begin{bmatrix} -\left(\nabla^2_{\varepsilon \mathbf{x}} L\right)^T \\ -u_1 \nabla_\varepsilon g_1^{\;T} \\ \vdots \\ -u_m \nabla_\varepsilon g_m^{\;T} \\ -\nabla_\varepsilon h_1^T \\ \vdots \\ -\nabla_\varepsilon h_p^T \end{bmatrix}$$

where L is the Lagrangian function for the problem.

4. SENSITIVITY ANALYSIS OF STOCHASTIC USER EQUILIBRIUM

An equivalent optimization formulation of the general SUE traffic assignment problem was established by Sheffi and Powell (1982):

$$\underset{\mathbf{x}}{\mathrm{Min}}\; z(\mathbf{x}) = -\sum_{rs} q_{rs}\, E\left[\min_{k \in \kappa(r,s)} \left\{ C_k^{rs} \right\} \middle| \mathbf{c}^{rs}(\mathbf{x}) \right] + \sum_a x_a t_a(x_a) - \sum_a \int_0^{x_a} t_a(\omega)\, d\omega$$

where q_{rs} is the total O-D flow between origin r and destination s;

C_k^{rs} is the perceived travel cost on route k between O-D pair r - s;

c^{rs} is the actual travel cost on route k between O-D pair r-s;

x_a is the flow on link a;

$t_a(x_a)$ is the travel cost on link a, assumed an increasing function of x_a only;

and $\kappa(r,s)$ is the set of routes connecting O-D pair r-s.

A notable feature of this formulation is that it is unconstrained, the solution automatically satisfying the flow conservation and non-negativity of path flows constraints. Thus, when applying the Fiacco sensitivity results to this problem, the **M** and **N** matrices defined in section 3 require only a consideration of $\nabla^2 z(\mathbf{x})$ (i.e. $\nabla^2 L$) and $\nabla^2_{\varepsilon x} z(\mathbf{x})$ (i.e.$\nabla^2_{\varepsilon x} L$). An expression for $\nabla z(\mathbf{x})$, and thence $\nabla^2 z(\mathbf{x})$, is readily deduced as (Sheffi, 1985, pp 318-319):

$$\nabla z(\mathbf{x}) = [\ -\Sigma_{rs}\ q_{rs}\ \mathbf{P}^{rs}\ \Delta^{rsT} + x\ ]\ .\nabla_x\ \mathbf{t}$$

$$\nabla^2 z(\mathbf{x}) = \Sigma_{rs}\ q_{rs}\ [(\nabla_x \mathbf{t}\ .\ \Delta^{rs})\ (-\nabla_c\ \mathbf{P}^{rs})\ (\nabla_x \mathbf{t}\ .\ \Delta^{rs})^T\ ] + \nabla_x\ \mathbf{t} + \nabla_x{}^2\ \mathbf{t}\ .\ \mathbf{R}$$

where $\nabla_x \mathbf{t}$ is the Jacobian of the link travel cost vector;

Δ^{rs} is the link-path incidence matrix for O-D pair r-s;

$\nabla_c\ \mathbf{P}^{rs}$ is the Jacobian of the route choice probability vector for O-D pair r-s;

$\mathbf{R}$ is a diagonal matrix, the a^{th} element of which is $(-\Sigma_{rs}\Sigma_k q_{rs} P_k{}^{rs} \delta_{a,k}{}^{rs} + x_a)$

$P_k{}^{rs}$ is the probability of using path k for O-D pair r-s;

and $\delta_{a,k}{}^{rs}$ is 1 if link a is on path k between O-D pair r-s and 0 otherwise.

The expressions above yield Fiacco's **M** matrix. The form of Fiacco's **N** matrix depends on the particular form of perturbations considered, and two cases are presented here:

Case (i) Changes in the O-D demand elements

$$\nabla_{x\varepsilon} z(\mathbf{x}) = [\ -\Sigma_{rs}\ \nabla_\varepsilon\ q_{rs}(\varepsilon)\ \mathbf{P}^{rs}\ \Delta^{rsT}\ ]\ .\nabla_x\ \mathbf{t}$$

Case (ii) Changes in the link cost function parameters

$$\nabla_{x\varepsilon} z(\mathbf{x}) = \Sigma_{rs}\ q_{rs}\ [(\nabla_x \mathbf{t}\ .\ \Delta^{rs})\ (-\nabla_c\ \mathbf{P}^{rs})\ \nabla_\varepsilon\ c^{rs}(\varepsilon)]$$

where $q_{rs}(\varepsilon)$ is the relationship between the O-D demand and the vector of perturbations;

and $c^{rs}(\varepsilon)$ is the relationship between the vector of route costs and the vector of perturbations.

Since the Jacobian of the link travel cost vector is relatively easily computed for common forms of performance function, the only challenging component of these expressions is the Jacobian of the choice probability vector ($\nabla_c \mathbf{P}^{rs}$). For the logit model, this can readily be obtained in analytic form, but since no closed form expressions for probit choice probabilities exist, alternative methods are required.

5. COMPUTATION OF THE PROBIT ROUTE CHOICE JACOBIAN

As described in Daganzo (1979), an efficient method for approximating probit route choice probabilities is by Monte Carlo simulation. That is to say, one performs repeated pseudo-random simulations from the given multivariate normal error structure. By the law of large numbers, as the number of replications approaches infinity, the fraction of times that an alternative has maximum utility will approach its choice probability. Such a method underlies the stochastic network loading step in the method of successive averages equilibrium algorithm, commonly used to implement probit traffic assignment (Sheffi, 1985). This method has been seen to be particularly attractive for such problems, where there are typically a large number of alternatives.

The simulation method above may be implemented for arbitrary multivariate normal error structures using only univariate normal pseudo-random numbers (Daganzo, 1979, p 49). In a traffic assignment context, the procedure becomes still more attractive when the correlations between alternative routes are assumed to be formed wholly due to the network overlap of independently-distributed normal *link* error distributions. In such a case, one can avoid enumerating all possible routes in advance, as they can be generated as needed during the simulation process.

It is therefore natural to consider the use of Monte Carlo techniques for the estimation of the probit route choice Jacobian, i.e. the derivatives of the route choice probabilities with respect to the route costs. A simple method of numerically evaluating these derivatives at a given point is finite difference approximation: each cost component is perturbed by a small amount in turn, and the effect on the choice probabilities computed. An estimate of any derivative is then the ratio of the change in the choice probability to the size of the cost perturbation. This approach is, however, potentially unreliable when the choice probabilities are estimated by simulation. In particular, the error in estimating a choice probability is a decreasing function of the number of times it is selected in the simulation (Daganzo, 1979, p 49-51). This implies that the difference in the *estimated* choice probability of such an alternative, between two independent simulation runs, is rather variable, and it is just such a difference which is used in the finite difference approximation of the derivative. A further undesirable property of the finite difference approach is that the signs of the estimated derivatives are not guaranteed to be logical. Although the impact of these problems might be lessened by the use of variance reduction techniques, the serious possibility remains that the error in the derivative estimates may be rather larger than that in the estimated choice probabilities.

Instead we adopt an approximation method for estimating the off-diagonal terms of the probit route choice Jacobian, $\nabla_c \mathbf{P}^{\mathbf{rs}}$, which Daganzo (1979) attributes to McFadden. An outline derivation of this result is provided in Appendix 2. Since this Jacobian is symmetric, it is only necessary to compute the upper triangular terms. For route k the derivative approximation with respect to the cost on an alternative route j, at a given route cost point $\mathbf{c}$ with a variance-covariance matrix $\mathbf{\Sigma}$, is of the form:

$$\left.\frac{\partial P_k}{\partial C_j}\right|_{\mathbf{C}=\mathbf{c}} = \sqrt{\frac{|\mathbf{\Sigma}(k,j)|}{2\pi\ |\mathbf{\Sigma}|}}\exp\left(\frac{A(k,j)}{2}\right) \mathrm{MNP}_k\left(\mathbf{c}(k,j), \mathbf{\Sigma}(k,j)\right) \qquad (k < j;\ k,j \in \kappa(r,s))$$

where $\kappa(r,s)$ is the subset of routes relating to O-D movement r-s, $MNP_k(\mathbf{c}(k,j), \mathbf{\Sigma}(k,j))$ is the multinomial probit (MNP) choice probability for alternative $k \in \kappa(r,s)$ corresponding to a model with $N_{rs} - 1$ alternatives (where the original model had $N_{rs} = |\kappa(r,s)|$ alternatives), mean cost vector $\mathbf{c}(k,j)$ and covariance matrix $\mathbf{\Sigma}(k,j)$; and where the matrix $\mathbf{\Sigma}(k,j)$, row-vector $\mathbf{c}(k,j)$, and scalar $A(k,j)$ are computed from the following scheme (with $\mathbf{c}$ assumed to be in row-vector notation):

1. Compute the inverse covariance matrix $\mathbf{\Sigma}^{-1}$ and the row vector $\mathbf{c\Sigma}^{-1}$.
2. Form the $(N_{rs} - 1)$-square matrix $\mathbf{D}(k,j)$ from $\mathbf{\Sigma}^{-1}$ by:
   (a) adding row j of $\mathbf{\Sigma}^{-1}$ to row k;
   (b) adding column j of the resultant matrix from (a) to its column k;
   (c) deleting row j and column j of the resultant matrix from (b).
3. Similarly, form the $(N_{rs} - 1)$ dimensional row vector $\mathbf{d}(k,j)$ from $\mathbf{c\Sigma}^{-1}$, by adding the j^{th} element to the k^{th}, and then deleting the j^{th} element.
4. Finally compute: $\mathbf{\Sigma}(k,j) = (\mathbf{D}(k,j))^{-1}$ $\quad$ $\mathbf{c}(k,j) = \mathbf{d}(k,j)\,\mathbf{\Sigma}(k,j)$

$$A(k,j) = \mathbf{d}(k,j)\,\mathbf{\Sigma}(k,j)\,(\mathbf{d}(k,j))^T - \mathbf{c\Sigma}^{-1}\mathbf{c}^T \quad .$$

In view of the symmetry of the Jacobian and the fact that the choice probabilities must sum to 1, then once all the off-diagonal terms have been determined, the diagonal terms may be found from:

$$\frac{\partial P_k}{\partial C_k} = -\sum_{j \in \kappa(r,s),\ j \neq k} \frac{\partial P_k}{\partial C_j} \quad .$$

The advantage of this approximation method in the present context is its feasibility for large networks, where the choice probability, $MNP_k(\mathbf{c}(k,j), \mathbf{\Sigma}(k,j))$, may be computed by Monte Carlo methods. (Note that the procedure is carried out independently for each origin-destination movement.)

In practice, this Jacobian is not computed for all available routes, but rather for the subset of used routes arising at the end of some equilibrium solution algorithm. This is a similar restriction to “active paths” as adopted by Tobin and Friesz (1988), as discussed in section 2, although here the level of the restriction is under the control of the modeller. That is to say, the restriction is simply due to the fact that a finite computation will be carried out in which, during the Monte Carlo process, there will still be some paths that have not been selected by the time the algorithm is terminated. By choosing to perform more iterations in computing the unperturbed SUE solution, more active paths will enter into the sensitivity analysis, and so the restriction is entirely under the control of the modeller.

It is worth noting here that although this approach ultimately requires path information, it is not specific to any equilibrium solution method, for example being applicable to solutions obtained by the method of successive averages, or the more recent step-length optimization methods (Maher and Huges, 1997). Indeed, even in the case of a link-based algorithm, the necessary path information may be obtained by using the converged SUE link costs, through the use of replicated Monte Carlo stochastic network loadings, carried out at the fixed SUE link costs.

Applications of this approach are reported in Clark and Watling (2000), where a close agreement is seen between the approximate linear solution and the “exact” (re-estimated) SUE solution. The method is also shown to be feasible for larger realistic networks (notably, the Sioux Falls network). However, while the network size is not a limitation, the network topology may be. The particular problem that arises, and its solution, is described in the following section.

6. THE LINEAR DEPENDENCE PROBLEM

Let us now consider the application of the methods described in section 5 in the context of the SUE sensitivity results described in section 4. By way of illustration, consider the figure of eight network (see Figure 1), where the O-D demand q = 1, and the link cost functions are:

$t_1 = 1 + x_1 \qquad t_2 = 2 + x_2 \qquad t_3 = 4 + x_3 \qquad t_4 = 8 + x_4$

[FIGURE 1 HERE]

The routes may be numbered (for reference) as follows:

Route 1 = Links 1,3 ; Route 2 = Links 1,4 ; Route 3 = Links 2,4 ; Route 4 = Links 2,3 .

Supposing further that the link errors follow independent normal distributions with variance equal to the free flow travel time, then we may write down both the link-path incidence matrix, and the implied route error variance co-variance matrix:

$$\boldsymbol{\Delta} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 5 & 1 & 0 & 4 \\ 1 & 9 & 8 & 0 \\ 0 & 8 & 10 & 2 \\ 4 & 0 & 2 & 6 \end{bmatrix}$$

Now it is evident from figure 1, the link-path incidence matrix and the variance co-variance matrix that there are linear dependencies between the routes within this network. Specifically, if c_k denotes the travel cost on route k, then $c_1 + c_3 = c_2 + c_4$. The effect is that the determinant $|\boldsymbol{\Sigma}|=0$, and so computing sensitivities by considering path choice probabilities as a function of path costs, by the method of section 5, is not then feasible. (It is noted in passing that Tobin & Friesz, 1988, too faced a linear dependency problem, though subtly different in nature: namely how to select from a convex set of possible DUE route flow solutions corresponding to a given DUE link flow solution.)

It should be emphasised that this linear dependence issue is not a fundamental problem of the general probit model, but is directly related to the way in which (a) route cost errors are assumed to be formed purely from link cost error components; and (b) the choice Jacobian is examined with respect to the implied route costs, rather than the link components. One "solution" to problem (a) would therefore be to require that the route cost errors are directly specified (as they are for the logit model, where such problems do not arise), with the requirement that the covariance matrix be of sufficient rank. This seems a rather unsatisfactory strategy however, since one of the main appeals of the probit model in the context of route choice is the manner in which the covariances may be naturally inferred from a combination of specified link cost error components and the network structure. An alternative approach is presented in Yai *et al* (1997), where the variance co-variance matrix contains an additional route specific error term which could, in principle, remove this non-invertibility problem. A pragmatic approach (which has not been investigated) would be to use Yai *et* al's approach to approximate pure link component models, by adding a very small route-specific component, though it might be expected that in practice an ill-conditioned numerical problem may still arise. An alternative which addresses problem (b) would be to instead work in terms of the Jacobian of route choice probabilities with respect to *link* costs. This is a rather natural way to address the problem if the error terms are indeed specified as link components. The disadvantage is that the appealing computational method, described in section 5, does not appear naturally to extend to estimate such a Jacobian.

The approach proposed, therefore, is to make an approximation that allows the method described in section 5 still to be applied: the approximation effectively removes the linear dependencies. Having first computed the unperturbed SUE solution with related path flows, the basic method involves selecting, for each origin-destination movement, the subset of linearly independent paths that in combination explains the greatest proportion of the origin-destination flow. In the figure-of-eight network, this might seem a rather crude approximation: as a pessimistic case, if in SUE the demand were approximately evenly split between the four routes, then the three linearly independent paths chosen would only explain around 75% of the origin-destination demand. There are, however, strong reasons to believe that this poor approximation is a particular facet of small networks, which would not be exhibited in large realistic networks. In such large networks, where there are a great many alternative routes, the linear dependencies are highly unlikely to arise in such a small subset of routes. In that case, then, the maximal linearly independent set is likely to involve a much greater number of routes and cover a much greater proportion of the origin-destination flow. This assertion has been investigated numerically in a number of networks, and particular results to justify it are given below.

The general procedure proposed is therefore as follows. Firstly, an SUE solution is computed for the unperturbed state, and the resulting estimated path flows are saved (if a path-based algorithm was used) or generated at equilibrium (if a link-based algorithm used). Then:

1. If, for any origin-destination movement, the paths contain linear dependencies, then select a subset of linearly independent paths that maximises the total flow on the selected routes (i.e. that explains the most origin-destination flow).

2. From step 1, a set of selected paths and a set of unselected paths will arise. For the unselected paths, fix the flow at the unperturbed SUE value, and load this as a fixed flow onto the links. (Such flows will be fixed during the sensitivity analysis). The selected paths are then assumed to be the active paths, to which the sensitivity analysis is then applied.

It cannot be guaranteed that there is a unique linearly independent set of paths for any given set of SUE paths, that is optimal in the sense of step 1. This is not, however, considered to be an important issue, since (as we shall illustrate in section 7) our numerical experience suggests that the magnitude of the unselected flows involved is sufficiently small to mean that this is unlikely to be significant.

In order to implement step 1, three solution methods were tested: enumeration, a greedy method and linear programming. The *enumeration* (exhaustive search) method works by firstly establishing, for the origin-destination movement under consideration, the rank, R, of the relevant variance co-variance matrix, and then generating all the possible subsets of R routes from the N available. Special-purpose algorithms are used to generate the required subsets in an efficient sequence (Nijenhuis & Wilf, 1975).

The second approach, namely the *greedy* method, works by starting with an empty set of selected routes, and then considers each route in turn, in descending order of flow. At each step, a route is added to the selected set if its inclusion would not create linear dependencies between the selected routes, otherwise it is discarded. This method is clearly not guaranteed to find an optimal solution, but its advantage lies in its computational simplicity. In cases where the first R routes considered turn out to be linearly independent, then the selected set will be optimal.

The third approach uses zero-one integer *linear programming* to maximise the total demand flow in the chosen sub-set. An indicator vector **s** is defined with a value of 1 at position k if route k is to be included and 0 otherwise. By way of example, consider the figure of eight network where SUE route proportions in the unperturbed state are estimated as **p**={0.509, 0.236, 0.081, 0.175}. The linear program formulation is:

$$\begin{array}{ll} \text{Maximise} & s_1\, p_1 + s_2\, p_2 + s_3\, p_3 + s_4\, p_4 \ \text{ with respect to } \mathbf{s} \\ \text{subject to} & s_1 + s_2 + s_3 + s_4 < 4\ . \end{array}$$

In this case the solution is **s**={1,1,0,1} which explains 91.9% of the demand flow. The method of singular value decomposition (Press *et al*, 1992) is used to provide information on whether linear combinations exist, and to determine what those linear combinations are. The general method used for solving the linear program was a branch and bound approach (Park, 1996).

Both the enumeration and linear programming techniques are guaranteed to yield an optimal solution, but the computational effort required is potentially great, especially for O-D pairs with a large number of routes. This is especially true for the enumeration method, which in order to select (say) 20 routes from 30 would require the enumeration of 30,045,015 subsets. If each enumeration took only 0.001 seconds to generate then it would take over 8 hours to generate all these subsets. It is unlikely that enumeration will ever be a practical option; for the purposes of the research study, it was at least useful in verifying the linear programming solution. The greedy method will, it is proposed, produce a good solution in a fraction of the time required for the other two methods, but it will not necessarily be optimal. In empirical experiments conducted with real-world networks, when the optimal solution has been known (through using the enumeration or linear programming method) the greedy method has always produced the same optimal solution.

To explore the practical implications of this exercise in sub-setting the total routes, the impact of an increase in the O-D flow computed using the approximation technique (greedy method) is compared with the re-estimated solutions (i.e. obtained from multiple runs of an SUE solution algorithm) for the figure of eight network. The re-estimated SUE solutions were obtained using the method of successive averages (Sheffi, 1985) with 1 inner *assignment* iteration per outer iteration, and 10 million outer *equilibrium* iterations[2]. The solution vector, the reduced route-link incidence matrix (for the linearly independent paths), the variance-covariance matrix and route choice Jacobian are given below:

---

[2] In this algorithm, the inner iterations are the number of Monte Carlo simulations used to estimate the probit choice fractions at fixed flows/mean costs, and the outer iterations handle the equilibrium feedback (the dependence of mean costs on flows). Sheffi (1985) illustrated that it is typically not efficient to perform more than one inner iteration per outer iteration.

$$\mathbf{x} = \begin{bmatrix} 0.6833 \\ 0.3168 \\ 0.7451 \\ 0.2549 \end{bmatrix} \quad \boldsymbol{\Delta}^{rs} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{\Sigma}_{rs} = \begin{bmatrix} 5 & 4 & 1 \\ 4 & 6 & 0 \\ 1 & 0 & 9 \end{bmatrix} \quad \nabla_c \mathbf{P}^{rs} = \begin{bmatrix} -0.2272 & 0.1909 & 0.0363 \\ 0.1909 & -0.2042 & 0.0133 \\ 0.0363 & 0.0133 & -0.0497 \end{bmatrix}.$$

The Jacobian of the link travel times, $\nabla_x \mathbf{t}$, is the identity matrix. The calculated values for the $\mathbf{M}$ and $\mathbf{N}$ and the product $\mathbf{M}^{-1}\mathbf{N}$ matrix are:

$$\mathbf{M} = \begin{bmatrix} 1.8552 & 0.1448 & 0.0086 & -0.0086 \\ 0.1448 & 1.8552 & -0.0086 & 0.0086 \\ 0.0086 & -0.0086 & 1.9550 & 0.0450 \\ -0.0086 & 0.0086 & 0.0450 & 1.9550 \end{bmatrix} \quad \mathbf{N} = \begin{bmatrix} 0.3416 \\ 0.1584 \\ 0.3725 \\ 0.1275 \end{bmatrix} \quad \mathbf{M}^{-1}\mathbf{N} = \begin{bmatrix} 0.6344 \\ 0.3656 \\ 0.7262 \\ 0.2738 \end{bmatrix}.$$

Figure 2 illustrates the correspondence between the solutions calculated from the linear approximations and the re-estimated exact solutions. The correspondence is good, even on those links which are part of the “neglected” route, namely links 2 and 4.

[FIGURE 2 HERE]

7. APPLICATION TO A LARGER NETWORK

The full sensitivity method has been applied to a number of hypothetical grid networks and the Sioux Falls network (see Clark & Watling, 2000), but for illustration we consider only a network representing the north-east of Leeds, cordoned from a network maintained by Leeds City Council, and containing some 123 links and 29 zones. Cost is assumed to be equal to time, with link travel time functions of the BPR form (Bureau of Public Roads, 1964). Perceived link travel times are assumed to be independent and normally distributed, with standard deviation equal to 0.3 of the free-flow travel time.

The first test involved a 10% (20 unit) increase in a single O-D flow. The calculation of the base probit SUE from 1,000 iterations by the method of successive averages took 1 minute on a 450MHz Pentium II PC. The calculation of the **M** matrix from the 230 non-zero O-D flows took a further 14 minutes, while the calculation time for the **N** matrix is negligible. The greedy method was used to select linearly independent subsets. For O-D pairs involving modest selection tasks (less than 13,360 possible combinations) the linear programming method was used to verify the optimality of the greedy solution.

Comparing the re-estimated equilbrium following the demand change with the linear sensitivity approximation, the average percentage absolute difference in link flows was 0.19%, with the largest error 2.41%; the distribution of these differences is illustrated in Figure 3. Figure 3 demonstrates that the error from using the approximation is small for the vast majority of links. Additionally, the average percentage difference in link travel times was only 0.01% with a maximum difference of 0.34%. On a network level, the modest 0.32% rise in total demand gave an increase in network travel time of 1.22% (approximation) against 1.12% (re-estimated).

[FIGURE 3 HERE]

Figure 4 illustrates the impact of the linear subsetting method, by showing the proportion of O-D flow explained by the selected routes, across the relevant O-D movements. The lowest proportion explained is 83%, whereas for 168 O-D pairs 100% is explained.

[FIGURE 4 HERE]

A second example involved changing the capacity on a link on a main arterial from 1,680 vehicles per hour to 2,280. The total network travel time for the unperturbed solution was 19,655 vehicle-minutes, for the approximate solution 19,405 minutes (1.27% less than the base), and for the re-estimated solution 19,395 minutes (1.32% less). The average absolute percentage difference between the approximate and re-estimated solutions for link flows was 1.08% and for link travel times was 0.10%, the maximum difference for link flows being 8.31% and for link travel times 5.14%.

In this example, the time noted above to calculate the sensitivity expressions may appear large, compared to the time to calculate a single SUE. However, this allows the impact to be estimated of any change to the 230 non-zero O-D flows, or any of the 4 link parameters on the 123 links, a total of $230 + 123 \times 4 = 722$ possible input data items to change. The sensitivity information is therefore obtained at a “rate” of one item every 1.16 seconds. In this respect, we would claim that the method is highly efficient, since to obtain a similar amount of information would otherwise require a considerable number of equilibrium algorithms to be solved.

8. CONCLUSION

This paper has outlined a technique for deriving a linear sensitivity result for the SUE model. The method is not limited to any particular form of random utility error structure, and in particular has been shown to be practical for a probit formulation. The technique described forms part of a study which is developing methods for estimating confidence intervals for the outputs from traffic assignment models. A companion part of the study is investigating the characteristics of the sampling and systematic variations of the input parameters to these traffic assignment models, in order to deduce sensible sampling distributions for the model inputs (O-D matrix, capacities, etc.). The sensitivity relationships may then be used to map the input sampling distributions to distributions for the output measures, such as link flows and travel times. There are clearly many other areas in which the SUE sensitivity analysis described here may be further developed. Two potential areas are the issue of link flow interactions, which in principle may be addressed by the fixed point sensitivity expressions derived in Appendix 1, and the use of improved (second order) sensitivity expressions.

ACKNOWLEDGEMENTS

The work presented in this paper was financed by the UK Engineering and Physical Sciences Research Council, under grant reference GR/L79540. The authors would like to thank the two anonymous referees whose observations helped to improve an earlier draft of this paper.

REFERENCES

Bell M.G.H. and Iida, Y. (1997). *Transportation Network Analysis.* John Wiley, Chichester.

Bureau of Public Roads (1964). *Traffic Assignment Manual*. US Department of Commerce.

Clark, S.D. and Watling, D. P. (2000). Probit Based Sensitivity Analysis for General Traffic Networks. *Transportation Research Record No 1733, p88-95.*

Dafermos, S.C. (1980). Traffic Equilibrium and Variational Inequalities. *Transportation Science* **14**, 42-54.

Dafermos, S.C. and Nagurney, A. (1984). Sensitivity analysis for the asymmetric network equilibrium problem. *Mathematical Programming* **28**, 174-184.

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

Davis, G.A. (1994). Exact local solution of the continuous network design problem via stochastic user equilibrium assignment. *Transportation Research* **28B**(1), 66-75.

Fiacco, A.V. (1983). *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming.* Academic Press, New York.

Hall, M.A. (1978). Properties of the Equilibrium State in Transportation Networks. *Transportation Science* **12**(3), 208-216.

Heydecker, B.G. (1983). Some consequences of detailed junction modelling in road traffic assignment. *Transportation Science* **17**, 263-281.

Leurent, F. (1996). An analysis of modelling error, with application to a traffic assignment model with continuously distributed values of time. *Proc 24th PTRC European Transport Forum*, September 2nd-6th 1996, Seminar D/E.

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

Magnanti, T.L. and Wong R.T. (1984). Network Design and Transportation Planning: Models and Algorithms. *Transportation Science* **18**, 1-55.

Maher, M.J. and Hughes, P.C. (1997). A probit-based stochastic user equilibrium assignment model. *Transportation Research* **31B**(4), 341-355.

Nijenhuis, A. and Wilf, H.S. (1975). *Combinatorial Algorithms.* Academic Press, New York.

Park, S (1996). A Collection of Computer Code for Operational Research Problems. Available at http://orly1.snu.ac.kr/software/index_software.html, visited on 12/12/2000.

Press, W.H., Vetterling, W.T., Teukolsky, S.A. and Flannery, B.P (1992). *Numerical Recipes in C*. Cambridge University Press. Cambridge.

Ran B. and Boyce D.E. (1993). *Dynamic Urban Transportation Network Models*. Springer Verlag, Berlin.

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

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

Smith, M.J. (1979). The Existence, Uniqueness and Stability of Traffic Equilibria. *Transportation Research* **13B**, 295-304.

Tobin, R.L. and Friesz, T.L. (1988). Sensitivity Analysis for Equilibrium Network Flow. *Transportation Science* **22**(4), 242-250.

Verlander, N.Q. and Heydecker B.G. (1994). Improved matrix estimation methods under congested conditions. Unpublished paper presented at *26th Annual Universities Transport Study Group Conference*, University of Leeds, UK, January 5th-7th 1994.

Watling D.P. (1996). Asymmetric problems and stochastic process models of traffic assignment. *Transportation Research* **30B**(5), 339-357.

Yai, T., Iwakura, S., and Morichi, S. (1997). Multinomial Probit with Structured Covariance for Route Choice Behavior. *Transportation Research* **21B**(3), 195-207.

Yang, H. (1995). Sensitivity Analysis for Queuing Equilibrium Network Flow and its Application to Traffic Control. *Mathematical and Computer Modelling* **22**(7), 247-258.

Yang, H. (1997). Sensitivity Analysis for the Elastic-Demand Network Equilibrium Problem with Applications. *Transportation Research* **31B**(1), 55-70.

Yang, H. & Bell, M.G.H. (1997). Traffic restraint, road pricing and network equilibrium. *Transportation Research* **31B**(4), 303-314.

Yang, H., Sasaki, T., Iida, Y. and Asakura, Y. (1992). Estimation of origin-destination matrices from link traffic counts on congested networks. *Transpn Research* **26B**(6), 417-434.

Yang, H., Yagar, S., Iida, Y. and Asakura, Y. (1994). An algorithm for the in-flow control problem on urban freeway networks with user-optimal flows. *Transportation Research* **28B**(2), 123-139.

APPENDIX 1: SENSITIVITY ANALYSIS OF SUE BY FIXED POINT FORMULATION

It is convenient to adapt the notation slightly:

- $\mathbf{x}$ column vector of link flows;
- $\tilde{\mathbf{P}}(\mathbf{t})$ matrix of link choice proportions as a function of the vector $\mathbf{t}$ of link costs;
- $\tilde{\mathrm{P}}_{an}(\mathbf{t})$ typical element of $\tilde{\mathbf{P}}(\mathbf{t})$, the proportion of flow for origin-destination movement n that uses link a when the link costs are $\mathbf{t}$;
- $\tilde{\mathbf{q}}$ column vector of origin-destination demand levels.

We consider the impact of perturbations to the link cost functions (demand perturbations can be handled in a similar way). Let $\mathbf{t}(\mathbf{x},\mathbf{s})$ denote the vector of link cost functions, dependent on some vector $\mathbf{s}$ of link attributes (e.g. capacities).

Consider the function:

$$\mathbf{f}(\mathbf{x},\mathbf{s}) \equiv \mathbf{x} - \tilde{\mathbf{P}}(\mathbf{t}(\mathbf{x},\mathbf{s}))\ \tilde{\mathbf{q}} \quad .$$

Let $\mathbf{x}^*(\mathbf{s})$ denote the SUE solution at a given value $\mathbf{s}$ of the parameter vector, so that:

$$\mathbf{f}(\mathbf{x}^*(\mathbf{s}),\mathbf{s}) = \mathbf{0} \qquad \text{for any given } \mathbf{s}.$$

Assuming differentiability of the involved functions, and regarding $\mathbf{s}$ now as a variable, a first order Taylor series expansion of $\mathbf{f}(\mathbf{x},\mathbf{s})$ in the neighbourhood of $(\mathbf{x},\mathbf{s}) = (\mathbf{x}^*(\mathbf{s}_0),\mathbf{s}_0)$ is:

$$\mathbf{f}(\mathbf{x},\mathbf{s}) \approx \mathbf{f}(\mathbf{x}^*(\mathbf{s}_0),\mathbf{s}_0) + \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{(\mathbf{x}^*(\mathbf{s}_0),\mathbf{s}_0)} (\mathbf{x} - \mathbf{x}^*(\mathbf{s}_0)) + \left.\frac{\partial \mathbf{f}}{\partial \mathbf{s}}\right|_{(\mathbf{x}^*(\mathbf{s}_0),\mathbf{s}_0)} (\mathbf{s} - \mathbf{s}_0)$$

where the derivative terms are the Jacobian matrices of $\mathbf{f}$ with respect to $\mathbf{x}$ and $\mathbf{s}$ respectively, evaluated at $(\mathbf{x}^*(\mathbf{s}_0),\mathbf{s}_0)$, which we shall henceforth denote $\mathbf{J}_1$ and $\mathbf{J}_2$.

The logic is that since $\mathbf{f}(\mathbf{x}^*(\mathbf{s}_0),\mathbf{s}_0)) = \mathbf{0}$ (as an SUE at $\mathbf{s}_0$), and knowing $\mathbf{s}_0$, $\mathbf{x}^*(\mathbf{s}_0)$, $\mathbf{J}_1$ and $\mathbf{J}_2$, then for some other given $\mathbf{s} \neq \mathbf{s}_0$ we can approximately solve the equilibrium condition $\mathbf{f}(\mathbf{x}(\mathbf{s}),\mathbf{s}) = \mathbf{0}$ for $\mathbf{x}(\mathbf{s})$ by substitution in the linear approximation above:

$$\mathbf{0} \approx \mathbf{0} + \mathbf{J}_1 (\mathbf{x}(\mathbf{s}) - \mathbf{x}^*(\mathbf{s}_0)) + \mathbf{J}_2 (\mathbf{s} - \mathbf{s}_0) \quad \Rightarrow \quad \mathbf{x}(\mathbf{s}) \approx \mathbf{x}^*(\mathbf{s}_0) - \mathbf{J}_1^{-1}\mathbf{J}_2 (\mathbf{s} - \mathbf{s}_0) .$$

Expressions for $\mathbf{J}_1$ and $\mathbf{J}_2$ in terms of the link cost Jacobian and choice Jacobian are readily derived (see Bell and Iida, 1997, pp 198-199).

Example

Consider a network consisting of two parallel links/routes serving a single origin-destination movement with demand 1 unit. Let the link cost functions be $t_1(x_1) = 1 + x_1^2$ and $t_2(x_2) = 2 + x_2$. Suppose that the link choice probabilities are derived from the logit form $\tilde{P}_1(\mathbf{t}) = (1+\exp(t_1 - t_2))^{-1}$, $\tilde{P}_2(\mathbf{t}) \equiv 1 - \tilde{P}_1(\mathbf{t})$. The unique SUE solution is $(x_1,x_2)=(0.6948,0.3052)$.

Now letting s denote the “capacity” parameter for link 1, and writing $t_1(x_1) = 1 + sx_1^2$ with $s_0 = 1$, then:

$$\mathbf{J}_1^{-1} = \begin{bmatrix} 0.8044 & 0.1407 \\ 0.1956 & 0.8593 \end{bmatrix} \qquad \mathbf{J}_2 = \begin{bmatrix} 0.1024 \\ -0.1024 \end{bmatrix}$$

and therefore:

$$\begin{bmatrix} x_1(s) \\ x_2(s) \end{bmatrix} \approx \begin{bmatrix} 0.6948 \\ 0.3052 \end{bmatrix} + \begin{bmatrix} 0.0679 \\ -0.0679 \end{bmatrix} \begin{bmatrix} s-1 \\ 0 \end{bmatrix}$$

Alternatively, using the equivalent optimization method described in section 4, we obtain the following expression for a change $\varepsilon = s - 1$:

$$\mathbf{M}^{-1} = \begin{bmatrix} 0.5789 & 0.1407 \\ 0.1407 & 0.8593 \end{bmatrix} \qquad \mathbf{N} = \begin{bmatrix} 0.1422 \\ -0.1024 \end{bmatrix}$$

and therefore:

$$\begin{bmatrix} x_1(\varepsilon) \\ x_2(\varepsilon) \end{bmatrix} = \begin{bmatrix} 0.6948 \\ 0.3052 \end{bmatrix} + \begin{bmatrix} 0.0679 \\ -0.0679 \end{bmatrix} \begin{bmatrix} \varepsilon \\ 0 \end{bmatrix}$$

namely, the same as the fixed point approach.

The fact that we have chosen a logit choice model to illustrate this point is purely for illustrative ease; precisely the same is true for the probit model.

APPENDIX 2: DERIVATION OF APPROXIMATION FOR PROBIT CHOICE JACOBIAN

Following the proof of Daganzo (1979, pp 72-73), a probit choice problem with three alternatives is considered here (the proof is easily generalised to an arbitrary number of alternatives).

If U_i denotes the utility perceived for alternative i, then the probit choice probability of choosing alternative i = 3, say, is:

$$p_3(\mathbf{V},\mathbf{\Sigma}) = \Pr\left(U_3 > \max(U_1,U_2) \mid E[\mathbf{U}] = \mathbf{V},\ \mathrm{var}(\mathbf{U}) = \mathbf{\Sigma}\right)$$
$$= \int_{u_3=-\infty}^{\infty} \int_{u_2=-\infty}^{u_3} \int_{u_1=-\infty}^{u_3} \phi\left(u_1,u_2,u_3 \mid \mathbf{V},\mathbf{\Sigma}\right) du_1 du_2 du_3$$

where $\phi(u_1,u_2,u_3 \mid \mathbf{V},\mathbf{\Sigma})$ denotes the density function of a multivariate normal variable with mean vector **V** and covariance matrix **Σ**. Note that we have to take a little care with the order of integration, since the region of integration does not have constant boundaries. Then, differentiating with respect to V_1, for example, yields:

$$\frac{\partial p_3}{\partial V_1} = \int_{u_3=-\infty}^{\infty} \int_{u_2=-\infty}^{u_3} \int_{u_1=-\infty}^{u_3} \frac{\partial \phi\left(u_1,u_2,u_3 \mid \mathbf{V},\mathbf{\Sigma}\right)}{\partial V_1} du_1 du_2 du_3 .$$

Now since by definition of the multivariate normal density, V_1 and u_1 appear in the same way in ϕ but with opposite sign, $\frac{\partial \phi}{\partial V_1} = -\frac{\partial \phi}{\partial u_1}$. But then by the Fundamental Theorem of Calculus

$$\int_{u_1=-\infty}^{u_3} \frac{\partial \phi\left(u_1,u_2,u_3 \mid \mathbf{V},\mathbf{\Sigma}\right)}{\partial u_1} du_1 = \phi\left(u_3,u_2,u_3 \mid \mathbf{V},\mathbf{\Sigma}\right) . \quad \ldots(*)$$

The right hand side of expression (*) should be carefully noted, with the (dummy) variable u_3 appearing twice as an argument.

Now, the general form of an n-dimensional multivariate normal density function is:

$$\phi\left(\mathbf{u} \mid \mathbf{V}, \boldsymbol{\Sigma}\right) = \left((2\pi)^n |\boldsymbol{\Sigma}|\right)^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}(\mathbf{u}-\mathbf{V})\boldsymbol{\Sigma}^{-1}(\mathbf{u}-\mathbf{V})^{\mathrm{T}}\right)$$

and the key term in the exponent may be expanded as

$$\begin{aligned}(\mathbf{u}-\mathbf{V})\boldsymbol{\Sigma}^{-1}(\mathbf{u}-\mathbf{V})^{\mathrm{T}} &= (\mathbf{u}-\mathbf{V})\boldsymbol{\Sigma}^{-1}(\mathbf{u}^{\mathrm{T}}-\mathbf{V}^{\mathrm{T}}) \\ &= \mathbf{u}\boldsymbol{\Sigma}^{-1}\mathbf{u}^{\mathrm{T}} - \mathbf{u}\boldsymbol{\Sigma}^{-1}\mathbf{V}^{\mathrm{T}} - \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{u}^{\mathrm{T}} + \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{V}^{\mathrm{T}} \\ &= \mathbf{u}\boldsymbol{\Sigma}^{-1}\mathbf{u}^{\mathrm{T}} - 2\mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{u}^{\mathrm{T}} + \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{V}^{\mathrm{T}}\end{aligned}$$

where the last equality exploits the symmetry of $\boldsymbol{\Sigma}^{-1}$. Based on (*) above, we now write $u_1 = u_3$, and aim for an equivalent representation of the matrix expansion above, in terms of $\hat{\mathbf{u}} = \left(u_2 \quad u_3\right)$ only. By expanding the relevant matrix multiplications, it can be shown that just such an alternative representation is

$$\mathbf{u}\boldsymbol{\Sigma}^{-1}\mathbf{u}^{\mathrm{T}} - 2\mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{u}^{\mathrm{T}} + \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{V}^{\mathrm{T}} = \hat{\mathbf{u}}\mathbf{Q}\hat{\mathbf{u}}^{\mathrm{T}} - 2\hat{\mathbf{u}}\mathbf{R} + \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{V}^{\mathrm{T}} \qquad (**)$$

where $\mathbf{Q}$ is obtained from $\boldsymbol{\Sigma}^{-1}$ by adding row 1 to row 3, adding column 1 to column 3 of the resultant matrix, and then deleting row 1 and column 1; and where $\mathbf{R}$ is obtained from $\mathbf{V}\boldsymbol{\Sigma}^{-1}$ by deleting the first element and adding it to the third. Now, expression (**) may alternatively be written in the form:

$$\begin{aligned}\hat{\mathbf{u}}\mathbf{Q}\hat{\mathbf{u}}^{\mathrm{T}} - 2\hat{\mathbf{u}}\mathbf{R} + \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{V}^{\mathrm{T}} &= (\hat{\mathbf{u}} - \mathbf{R}\mathbf{Q}^{-1})\mathbf{Q}(\hat{\mathbf{u}} - \mathbf{R}\mathbf{Q}^{-1})^{\mathrm{T}} - \mathbf{R}\mathbf{Q}^{-1}\mathbf{R}^{\mathrm{T}} + \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{V}^{\mathrm{T}} \\ &= (\hat{\mathbf{u}} - \hat{\mathbf{V}})\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\mathbf{u}} - \hat{\mathbf{V}})^{\mathrm{T}} - \mathbf{A}\end{aligned}$$

where

$$\hat{\mathbf{V}} = \mathbf{R}\mathbf{Q}^{-1} \qquad \hat{\boldsymbol{\Sigma}}^{-1} = \mathbf{Q} \qquad \mathbf{A} = \mathbf{R}\mathbf{Q}^{-1}\mathbf{R}^{\mathrm{T}} - \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{V}^{\mathrm{T}} .$$

Tying this rearrangement together with result (*), we have:

$$\begin{aligned}\phi\left(u_3, u_2, u_3 \mid \mathbf{V}, \boldsymbol{\Sigma}\right) &= \left((2\pi)^3 |\boldsymbol{\Sigma}|\right)^{-\frac{1}{2}} \exp(\tfrac{1}{2}\mathbf{A})\ \exp\left(-\tfrac{1}{2}(\hat{\mathbf{u}} - \hat{\mathbf{V}})\hat{\boldsymbol{\Sigma}}^{-1}(\hat{\mathbf{u}} - \hat{\mathbf{V}})^{\mathrm{T}}\right) \\ &= \frac{\left((2\pi)^3 |\boldsymbol{\Sigma}|\right)^{-\frac{1}{2}}}{\left((2\pi)^2 |\hat{\boldsymbol{\Sigma}}|\right)^{-\frac{1}{2}}} \exp(\tfrac{1}{2}\mathbf{A})\ \phi\left(\hat{\mathbf{u}} \middle| \hat{\mathbf{V}}, \hat{\boldsymbol{\Sigma}}\right) = \left(\frac{|\hat{\boldsymbol{\Sigma}}|}{2\pi|\boldsymbol{\Sigma}|}\right)^{\frac{1}{2}} \exp(\tfrac{1}{2}\mathbf{A})\ \phi\left(\hat{\mathbf{u}} \middle| \hat{\mathbf{V}}, \hat{\boldsymbol{\Sigma}}\right)\end{aligned}$$

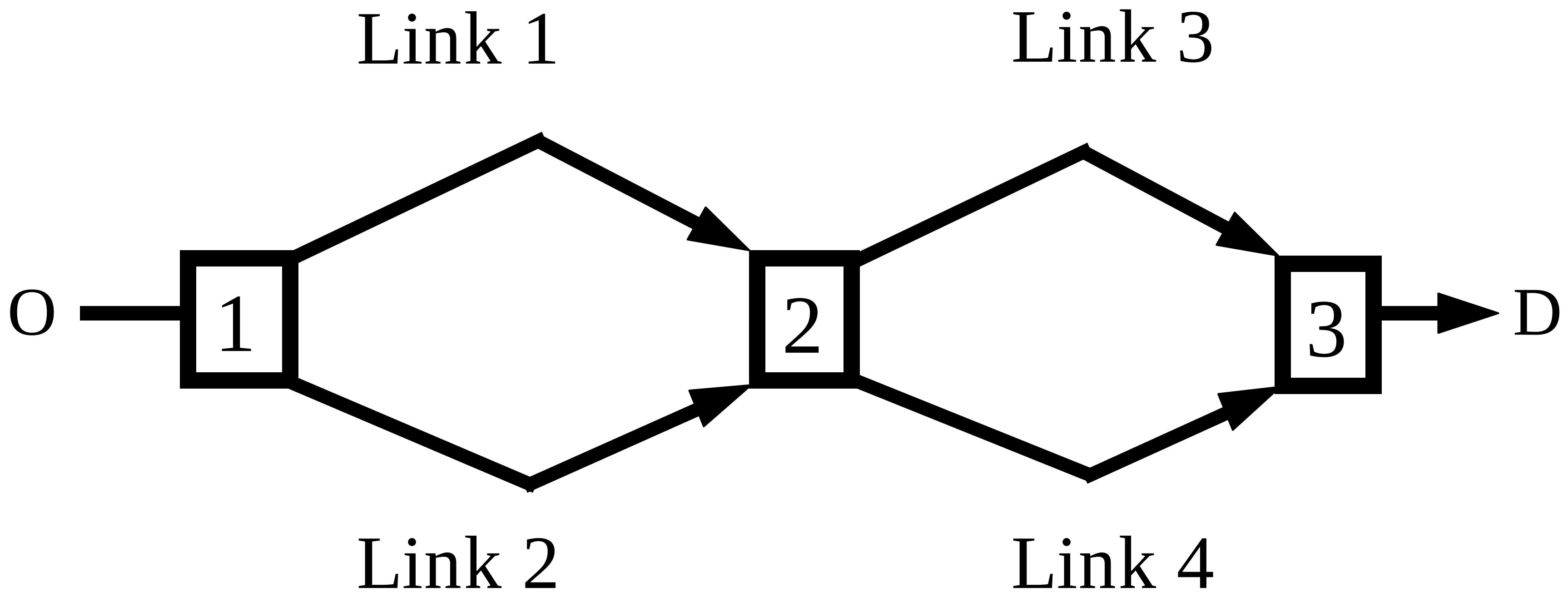
where we have introduced an additional leading term in order to write the expression in the form of a multivariate normal density for $\hat{\mathbf{u}}$. Hence,

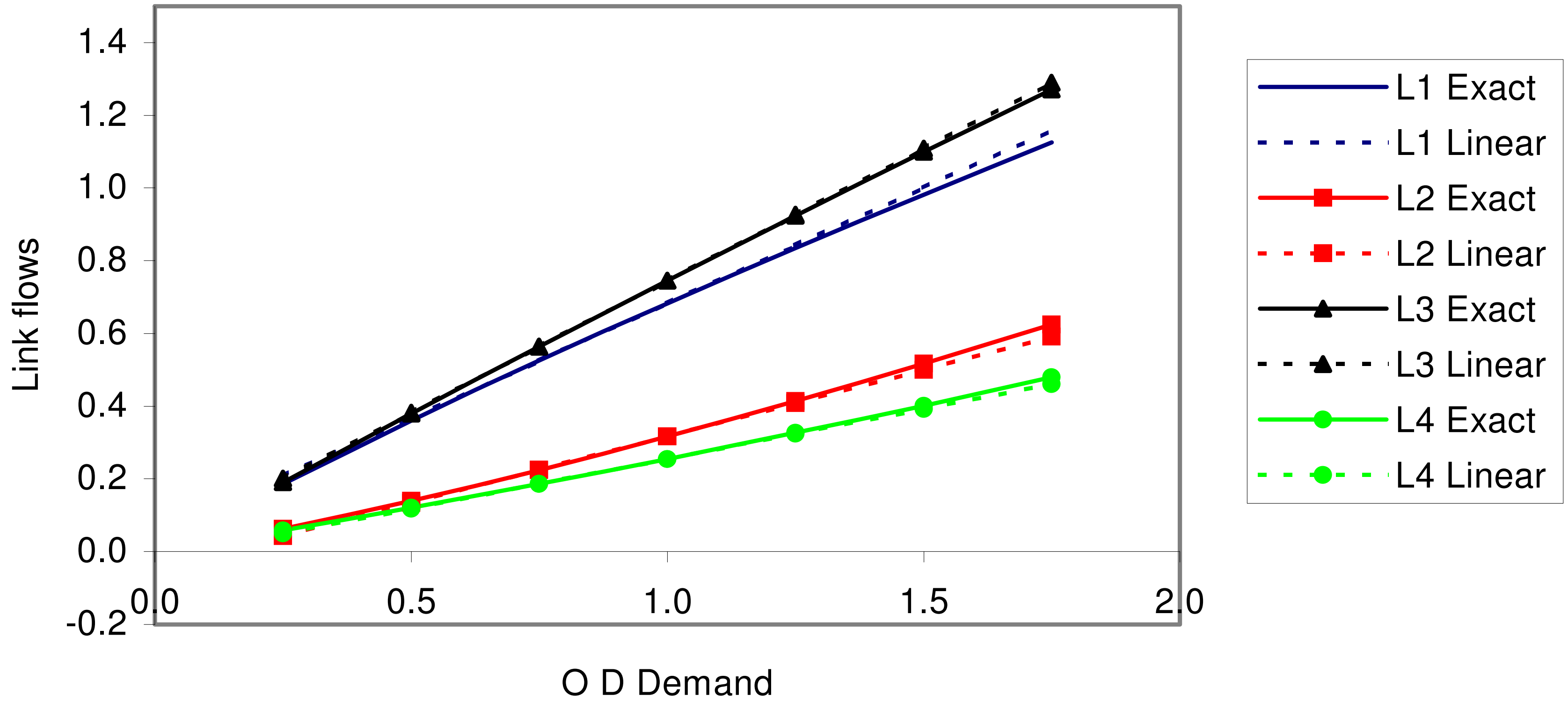
$$\frac{\partial p_3}{\partial V_1} = -\left(\frac{|\hat{\boldsymbol{\Sigma}}|}{2\pi|\boldsymbol{\Sigma}|} \right)^{\frac{1}{2}} \exp(\tfrac{1}{2}\mathbf{A}) \int_{u_3=-\infty}^{\infty} \int_{u_2=-\infty}^{u_3} \phi\left(u_2, u_3 \middle| \hat{\mathbf{V}}, \hat{\boldsymbol{\Sigma}}\right) du_2 du_3$$

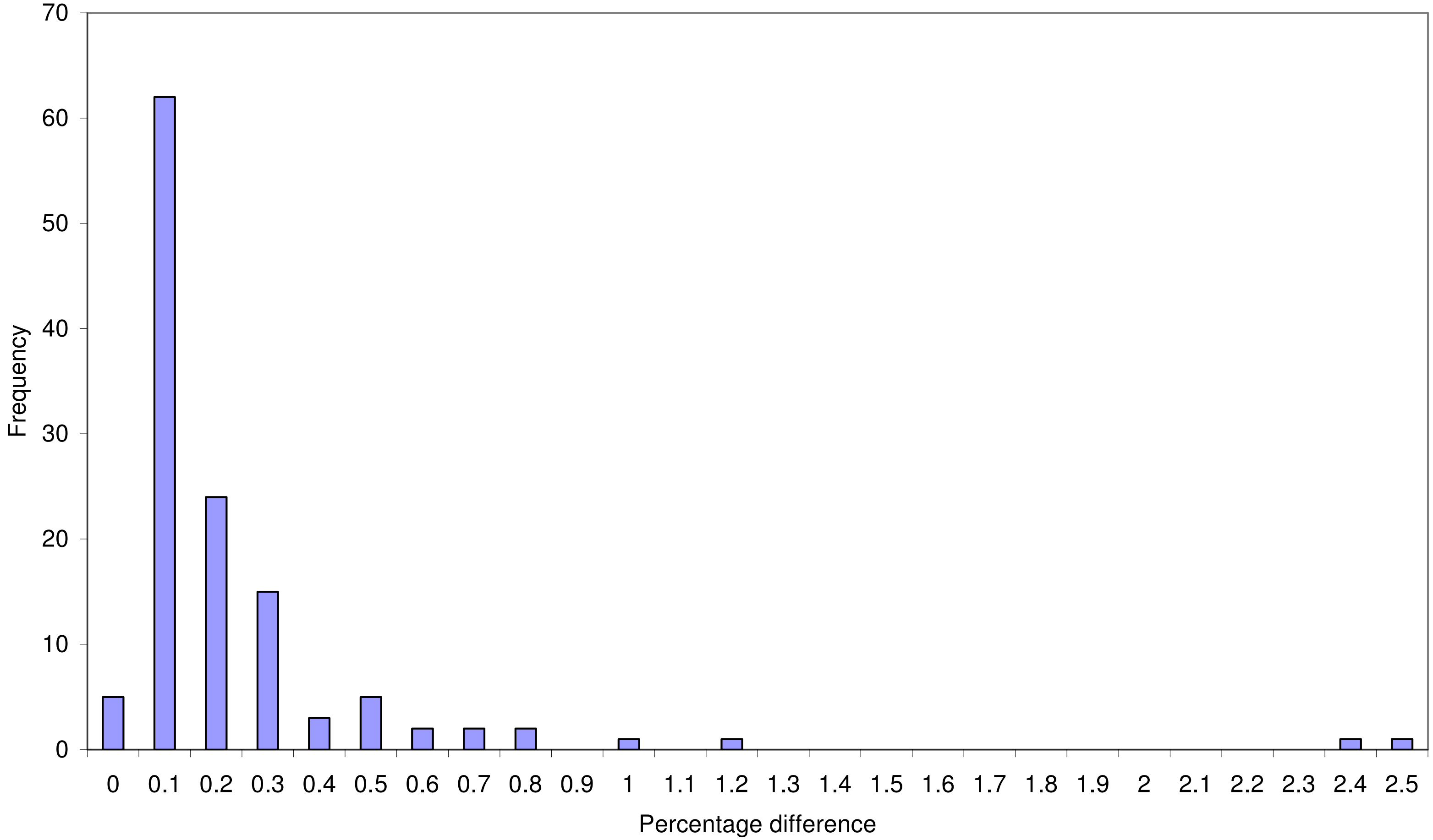
$$= -\left(\frac{|\hat{\boldsymbol{\Sigma}}|}{2\pi|\boldsymbol{\Sigma}|} \right)^{\frac{1}{2}} \exp(\tfrac{1}{2}\mathbf{A}) \; p_3(\hat{\mathbf{V}}, \hat{\boldsymbol{\Sigma}}) \; .$$

That is to say, in order to determine the required derivative, we need to determine the choice probability $p_3(\hat{\mathbf{V}}, \hat{\boldsymbol{\Sigma}})$ for a problem with alternative 1 deleted (the 1 arising from the fact that we are differentiating with respect to alternative 1's utility). This approach is valid for determining any off-diagonal entry of the choice Jacobian.

The expression above is exact, no approximation is involved. In practice, however, the choice probability for the modified problem must itself be estimated, say by Monte Carlo simulation. Nevertheless, there is an advantage relative to finite difference approximation that the error in estimating the derivative is of the same order as the error in estimating the choice probability. If, as in the current application, one also needs to estimate choice probabilities for the full problem via Monte Carlo techniques, then the choice probabilities for the modified problems are readily computed during this process, meaning that the derivatives may be rather efficiently computed.


Link 1
Link 3
O
1
2
3
D
Link 2
Link 4


1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
Link flows
0.0
0.5
1.0
1.5
2.0
O D Demand
L1 Exact
L1 Linear
L2 Exact
L2 Linear
L3 Exact
L3 Linear
L4 Exact
L4 Linear


Frequency
70
60
50
40
30
20
10
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2 2.1 2.2 2.3 2.4 2.5
Percentage difference

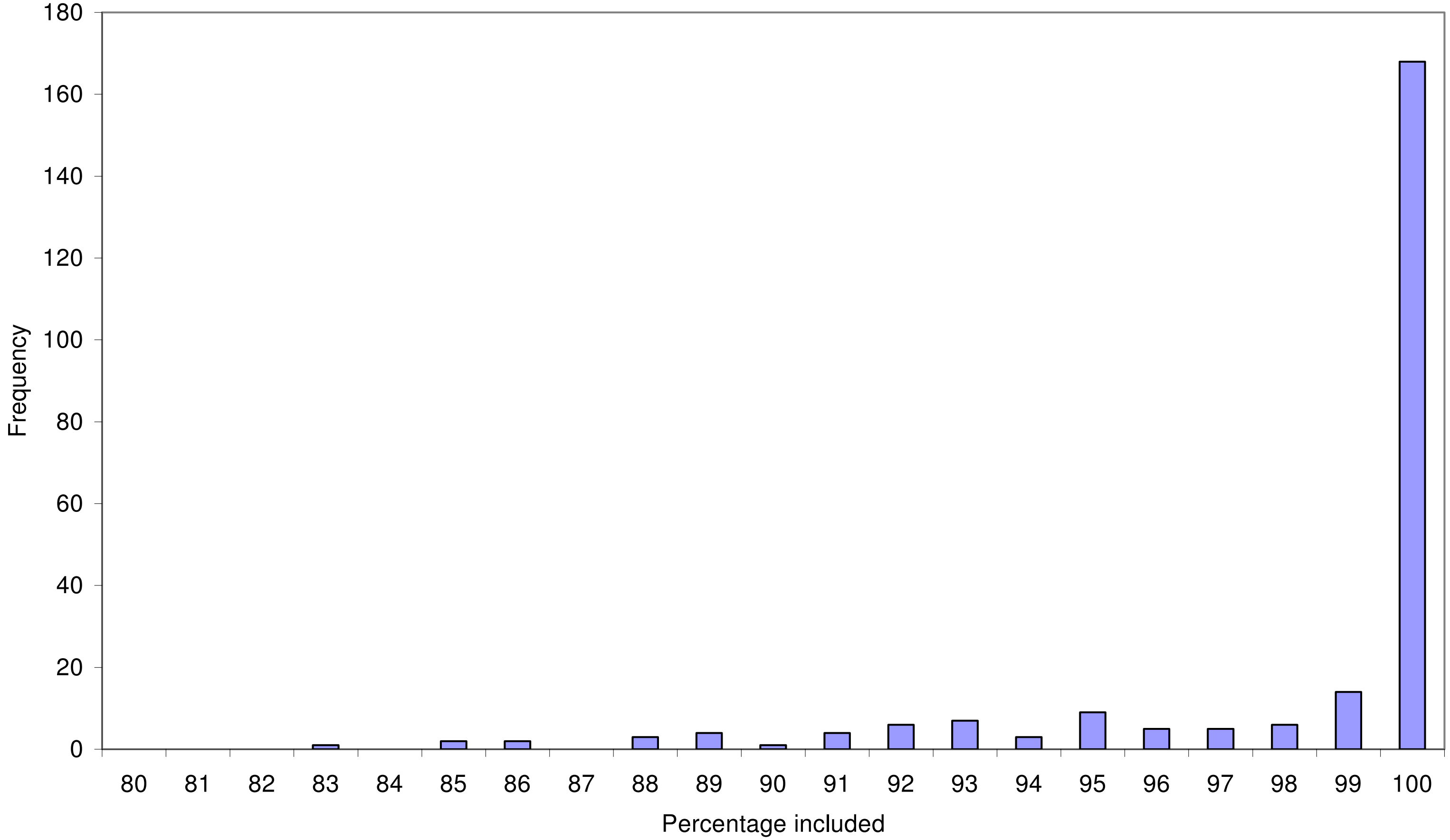

180
160
140
120
100
80
60
40
20
0
Frequency
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
Percentage included

Figure Captions:

Figure 1 : Figure-of-eight network

Figure 2 : Comparison between exact and re-estimated solutions (figure-of-eight network)

Figure 3 : Distribution of percentage differences in link flows for approximate and re-estimated solutions (Headingley, demand perturbation)

Figure 4 : Distribution of the percentage of O-D flow included in the senstivity analysis, across OD movements (Headingley)