Article:

Connors, RD orcid.org/0000-0002-1696-0175 and Watling, DP orcid.org/0000-0002-6193-9121 (2015) Assessing the demand vulnerability of equilibrium traffic networks via network aggregation. Networks and Spatial Economics, 15 (2). pp. 367-395. ISSN 1566-113X

https://doi.org/10.1007/s11067-014-9251-9

Assessing the Demand Vulnerability of Equilibrium Traffic Networks via Network Aggregation

Authors: Richard D. Connors, David P. Watling
Affiliation: Institute for Transport Studies, University of Leeds, Leeds, LS2 9JT, United Kingdom

Corresponding Author:
Email: R.D.Connors@its.leeds.ac.uk
Tel: 0113 3431799
Fax: 0113 343 5334

Abstract
Studies of network vulnerability typically focus on changes to the supply side; whether considering a degradation of link capacity or complete link failure. However, the level of service provided by a transport network is also vulnerable to increases in travel demand, with the consequent congestion causing additional delays. Traffic equilibrium models can be used to evaluate the influence of travel demand on levels of service when interest is restricted to only a small number of pre-specified demand scenarios. A demand-vulnerability analysis requires understanding the impact of unknown future changes to any possible combination of OD demands. For anything but the smallest networks, this cannot be accomplished by re-computing network equilibrium at all possible demand settings. We require a representation of the functional relationship between demands and levels of service, avoiding the need to re-evaluate the equilibrium model. This process—of collapsing the demand and network representations onto a single, coarse-level network with explicit functional relationships—is referred to here as 'network aggregation'. We present an efficient method for network aggregation for networks operating under Stochastic User Equilibrium (SUE). In numerical experiments, we explore the nature and extent of the aggregation errors that may arise.

Keywords
Demand Vulnerability. Network Aggregation. Sensitivity Analysis. User Equilibrium. Stochastic User Equilibrium.

1 Introduction and Review

Real-life transport systems are forever in a state of flux, evolving over different time-scales. Accidents, breakdowns and severe weather may lead to short-term, unplanned reductions in capacity, and special events may lead to unusual levels of demand, all of which lead to unpredictable traffic conditions on a particular day for both travellers and road operators. In the medium-term, routine maintenance of roads/utilities and seasonal changes in demand patterns may all lead to periodic changes to traffic patterns. In the longer-term, capacity improvements and growth/decay in the overall demand for travel will lead to a changing systematic trend in traffic patterns.

Looking to the past, traffic authorities worldwide have taken a growing interest in measuring, historically, the extent to which their transportation systems have been able to accommodate such changes while still delivering an acceptable level-of-service. This has led to several schemes around the world for monitoring the repeatability of journey times, historically (see the review of Watling & Balijepalli, 2012). It has also led to transport planners seeking to design transport systems which are relatively robust to such changes that may occur in the future, by somehow taking account of such eventualities well before their occurrence and scale is actually known. The focus of the present paper is on such a future-oriented approach to robust planning.

While other possibilities exist, two broad classes of approach have dominated the transportation literature for assisting planners in addressing such kinds of problem. Reliability methods are typified by an approach in which: (i) probability distributions are first specified of the causes (e.g. capacity, demand, weather) or impacts (e.g. travel time variability), (ii) a network/behavioural model is assumed for how travellers might respond in such a changing environment, considering the interaction of their choices with congestion patterns, and (iii) the consequential distribution of measures of system performance (or moments or quantiles thereof) are inferred. In vulnerability methods, on the other hand, we avoid the need to posit probability distributions, and instead aim to identify parts of the system that are weak or could most contribute to degradation in system performance. A disadvantage of vulnerability methods is that they do not, therefore, provide a prediction of the likely future states of the network, as reliability methods do. On the other hand, the advantage of vulnerability methods is that they effectively remove the need to associate probabilities to the system states identified in element (i) of the reliability approach. Removing the need to specify the probabilities of future events becomes more attractive, the further in the future the robust planning horizon considers, since the uncertainty in specifying such probabilities is much greater in such cases.

In the field of reliability analysis, a wide range of tools now exist which vary in terms of the types of variability modelled, the kinds of network/behavioural assumptions and models adopted, the algorithms used for implementing the methods, and the measures of system reliability impact considered. The field was originally motivated by considering variations in link capacity, and this has continued to be a strong theme running through the research methods. The origins can be traced to methods in which the system measure is the probability of an OD pair to remain connected by any available path, assuming independent link failures, as described by Bell & Iida (1997), ranging to later work which allows multiple link states, correlated across links, and alternative behavioural models for how travellers might adjust (e.g. Sumalee & Watling, 2008). In such methods, the concept of connectivity is generalised such that the measure is the probability of the travel time for an OD movement to exceed some threshold (connectivity obtained in the limit, as the threshold tends to infinity). Taking a somewhat different stance on the system measure of reliability, Yang et al (2000) and Chen et al (2002) instead proposed calculating the probability of catering for a given level of demand in the face of stochastic network capacities. In contrast, other papers have sought to highlight the role played in system performance by variations in demand about some mean level, rather than variations in capacity (e.g. Clark & Watling, 2005; Shao et al, 2006; Nakayama & Watling, 2014), and indeed several methods now exist that allow both demand and capacity to be randomly varying (e.g. Lam et al, 2008; Sumalee et al, 2011; Uchida, 2014), including the impacts of weather.

In the field of vulnerability analysis, a corresponding wide array of techniques now exist, again primarily focused on the impacts of degradations in capacity (Bell, 2000; D'Este & Taylor, 2001; Nicholson & Dalziell, 2003; Taylor et al, 2006; Jenelius et al, 2006; Szeto et al, 2006; Yang & Qian, 2012; Sullivan et al, 2013). As with reliability methods, these papers cover a diverse range of assumptions, techniques and measures of impact, though with the general theme that they aim to find the weakest part of the system by considering how it might perform, when allowing for travellers to readjust to the degraded situation. Similarly to reliability methods, there has also been an emerging interest in considering the role of demand in the vulnerability of congested networks. Berdica (2002) considered interruptions to critical links, general capacity reductions and variability of the demand, using equilibrium models to numerically assess the sensitivity of networks to such potential interruptions. Chen et al (2012) considered the impact of variability in OD demand particularly for large-scale systems. Ho et al (2013), on the other hand, used a continuum model to explore vulnerability at a regional level. Watling and Balijepalli (2012) considered the twin problem of demand variability and demand growth, developing a link-level indicator to parallel those considered for capacity changes (Knoop et al, 2012).

In the present paper, we shall focus on the role of OD demand on network vulnerability. Motivated partly by the observations in Watling and Balijepalli (2012), we shall specifically consider how demand growth might affect the performance of a network. Our interest is not so much in variability (either in capacity or OD demand), which might occur in the short to medium term, but rather in longer term trends in demand patterns. Typically, such patterns may be expected to unfold over many years, although in rapidly-emerging economies we are can observe major changes to demand patterns over shorter periods. In essence, our take on vulnerability is to consider to what extent the existing network capacity can accommodate future growth, without travel times/costs growing at an unacceptable rate.

The main methodological contribution of our paper is developed from the observation that we may link such an analysis of demand vulnerability to the methods developed in two existing fields of transportation analysis, namely network aggregation and sensitivity analysis. In order to understand such connections, it is useful to first imagine numerically solving a great number of network equilibrium problems for a given network, with the problems differing by making adjustments in some defined to the levels of demand in the OD matrix. For simplicity's sake, let us imagine that the Deterministic User Equilibrium (DUE) is used with separable link travel cost functions, then for each OD demand matrix instance we obtain a unique travel cost (on any used path) at the OD level. If there are n OD demand movements, then by varying the OD demand we obtain for each OD movement a functional relationship between that movements OD travel cost and the n OD demand levels; let us call these the OD cost functions, which are clearly non-separable. We could then draw a network in which all OD movements are directly connected by a single link, and associate the corresponding OD cost function with each such link; we refer to this as the 'aggregate network'. Indeed, as we shall show, this kind of conceptual development is not specific to the DUE model, and in fact we can develop a kind of 'aggregate network' also based on the Stochastic User Equilibrium (SUE) model, which it transpires has some mathematical advantages in our subsequent analysis.

While this conceptual link to network aggregation is useful in theory, it seems to have two major drawbacks. Firstly, it seems to require that we solve an enormous number of equilibrium problems, which would not seem to be practicable in real-life systems. Secondly, even if we can overcome these numerical problems, we would then present planners with a functional relationship from and to a Euclidean space that is as large in dimension as the number of OD movements, and it is unclear how they might use this information. Therefore, having established the theoretical link to network aggregation (section 2), and the notation to be subsequently adopted (section 3), we numerically explore the nature of these relationships in simple cases (section 4), in order to motivate our approach to both efficiently estimate (section 5) and to illustrate/explore the aggregate relationships in the context of vulnerability (section 6). Our solution approach in section 5 utilises the connection to sensitivity analysis, whereby the gradient vector of the relationship at a single equilibrium solution is seen to be (perhaps surprisingly) sufficient to construct a reasonable approximation to the whole aggregate relationship. In section 6, we further utilise sensitive analysis to develop a simple illustration of point vulnerability through a demand vulnerability matrix, with the idea that such a matrix might be examined at different point of future growth. Finally, in section 7, we draw conclusions from the research and examine future research directions.

2 Aggregation in Transport Network Models

Transport analysis must span many *scales*. For example, at one extreme its objective may be the design of traffic signals at an urban intersection for bus priority, or the shock-wave analysis on traffic flow of a motorway incident. At a wider geographical level, transport planning over a city network may be required to examine the impact on route choice and travel demand of measures such as road user charging or new retail developments. At a still wider level, the analysis may need to address problems of regional or national impact, such as the introduction of a national road user charging system. An element in all such analyses is that as geographical scope of the problem changes, so does the scale of the models used, in terms of the fidelity of network representation, the size of the zones over which trip demands are modelled, and the level of disaggregation of traveller responses by, say, socio-economic group. The question then naturally arises: are these models all *consistent* in some sense, across the different aggregation scales?

Focusing specifically on the application of network equilibrium route choice models, as are commonly used in transport planning exercise, this issue of differing aggregation scales may manifest itself in a number of ways, but a common feature is that we would rather not have to specify in advance the travel demand matrix:

For example, our interest may be in modelling a defined urban area, but we are aware of the potential for through-trips that may originate and terminate outside the urban area. Such trips may be sensitive to policies over a wider ('regional') scale, which may cause them to re-route so as not to use the urban area at all; or alternatively, congestion at the regional scale may lead to re-routing that would be seen as new trips through the urban area. Transport planners may aim to marry these issues together by using the outputs from a regional-scale model as inputs of demand to the urban-scale model, even though the different models may typically represent the network and congestion phenomena in different ways, thus leading to inconsistencies between the two models in terms of the estimated travel time to traverse the urban area.

A different kind of problem emerges even when we consider only one geographical area in isolation, but wish to explore the vulnerability of the area to unknown potential changes in the trip demands. In this case the difficulty is that while in principle we may re-run our model in many scenarios in which the demands are changed, there are so many such scenarios it is difficult to know how to select them without it becoming an excessive computational exercise.

In such cases, the common issue is that essentially we wish to understand how changes to the (urban) trip matrix will manifest themselves in changes to travel times to make those trips. This information is embedded implicitly in the equilibrium relationship, but we seemingly cannot extract this embedded information without performing excessive computational experiments. We refer to this as an 'aggregation problem' as our view is that its resolution is understanding the embedded functional relationships between the travel demand matrix and the trip travel times, but at a scale that does not require us to resort to re-running the fine (network link) level model for each demand instance.

Such questions regarding aggregation are not new. In the landmark review of network equilibrium approaches given by (Friesz, 1985), he notes an '*interest in network aggregation [that] stems primarily from the high computational cost of solving large network equilibrium and design models*'. Purely from a viewpoint of computational speed, it is interesting to note that such a motivation is still potentially valid today, even given on-going increases in computer power: this is particularly the case for network design problems, where one may wish to find network equilibrium flows a great many times as part of an algorithm to evaluate, for example, the efficacy of alternative road pricing cordon designs (Sumalee, 2004). Yet even in cases where computational speed is not an obstacle – as witnessed by the increasing number of large-scale regional and national network models in operation containing tens of thousands of links – there still exist basic questions of consistency of scale. For example, even if we could do so, would it be desirable to run a national model each time we wished to adjust the traffic signals at one junction? Clearly this would ensure that the urban, regional and national scales are consistent, but would we believe impacts at this level, and how would we deal with validation of the model? In any case, we cannot neglect that fact that in practice there are a wide variety of transport modelling software packages in use, which have different capabilities directed at different kinds of problem. Many packages that are able to deal with land-use/transport interactions or demand modelling in a sophisticated way are not able to deal with network modelling: and interfacing such software programs is both time-consuming and generates problems of its own in ensuring convergence to self-consistent solutions.

Therefore there is a justifiable *need* to examine and understand the problem of aggregation scale. In transport we use the term *aggregation* to refer to the level of detail included in network and behavioural models, it is also used to refer to methods used in order to summarise characteristics of detailed models for larger scale analyses. Looking to the transport literature, it is possible to identify several aggregation themes that have attracted attention.

Firstly, there is *decision aggregation* (dating back to Theil (1957)). Aggregation here is with respect to the distribution of (socio-economic) attributes across the population of decision makers: with a continuous distribution of different individuals, determining the number of people choosing each alternative involves an integral. This introduces a computational burden into the equilibrium assignment of transport network analysis; aggregation in this context has been investigated for both logit (Sheffi, 1984) and probit (Bouthelier & Daganzo, 1979) models. However, these approaches to the aggregation of decision makers do not shed much light on the issue of network aggregation.

Secondly, there is *traffic aggregation*, which covers a large spectrum of approaches (indeed it could be argued that such a class also incorporates the network aggregation methods described below, though we have chosen to separate these classes). At one extreme, this class contains an analysis of traffic flow dynamics on a single link, leading to theory on the aggregation of individual car-following models to give fluid flow PDEs (Lighthill & Whitham, 1955). Recent work in this category: following Herman and Prigogine (1979) and Ardekani and Herman (1987), Daganzo (2007) employs the two-fluid model to aggregate dynamic urban flow into zones, resulting in a multiple-reservoir model. This representation approaches aspects of the area speed-flow and continuum models mentioned below.

Thirdly, *zonal aggregation* considers the method by which continuous space is partitioned into discrete zones; in transport modelling this is typically for the representation of trip demand patterns. Simplification of trip origins and destinations into representative (centroids of) traffic agglomeration zones (TAZs) is a particular instance of the modifiable areal unit problem; the subject of many papers outside of transport network analysis journals. The effect of zonal aggregation on parameter estimates and goodness-of-fit in spatial interaction models (without a network) has received attention for decades (Openshaw, 1977; Batty & Sikdar, 1982a, 1982b, 1982c, 1982d; Viegas, Martinez & Silva, 2009). In a pair of papers Daganzo (1980a, 1980b) proposes a theoretical method to model a TAZ with many centroids (rather than just one as is usually the case), and further shows how to compute the UE flows under this representation, dealing with the very large number of centroids this method introduces into the network assignment problem.

Fourthly and finally, there exists the problem of *network aggregation*, where a discrete network of congestible links is somehow simplified or summarised (as this class of problem is the focus of the present paper, we will review it in more detail here). Practitioners face such a problem commonly, where they will often need to move between different levels of network fidelity for policies at different geographical scales and/or with different kinds of short-to-long term response represented. However, there is little systematic guidance on how the resulting inconsistencies might be minimised; in this respect, in terms of the state of practice, the comments of Friesz (1985) would still seem to ring true today: 'aggregation has been practised in an essentially *ad hoc* fashion since the advent of widespread use of network transportation planning models'. However, we know from empirical studies that both the method and level of aggregation may significantly affect model predictions and policy conclusions (Bovy & Jansen, 1983; Sbayti, El-Fadel & Kaysi, 2002; Chang, Khatib & Ou, 2002). Thus, in order systematically to analyse the topic of network aggregation, we need to consider both the type of representation used and the process by which such a representation is estimated. In this way, existing methods of network aggregation can be split into four main classes[1]:

Aggregation of uncongested networks. In cases where a network representation with uncongestible links is appropriate, several systematic methods exist, for example in the geography and planning literature (Xu & Sui, 2007); Xie and Levinson (2007) review such measures and analyse their relevance to transport. Rogers et al. (1991) provide a corresponding review of network aggregation studies existing in the operational research literature, again for constant link cost problems. As one example of such approaches, Zipkin (1980) shows how to simplify a linear minimum-cost network-flow optimisation problem by replacing groups of nodes with aggregate nodes, and derives bounds for the solution to the aggregated problem.

Construction of Area-wide Congestion Relationships. In approaches which develop area speed-flow relationships, the intention is to represent a steady-state relationship between (a) the total trip demand in an area within some time period of the day (with trips possibly weighted in some way, e.g. by trip length), and (b) the average speed of vehicles as they move within this area (e.g. Herman & Prigogine, 1979; May, Shepherd & Bates, 2000; Hills, 2001; May, Shepherd & Bates, 2001). More recently, there has been a focus on traffic flow dynamics within areas, with macroscopic fundamental diagrams aiming to represent the relationship between (a) the traffic density (veh/km/lane) in an area, and (b) the space-mean flow of vehicles (veh-km/hr) in that area (e.g. Daganzo & Geroliminis, 2008; Tsekeris & Geroliminis, 2013).

Link Extraction Methods. In such approaches, the aim is to discard 'unimportant' links and nodes in the network (strictly speaking, then, their aim could be said to be simplification as opposed to aggregation).

[1] Our classification covers only those approaches which begin with a discrete graph representation and aim for some kind of simplified representation from that graph. Hence, although they have some relation, we exclude from this classification continuum models, whereby a dense urban network is represented as a continuum, with flows comprising a vector field (e.g. Dafermos, 1980; Wong, 1998; Daniele, Idone & Maugeri, 2003).

Chan (1976) proposes such a (heuristic) link extraction method, though it is restricted to networks with constant cost links. The only analytical link extraction method that has been proposed for congested transport networks is that of Hearn (1984). In this approach, a 'focus area' is first identified within a large detailed network, assumed to be operating under user equilibrium. A transfer decomposition method then splits the User Equilibrium (UE) objective function into a pair of interrelated mathematical programs, one relating to the links of the focus area and the other to its complement, and a set of artificial links introduced to connect the two decomposed parts of the original network.

Link Abstraction Methods. In such approaches the aim is to replace sets of network nodes/links with representative abstract nodes/links. Simple network substructures, such as two links in series or parallel, can easily be replaced by a single representative link; implemented network-wide, this gives an effective method that can be used to simplify large networks (Wright, Xiang, Waller, et al., 2010). Boyles (2012) presents an analytical method for the examination of a sub-network of interest by aggregating the remainder of the network. Like Hearn's link extraction method described above, Boyles's method is based on the analysis of the UE model using a 'focus area' within a larger network. Analytical sensitivity analysis is then used to generate a representative trip matrix for the network outside of the focus area sub-network.

The approach presented in the current paper falls within this final class of link abstraction methods. Like Boyles (2012) our focus will be on implicitly capturing the effect of route choice within the aggregated area/relation; that is to say, our interest is in the fact that as demand increases, self-organising principles of route choice imply that travellers attempt to mitigate potential congestion increases within an aggregated area through re-routing at the fine level (where possible), but that this is a complex, high-dimensional and highly non-linear effect to capture, especially without an explicit fine-level network to represent it. In order to do so, we accept the compromise of modelling link congestion through steady-state relationships between travel time and flow, and so do not aim to capture dynamic phenomena (the kind that would be evident from an approach based on macroscopic fundamental diagrams, for example). Our analysis departs from that of Boyles in our use of the Stochastic User Equilibrium (SUE) model for route choice. This means that we derive smooth relationships throughout, unlike the UE case. Furthermore, and importantly, the resulting aggregate relationships are analytically derived from, and are unique with respect to the underlying detailed network model. Uniqueness is an important property as it avoids the arbitrariness noted above of network aggregation/simplification methods adopted in practice. The fact that we link our aggregate relationships uniquely to the underlying fine-scale network also distinguishes our work in approach from the more empirically-motivated, area-wide congestion relationships described above. For example, even though several of the papers cited above estimate area speed-flow relationships from detailed (micro-simulation) models, this is done by correlating two aggregate measures of the micro-simulation under user-defined scenarios. There is no sense in which these are the *unique* area speed-flow relationships associated with the fine-level model, the estimated relationships will depend on the (arbitrary) method by which the scenarios are generated.

3. Definitions & Notation

We represent the road network by a directed graph consisting of N links labelled $i = 1,2, \ldots, N$; a demand vector $\boldsymbol{q}$, with entries $q^r \geq 0$ $[r = 1, \ldots, R]$, representing the travel demand on the r^{th} origin-destination (OD) movement, which is served by a non-empty set of acyclic paths K^r. An assignment of flows to all paths is denoted by $\boldsymbol{f}$, with elements $f_k^r \geq 0$ giving the flow on the k^{th} path connecting the r^{th} OD. The assignment is feasible for demand vector $\boldsymbol{q}$ if and only if

$$\sum_{k\in K^r} f_k^r = q^r\ \forall r\ \text{ and } f_k^r \geq 0\ \forall k, r \tag{1}$$

The (closed, convex) set of feasible path flows thus defined is denoted F. The link travel costs are single-valued functions, $t_i(\boldsymbol{x})$, of the vector of link flows, $\boldsymbol{x} = \Delta \cdot \boldsymbol{f}$, where the link-path incidence matrix, Δ, has elements, $\delta_{ik}^r \in \{0,1\}$, denoting the links i that are part of path k, ordered by OD movement r. The mapping between flows and costs arises from the standard link-additive model:

$$\boldsymbol{c}(\boldsymbol{x}) = \Delta^T \boldsymbol{t}(\Delta \boldsymbol{f}\,) \tag{2}$$

The foundation text for deterministic network equilibrium is Wardrop (1952), who proposed that travellers will change route in order to reduce their travel cost. The consequence is that, at equilibrium, all used routes have equal cost for each OD movement, with no unused route being cheaper. This condition is referred to as User Equilibrium (UE). The path flow vector $\boldsymbol{f}^*$ is a solution to UE (Smith, 1979) if it satisfies (6) and

$$\boldsymbol{c}(\Delta \cdot \boldsymbol{f}^*)^T(\boldsymbol{f} - \boldsymbol{f}^*) \geq 0\ \forall \boldsymbol{f} \in F \tag{3}$$

If we begin with a less constrained representation of individual choice behaviour, we can derive a generalization of the UE paradigm. Define the *perceived utility* of path k (OD movement r) to be the sum of the systematic utility $v_k^r = -\theta c_k^r(\boldsymbol{x})$ and a continuous random variable ϵ_k^r:

$$U_k^r(\boldsymbol{c};\theta) = v_k^r(\boldsymbol{c};\theta) + \epsilon_k^r = -\theta c_k^r(\boldsymbol{x}) + \epsilon_k^r \tag{4}$$

Where $\theta > 0$ is a scaling parameter, $v_k^r = \mathrm{E}[U_k^r]$, $\mathrm{E}[\epsilon_k^r] = 0$, $\mathrm{Var}[U_k^r] = \mathrm{Var}[\epsilon_k^r] = \sigma_k^r$, and the residuals have joint PDF (probability density function) $\phi(\boldsymbol{\epsilon}) \geq 0$.

The probability that a traveller chooses path k (from the path set K^r) is given by the probability that the perceived utility of this path is greater than the alternatives in K^r

$$P_k^r(\boldsymbol{c};\theta) = \Pr\big(U_k^r(\boldsymbol{c};\theta) > U_j^r(\boldsymbol{c};\theta)\quad \forall j \in K^r\big) \tag{5}$$

The *satisfaction* refers to the expected maximum utility over all the alternatives. On the r-th OD this is

$$S^r = E\left[\max_{k\in K^r}\{U_k^r(\boldsymbol{c};\theta)\}\right] \tag{6}$$

It will be more convenient to use the satisfaction expressed in cost units. We define the *composite cost* to be $W^r = (-1/\theta)S^r$, so that W^r expresses a degree of *deterrence* to travel movement, which it will be convenient to compare with the generalised cost used in deterministic UE models. Assuming that the random residuals have non-zero, finite variances and that the PDF, $\phi(.)$, is continuously differentiable, strictly positive and independent of the systematic utility (Cantarella, 1997) then

$$\frac{\partial W^r}{\partial c_k^r} = P_k^r \tag{7}$$

A *stochastic user equilibrium* (SUE) is a feasible path flow vector, $\boldsymbol{f}^*$, such that

$$\boldsymbol{f}^* = \boldsymbol{q} \cdot \boldsymbol{P}\big(\boldsymbol{c}(\Delta \cdot \boldsymbol{f}^*)\big) \tag{8}$$

In addition to the assumptions made above, if the link cost functions are strictly monotone increasing then the SUE link flows and path flows are unique (see Cantarella, 1997). With these assumptions in hand, we henceforth use $\boldsymbol{x}^* = \Delta \boldsymbol{f}^*$ to denote the SUE link flow vector.

Note that, at SUE, all paths are used even though they do not have equal cost. Hence there is no unique "OD travel cost" for used routes, as is the case with UE. The analogous quantity is the (unique) cost-based OD composite cost, W^r, as defined above. Since not all travellers use the (actual) minimum cost path at SUE, the average experienced SUE travel cost for the OD pair is always in excess of the minimum deterministic OD travel cost. These inequalities give (for any $\theta > 0$ and $\boldsymbol{c} \geq 0$):

$$W^r = -\frac{1}{\theta} E\left[\max_{k\in K^r}\{U_k^r(\boldsymbol{c};\theta)\}\right] \leq \min_k\{c_k^r\} \leq \sum_{k\in K^r} P_k^r c_k^r \tag{9}$$

This is a property of the composite cost that it will be important to recall, in order to fully understand our later numerical results, especially when comparing UE with SUE. Note that in the limit as the ϵ-variances reduce to zero, all terms in (9) become equal: the perceived OD travel cost become the minimum (actual/measured) OD travel cost, and we recover the common OD travel cost of UE (see Sheffi, 1985).

4. Motivating examples: Exploration of implicit aggregate relationships under the UE and SUE models

In this section we present a series of examples that are intended to explain the basis for our later analysis. While in the remainder of the paper there are particular differences that emerge between using the SUE or UE models, this choice makes no material difference to the main issues we wish to illustrate in the present section. Since under the UE model we can derive explicit relationships for some small networks (that are not available for SUE), we begin by focussing on results obtained under the UE model; in later sections we will adopt an SUE approach, but will use some of the numerical results obtained in the present section for comparison between SUE and UE.

Regardless of whether a UE or SUE paradigm is adopted, the standard network representation used in transport analysis can be seen as comprising two networks as shown in Figure 1a. The (upper) network of desired travel movements (arrows) between origin and destination nodes must travel through the (lower) congestible supply network of road links. The network aggregation process that we adopt can be viewed as collapsing these two networks into a single equilibrated network of multidimensional OD cost functions as illustrated in Figure 1b. It is these cost functions that will be used to reveal the demand vulnerabilities of the network. We aim to illustrate this through some examples. In these small examples it is sufficiently easy and fast simply to re-solve the UE problem many times for different instances of the OD demand matrix; we focus on develop more practical techniques for larger networks in subsequent sections.

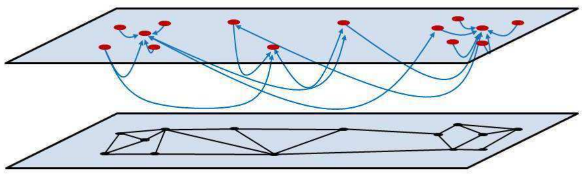

Figure 1a: The demand (upper) and supply (lower) networks

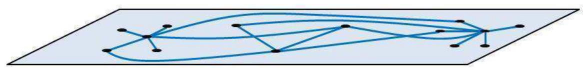

Figure 1b: The single layer OD equilibrium network with multidimensional cost functions

4.1 UE: Analytic results for an N-parallel link network with linear travel cost functions

To illustrate the basis for network aggregation set out in this paper, we begin by considering a network of N parallel links under user equilibrium (UE), serving demand $q > 0$, and with the i-th link having travel cost function

$$t_i(x) = a_i + b_i x_i.$$

Without loss of generality $0 \leq a_1 \leq a_2 \leq \cdots \leq a_N$ and $0 < b_i$. Under any demand $q > 0$, the UE link flows are unique (conditions given in Smith, 1979). If demand is sufficiently low, the user equilibrium solution is that all flow is accommodated by link 1 (the link with least free flow travel cost). This happens if

$$a_1 + b_1 q \leq a_2 \Leftrightarrow q \leq \frac{a_2 - a_1}{b_1} \tag{10}$$

As the demand increases new links become active in order of increasing free flow travel cost. Each additional link carrying flow requires the specification of the OD travel cost function to be extended by an additional demand 'regime'.

Over its domain, $q > 0$, the equilibrium flow on each link, $x_i^*(q)$, is continuous, differentiable except at the intersection of the linear parts, and directionally differentiable everywhere. Link 1 is used at all levels of demand, and so the equilibrium travel cost that any user would experience can be found by substituting $x_1^*(q)$ into the travel cost flow function for link 1. According to the definition of UE all used routes will have the same travel cost, thus under any UE flows we can refer to the UE travel cost $T^*(q)$ for any traveller using the network, at any demand q.

Within each demand regime, the OD travel cost function is linear. Collecting them together gives a unique, piecewise-linear specification of the OD travel cost function under UE that covers the entire domain, $q > 0$. As a composition of continuous functions, $T^*(q)$ is continuous over its domain. It is differentiable everywhere except intersection points, and directionally differentiable everywhere.

Active Links	Demand Regime	OD Travel cost
1	$0 < q$	$T^*(q) = b_1 q + a_1$
2	$\frac{1}{b_1}(a_2 - a_1) \le q$	$T^*(q) = \frac{1}{b_1 + b_2}(b_1 b_2 q + b_2 a_1 + b_1 a_2)$
3	$\frac{1}{b_1 b_2}\,(b_1[a_3 - a_2] + b_2[a_3 - a_1]) \le q$	$T^*(q) = \frac{(b_1 b_2 b_3 q + b_2 b_3 a_1 + b_3 b_1 a_2 + b_1 b_2 a_3)}{b_1 b_2 + b_2 b_3 + b_3 b_1}$
Define $B^N = \prod_{i=1}^{N} b_i$ and $B_j^N = \prod_{\substack{i=1 \\ i \ne j}}^{N} b_i = \frac{\prod_{i=1}^{N} b_i}{b_j} = \frac{B^N}{b_j}$		
N	$\frac{1}{B_N^N}\left(\sum_{j=1}^{N-1} B_j^{N-1}[a_N - a_j]\right) \le q$	$T^*(q) = \frac{1}{\sum_{j=1}^{N} B_j^N}\left(B^N q + \sum_{j=1}^{N} B_j^N a_j\right)$

Table 1: N-parallel Links OD Travel cost Function

The aggregate representation of this N parallel link network under UE is a single OD link with the travel cost function as set out above. This gives a unique, consistent, aggregate representation of our original N-link network under UE, in the sense that both the aggregate and disaggregate networks give rise to exactly the same OD travel cost for any traveller, at any level of demand q. The aggregate network gives exact predictions of OD travel costs at all levels of demand.

In terms of the OD cost, and its susceptibility to change under different demand scenarios, the N-parallel link network is thus equivalent to a network model consisting of a single link carrying flow q and with link travel cost function given in Table 1. This simple single OD network paves the way for more general results to follow. An explicit 2-link example:

$$a_1 = 20, b_1 = \frac{1}{600}, \qquad a_2 = 25, b_2 = \frac{1}{900}$$

In Figure 2, the OD cost coincides with the link 1 cost (blue solid line). Link 2 cost (red dashed line) is initially higher and hence this link unused. Link 2 activates at demand $q = 3000$, since $t_1(3000) = t_2(0)$. The OD cost curve is non-differentiable at the link activation point $q = 3000$.

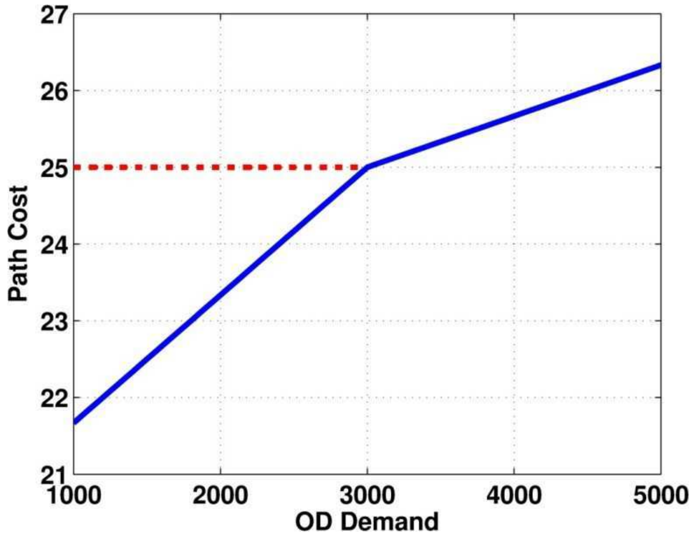

Figure 2: UE 2-link network. OD cost (blue solid line) and link 2 cost (red dashed line)

4.2 UE: More General Networks

For more complex networks it is (typically) not possible to write down the OD travel cost function in closed form[2]. Nevertheless, the OD travel cost under UE is a well-defined single valued function of the demand, even in the case of large networks with multiple OD movements and non-linear link travel cost function (so long as the UE link flows are unique; $\nabla_x t$ being positive definite is sufficient, for example, see Smith (1979)). Moreover, the OD travel cost is continuous and directionally differentiable, being differentiable except (Patriksson, 2004) at the finitely many transition points between demand regimes (where new routes activate).

The OD travel cost functions are straightforward to evaluate at any given demand level, by solving for UE. However, understanding the demand vulnerability of the network requires examining the behaviour of the multivariate OD cost functions across a range of demands. Tackling this by multiple re-evaluations of UE would be prohibitively computationally expensive except for very small networks; this is an issue we address in later sections. For the moment, however, we consider a small example 'toy city network', as shown in Figure 3, where re-computing UE solutions many times is not problematic.

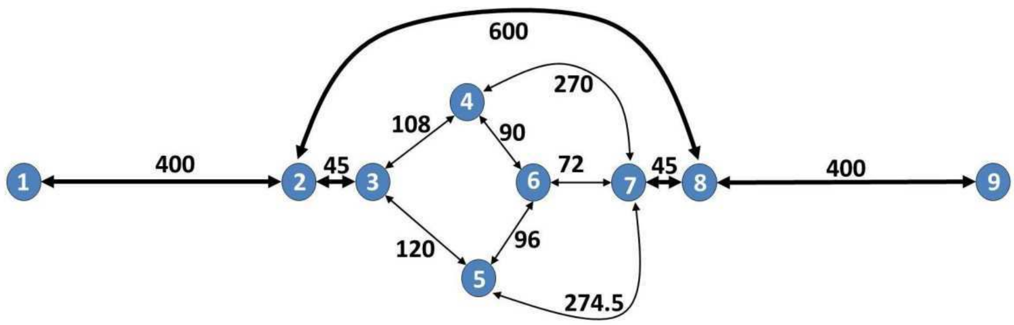

Figure 3: Toy City Network

The black arrows indicate the (two way) road links with free flow travel cost indicated. Link cost functions are BPR-type with details given in table 3 in the appendix. The ten OD movements are shown in Figure 4 with

[2] Indeed the equilibrium flows themselves can, in favourable circumstances, only be expressed as the solution to an optimisation problem.

the demands levels indicated next to the destination (e.g. 850 from 8 to 6). For reference the OD movements are grouped into URBAN movements within the "city", COMMUTER movements to/from the outskirts and THROUGH traffic that may go through the city or use the bypass.

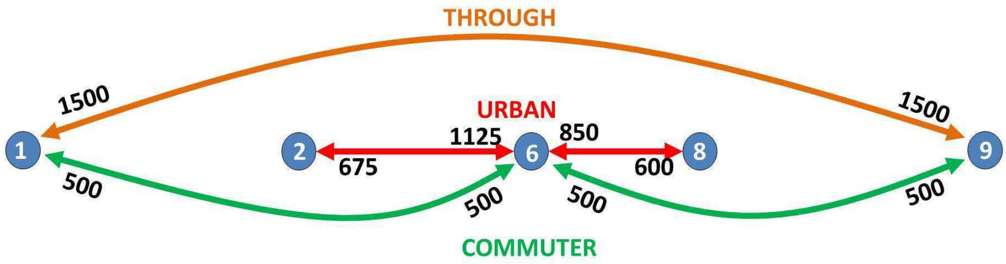

Figure 4: Toy City OD Movements

For convenience we number these ten OD movements in Table 2

Table 2: Toy City OD indices

OD	6→1	9→1	6→2	1→6	2→6	8→6	9→6	6→8	1→9	6→9
OD Index	1	2	3	4	5	6	7	8	9	10

We now consider changing the demand only on OD 1→6, from the base demand $Q = 500$ to $Q + \Delta Q$ with $\Delta Q \in [-300,300]$, fixing the other OD demands.

To reveal the impact of this demand change on each OD cost, we re-solve UE at each demand level (graphs in Fig 5 are based on interpolating 51 UE evaluations). We hence obtain numerical estimates of the underlying OD cost functions which we know to be single valued and continuous under standard assumptions (Patriksson, 2004). As we would expect, the network has interdependent OD demand vulnerabilities. Changing the demand level on OD 1→6 affects the cost on some of the other OD movements, as shown in Figure 5. Similar plots arise when changing the demand on other OD movements; each OD cost is a function of all ten OD demand levels. It is these continuous, multivariate single-valued functions, whereby the equilibrated OD cost is a function of all OD flows, that we aim to efficiently capture with our aggregation method. These functions implicitly take account of the assumed routing principle (UE in this case).

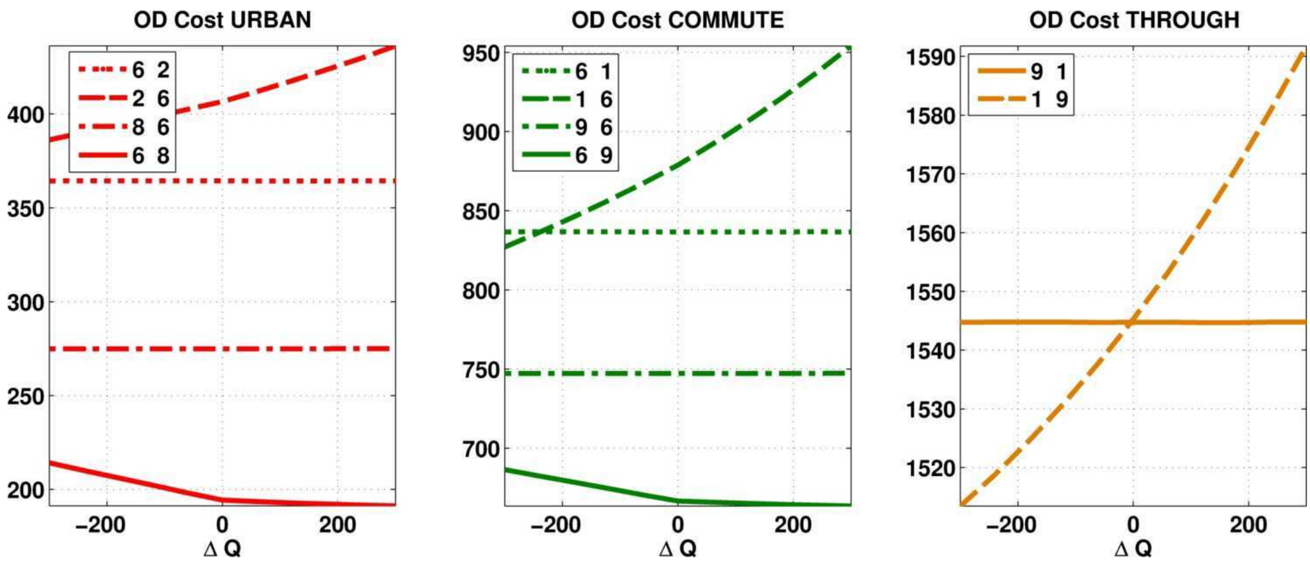

Figure 5: UE Toy City Network: increasing OD 1→6

4.3 The Case of SUE: A two-link network example

In sections 4.1 and 4.2, we saw how it was possible under the UE model to derive explicit, analytic aggregate relationships for small networks, though this becomes infeasible for larger scale problems. In the case of SUE, even in the simplest case of a two-link network, deriving explicit analytic relationships is not feasible. On the other hand, with some modifications, similar principles can be adopted for the case of SUE, and so the purpose of the present section is to illustrate how the UE-based principles are adapted, even though we cannot have recourse to the same analytical examples as for UE.

Following the definition of SUE in equations (4)-(8), we focus on the specific case of logit SUE where the random terms ϵ_k^r are independent and identically distributed, following the standard Gumbel distribution. The equilibrium link flows cannot be written in closed form even for the simple two link network considered above. The logit SUE flows can only be implicitly defined as satisfying

$$x_1^* = q \frac{\exp\left(-\theta\, c_1(x_1^*)\right)}{\exp\left(-\theta\, c_1(x_1^*)\right) + \exp\left(-\theta\, c_2(q - x_1^*)\right)} \tag{11}$$

With $x_2^* = q - x_1^*$. Moreover, unlike the case of UE, for SUE both paths/links have non-zero flow at all demand levels and the costs on the two links are not equal at equilibrium. The composite cost is

$$W^*\left(c(q,\theta)\right) = -\frac{1}{\theta}\ln\left[\exp\left(-\theta c_1(x_1^*)\right) + \exp\left(-\theta c_2(x_2^*)\right)\right] \tag{12}$$

The SUE function $W^*\left(c(q,\theta)\right)$ is the analogue of the UE function $T^*(q)$ (see Eq (5), Table 1, Figure 5), again being uniquely determined by the underlying network. For the 2-link network we can easily compute and hence plot this function. Note that the logit dispersion parameter, θ, corresponds to a standard deviation in the perceived path costs of $\sigma = \pi/\theta\sqrt{6}$. The OD composite cost functions for $\theta = 5$ and $\theta = 0.5$ (corresponding to standard deviations in perceived path costs of ~1% and ~10% of the free flow costs) are given by the solid blue lines in figures 6 and 7. The green dashed and red dotted lines show the (mean perceived) costs on paths 1 and 2 respectively.

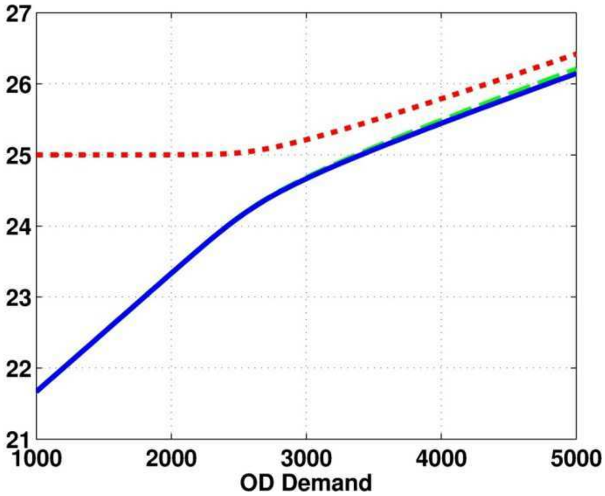

Figure 6: SUE 2-link network with $\theta = 5$.
OD composite cost (solid blue), link 1 cost (dashed green), link 2 cost (dotted red)

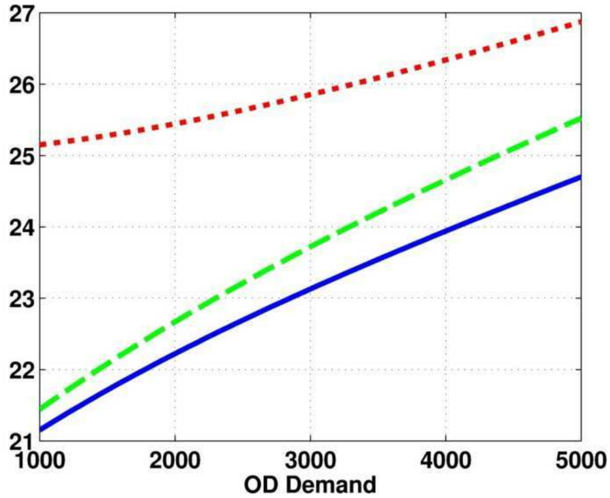

Figure 7: SUE 2-link network with $\theta = 0.5$.
OD composite cost (solid blue), link 1 cost (dashed green), link 2 cost (dotted red).

In Figure 6, careful inspection will reveal that there are in fact three curves, one of the curves almost indistinguishable from the solid (blue) line. Figure 6 is the SUE analogue of Figure 2 given earlier for UE, a key difference in the SUE being that both routes are used at all levels of demand. The two figures are also comparable in how they measure “OD cost”: in Figure 2 this is the solid blue line, which coincides with the cost on the link that is used at all levels of demand (since this is always a minimum cost route). That is to say, Figure 2 shows three curves, but two of them coincide. In Figure 6 the OD cost is replaced by the composite cost function, again represented by a solid blue line. At this high value of theta it is almost indistinguishable from the curve showing the cost on link 1 (green dashed line, barely visible beneath it). Thus we can see that, as expected for high values of theta, our aggregate SUE relationship closely resembles that of the UE case.

However, we need to take a little care in the interpretation of the composite cost function. A tempting interpretation might be that it is somehow representing an average of the actual costs on the used routes, but this is not true. This is evident from the inequalities (9) noted at the end of section 2, and can be seen clearly from Figure 7 based on a lower value of the scaling parameter (hence relatively more variance in utility). Consistently with the inequalities (9), Figure 7 shows how the composite cost function (blue line) is always below the minimum of the actual route costs (the other two lines). This is reflecting the fact that the logit model is capturing the benefits of variety; it is simple to see from the composite cost expression, that even in the case of two alternatives with equal systematic/mean cost, the composite cost (expressed in cost terms, as we do throughout this paper) is always less than this equal mean cost. This is important to understand in the sense of how the OD composite cost might be used and interpreted, when derived.

This simple example illustrates the main conceptual issues in transferring from a UE-based to SUE-based analysis. For general networks with R OD movements, each of the OD composite costs will be a function of all R OD demand flows. Individual values of all OD composite costs can be computed by solving SUE at a specific OD demand matrix. However, as was the case for UE, the computational burden of resolving SUE many times would make this infeasible for all but small networks. In the remainder of the paper we show how the functional form of the SUE-based aggregate relationships can be efficiently approximated.

5. Methodology for constructing SUE-based aggregate relationships

The examples above aim to illustrate that the computing the vulnerability of a network’s performance to changes in OD demands is (in general) not feasible by re-calculating traffic equilibrium at every combination of demand settings. Analytic approximation of the multidimensional OD demand-performance relationships, i.e. network aggregation, offers a solution. Our objective now is to present a formal method for network aggregation, which is applicable to large-scale networks; in particular, we will aim to produce estimates of the full, equilibrated OD cost *versus* OD matrix relationships using only a *single* run of an equilibrium algorithm. As we noted earlier, for general networks under SUE, the OD composite costs can in principle be computed by re-solving SUE many times; just as was the case for the analogous OD travel cost function, $T^{r*}(\boldsymbol{q})$, under UE. This is computationally expensive, prohibitively so for large networks with many OD movements. For the case of SUE we can take advantage of established sensitivity analysis results for the equilibrium flows, and in this way efficiently derive approximations to the OD composite cost functions.

We wish to approximate the changes in costs across the network as OD demands change under the SUE route choice model. Our method stems from the sensitivity analysis derived in Davis (1994) and provides approximations to the variation in SUE link flows under perturbations to network parameters. Consider the OD demands, $\boldsymbol{q}$, as functions of $\boldsymbol{s}$: $\boldsymbol{q} = \boldsymbol{q}(\boldsymbol{s})$. Following the analysis in Connors et al (2007), we find that

$$\boldsymbol{x}^*(\boldsymbol{q}+\boldsymbol{s}) \approx \boldsymbol{x}^*(\boldsymbol{q}) + (\boldsymbol{I} - \Delta \cdot \boldsymbol{q} \cdot \nabla_c \boldsymbol{P} \cdot \Delta^T \cdot \nabla_{\mathbf{x}} \boldsymbol{t}\,)^{-1} \Delta \cdot \mathbf{P} \cdot \nabla_{\mathrm{s}} \mathbf{q}(\mathbf{s}) \cdot \mathbf{s} \tag{13}$$

Evaluating the Jacobian matrices of link costs with respect to link flows, $\nabla_{x}\boldsymbol{t}$, and path choice probabilities with respect to path costs, $\nabla_{c}\boldsymbol{P}$, at the unperturbed SUE solution, we can collect terms together into a constant matrix $\boldsymbol{A}$ multiplying the perturbation parameter $\boldsymbol{s}$

$$\boldsymbol{x}^{*}(\boldsymbol{q}+\boldsymbol{s}) \approx \boldsymbol{x}^{*}(\boldsymbol{q})+\boldsymbol{A}\cdot\mathbf{s} \tag{14}$$

The vector $\boldsymbol{s}$ will have zeros in those positions where OD demands are not being perturbed.The sensitivity analysis result (18) therefore allows us to write the SUE link flows as a function of the OD demand perturbations. Using (18) we can compute approximate link flows and hence link costs, sum them accordingly to get path costs and hence calculate the composite cost for each OD as a function of the OD demands:

$$W^{r}=W^{r}\left(\boldsymbol{c}\big(\boldsymbol{x}^{*}(\boldsymbol{q}+\boldsymbol{s})\big)\right) \approx W^{r}\big(\boldsymbol{c}(\boldsymbol{x}^{*}(\boldsymbol{q})+\boldsymbol{A}\cdot\boldsymbol{s})\big) \tag{15}$$

For logit SUE, where the random terms in are drawn from Gumbel distributions, the composite cost function can be written in closed form. We can therefore write (an approximation to) the composite cost for each OD explicitly as a function of the OD demand perturbations:

$$W^{r}\left(\boldsymbol{c}\big(\boldsymbol{x}^{*}(\boldsymbol{q}+\boldsymbol{s})\big)\right) \approx \ln\left(\sum_{k\in K^{r}}\exp\big(-\theta c_{k}^{r}(\boldsymbol{x}^{*}(\boldsymbol{q})+\boldsymbol{A}\cdot\boldsymbol{s})\big)\right) \tag{16}$$

The composite costs (21) can thus be constructed for any new OD demand profile $\boldsymbol{q}+\boldsymbol{s}$; they depend on the OD demand perturbations, $\boldsymbol{s}$, and through the sensitivity analysis incorporate the 'network effect' where changes in demand on any one OD movement causes congestion and hence re-routing of traffic on other OD movements. Both the underlying OD composite cost function (16), and the approximation to it given by (16), are uniquely defined by the underlying network. While (16) is an exact aggregate representation, (16) is an approximation in the degree to which demand perturbations have a non-linear impact on the SUE link flows. Noting that this approximation is based on the truncated Taylor expansion shown in (13), one might expect it would only be accurate for very small perturbations, $\boldsymbol{s}$. However, as anticipated as a result of the demand-perturbed numerical experiments in Clark & Watling (2002), the range of validity for this approximation may be much wider, and this issue is further examined numerically below.

6. Numerical Experiments

6.1 Example 1: The 2-link network

We consider the same two link network as before (section 4.1 and 4.4) under different levels of demand $Q=Q_0+\Delta Q$ with $Q_0=2000$ and $\Delta Q\in[-1000,3000]$. At each demand level we re-solve SUE (as we did in section 3.4) and compare these results with those based on a sensitivity analysis (SA) executed at Q_0, using the methodology described in section 5. The link flows predicted by the SA, x_{SA}, at each demand setting are used to compute the OD composite cost and the SUE gap function, to indicate how good an approximation they provide to the true SUE link flows. In the figures 8 and 9 we plot with solid blue lines the accurate solutions from re-solving SUE at each demand setting. The black dotted lines come from estimating the flows via sensitivity analysis executed at Q_0. Two values of the logit dispersion parameter are shown, $\theta=5, 0.5$, as before.

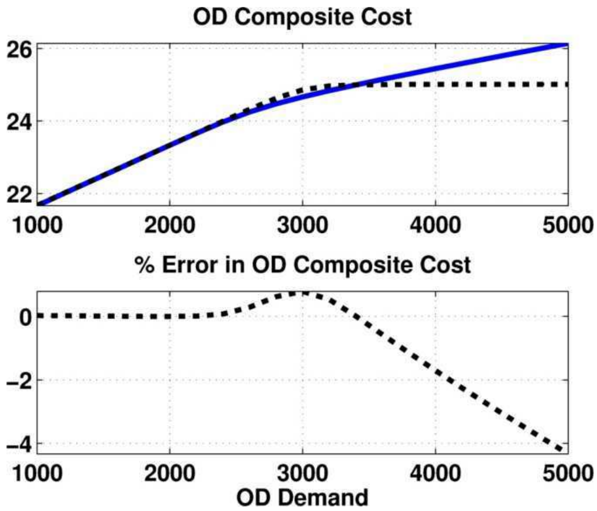

Figure 8: SUE 2-link network approximated OD travel cost function with $\theta = 5, Q_0 = 2000$

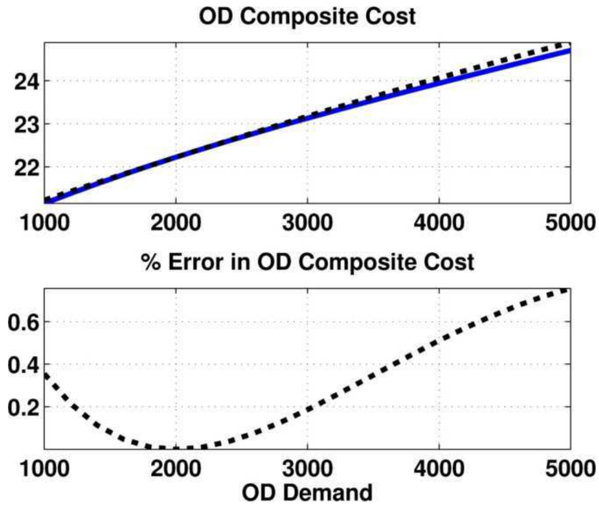

Figure 9: SUE 2-link network approximated OD travel cost function with $\theta = 0.5, Q_0 = 2000$

The solid blue lines show the true SUE solution(s) and the black dashed lines correspond to approximate SUE solutions. With little dispersion ($\theta = 5$) the behaviour of SUE is rather close to that of UE (compare Figure 2 & 6). Recall that the OD cost for UE is non-differentiable where the path set changes at $Q = 3000$. Here the SUE sensitivity analysis is computed at $Q_0 = 2000$, and the SA approximated OD composite cost remains good (<1% error) over the substantial range $Q \in [Q_0 - 1000, Q_0 + 1500]$. (We remark in passing that in order to have applied a UE-based sensitivity approach in the same context, evaluated at only a single UE solution, we would need to have entirely ignored any path set change.) As the variance of SUE perceived costs increases (to $\theta = 0.5$), there is a systematic improvement in the SA approximation (note y-axis scale in the lower plot of Figure 9).

6.2 Example 2: Toy City network changing a single OD demand

We consider the toy city as shown in Figures 3 and 4, with the same parameters as before though now under SUE. We set the logit parameter for each OD to be 0.03 so that the weighted average path cost at SUE is 914.2 and the approximate standard deviation in perceived path cost is 42.8 or 4.7%. The sensitivity analysis matrix $\boldsymbol{A}$ from (13) is 24x10 since there are 24 links and 10 OD movements. It is computed only once (at the base demand levels). The aggregate network has all ten OD movements represented, and the non-separable OD composite cost functions are each functions of the ten OD demand levels. The approximate OD composite cost functions derived from the sensitivity analysis are nonlinear multivariate functions with fixed coefficients.

To illustrate the performance of the multivariate aggregate network representation, we first perturb the single OD $1 \to 6$ over the range $\Delta Q \in [-300,300]$. Since the base demand on this OD is 500, this is a perturbation of up to 60%. As OD $1 \to 6$ is perturbed, the composite cost changes on this OD and on all other ODs, as shown in the upper plots of Figure 10. Increasing demand on OD $1 \to 6$ increases the OD composite cost most noticeably on ODs $1 \to 6, 2 \to 6$ & $1 \to 9$. The cost on ODs $6 \to 8$ & $6 \to 9$ decrease resulting from through traffic diverting to use the bypass (See Figure 4). Note that OD costs on "right to left" movements (e.g. ODs $6 \to 2, 8 \to 6$ & $9 \to 1$) are unaffected.

For each OD the SA approximation errors across this range of perturbations are never more than 0.5%; these are shown by dotted black lines in the upper plots which are mostly obscured by the "true" lines from re-solving SUE at each new demand level. The lower plots show the % error in the SA approximated OD costs, the worst being OD $6 \rightarrow 8$.

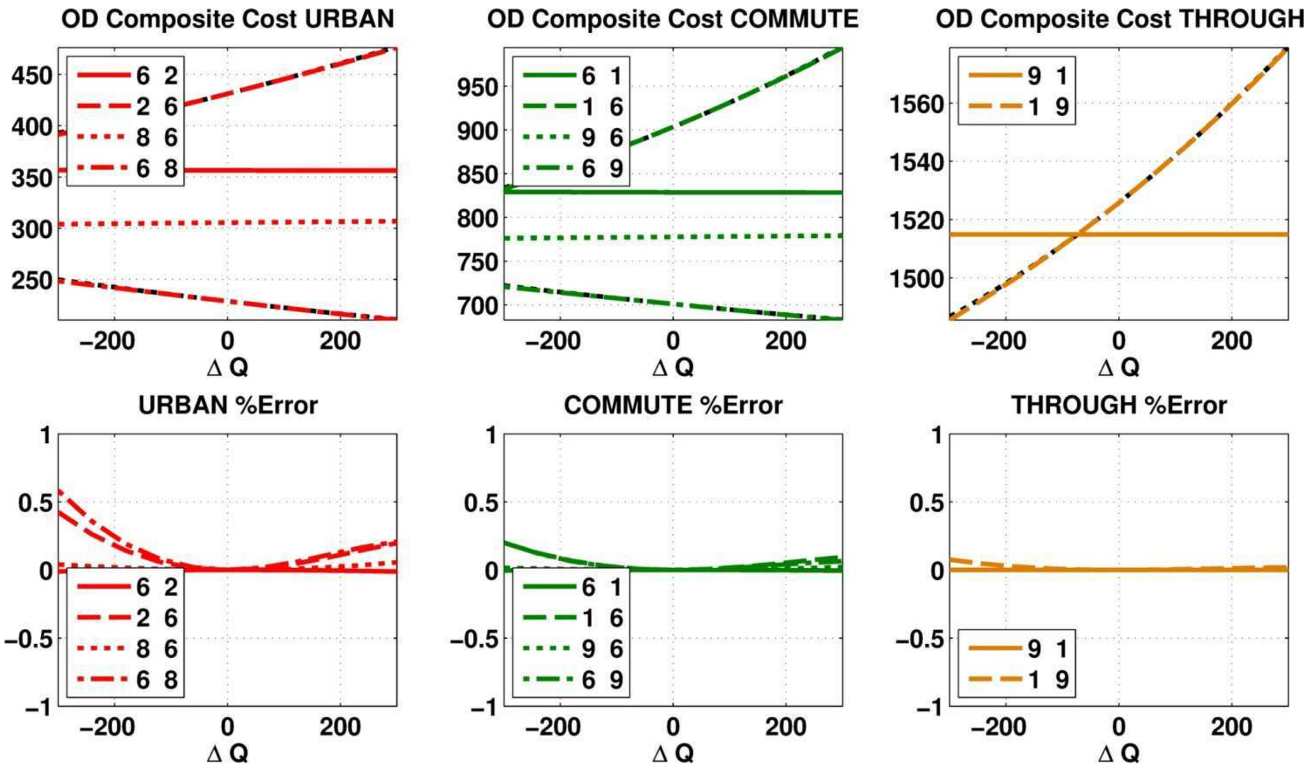

Figure 10: Toy City SUE: Accuracy of the SA approximated OD costs

The same sensitivity analysis approximation used above captures the network re-routing effects induced by changing congestion. From the SA results, Figure 11 shows how path flows for OD $1 \rightarrow 9$ are affected by demand increasing on OD $1 \rightarrow 6$, with paths numbered in the legend. Path 89 uses the bypass, whereas paths 85 & 86 go via node 6 and paths 87 & 88 use the "inner" ring traversing nodes 3-4-7 and 3-5-7 respectively.

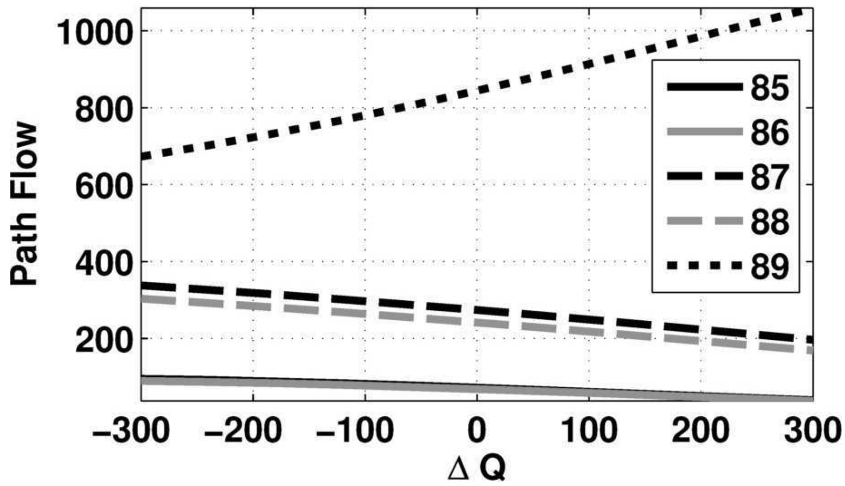

Figure 11: Path Flows on OD 1→9 as demand increases on OD 1 → 6; flow diverts to the bypass

As the demand on OD $1 \rightarrow 6$ is increased and causes congestion on the city centre links, so traffic on other OD movements diverts onto the bypass in order to avoid the city centre. The impact of such interactions on

OD costs are captured by the multidimensional OD cost functions of the aggregated network. These results display the expected consequences of a network equilibrium model, but highlight the interaction between OD demands and OD costs, and show the accuracy of the functional relationship between OD demands and costs given by the SA approximation (16).

The SA described above can be used to get a snap shot of various network characteristics at the current demand level. The *demand vulnerability matrix* shows the sensitivities of each OD cost to each of the network wide OD demands. Figure 12 shows for each OD pair (numbered as in Table 2) the implied point elasticities, namely the percentage change in OD composite cost resulting from a 1% demand increase on each OD in turn. The positive diagonal elements show that OD composite cost increases with OD flow, but there are significant off-diagonal interactions, including OD composite cost reductions.

The demand vulnerability matrix helps to identify the most influential OD trip demands on the composite cost for any given OD movement. For example, considering the OD movement 6→8, which referring to Table 2 corresponds to row 8 of the matrix in Figure 12, we can see that the most influential OD movements are from column 5 (OD 2→9) and 10 (6→9), whereas OD movement 9→6 (row 7) is influenced by column 6 (OD 8→6). Having identified the most relevant aggregate dependencies, the functional relationship between demand and composite cost can immediately be plotted using (16), as shown for this example in Figure 13.

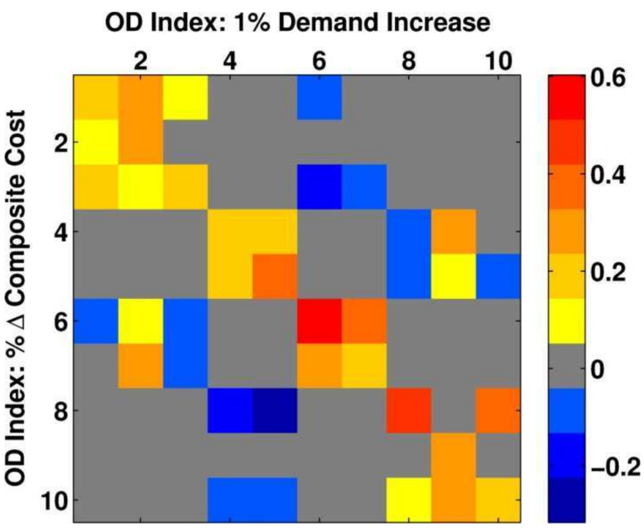

Figure 12: Demand vulnerability of the Toy City network (OD indices displayed as in Table 2)

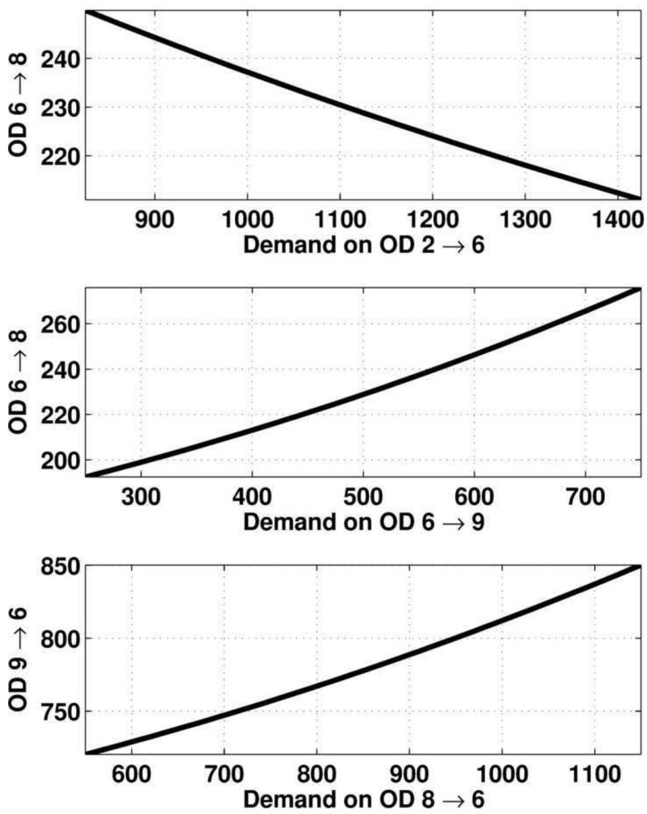

Figure 13: Composite cost as a function of demand for off-diagonal terms

6.3 Example 3: Toy City Urban Aggregation

The aggregation method described naturally allows other approaches to summarising the interactions between demand and OD travel cost. For example to illustrate the impact on OD costs of changes to city centre demand levels. We can aggregate the urban part of the Toy City network, resulting in the network shown in Figure 14, though here the links are no longer simple supply links. We seek to represent only the *commuter* and *through* OD movements, while taking into account the impact of urban congestion and the urban road network (including the bypass).

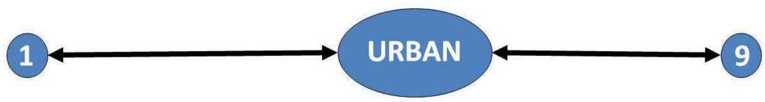

Figure 14: Reduced network with urban area aggregated

As above, at the base demand we compute the SUE flows. In the full resolution network, the composite cost on each OD movement r is a function of the four urban demands, q_U^i, the four commuter demands, q_C^i, and the two through demands, q_T^i.

$$W^r = W^r\left(\boldsymbol{c}\left(\boldsymbol{x}(q_U^1, q_U^2, q_U^3, q_U^4, q_C^1, q_C^2, q_C^3, q_C^4, q_T^1, q_T^2)\right)\right)$$

The network simplification step is to group together the impact of the urban traffic level on the *through* and *commuter* OD movements. For the SA, we consider a simultaneous uniform additive perturbation to each of the urban demands.

$$\boldsymbol{q} = \left(q_U^1, q_U^2, q_U^3, q_U^4, q_C^1, q_C^2, q_C^3, q_C^4, q_T^1, q_T^2\right)^T = \boldsymbol{q_0} + \boldsymbol{\Delta}\mathbf{Q}\ (1,1,1,1,0,0,0,0,0,0)^T$$

The single parameter $\boldsymbol{\Delta Q}$ indicates the change in the total amount of urban traffic, compared with the base case. While every OD cost is undoubtedly a function of all OD demands, we optimistically make the approximation that each OD cost can be approximated by the univariate function $W^r\left(\boldsymbol{c}(\boldsymbol{x}(\boldsymbol{q_0} + \boldsymbol{\Delta Q}))\right)$ from (16), having generated the relevant Jacobian matrices for this case, which are easily derived from (13).

Note that the urban OD demands and consequent congestion, and the contribution of the bypass are already encoded into the *commuter* and *through* OD cost functions derived above from the sensitivity analysis in (13). We perturb the four URBAN ODs simultaneously (additively) by the same amount, $\Delta Q \in [-300,300]$ and show the OD cost changes in Figure 15 below. As before, the upper plots show the change in the OD composite costs for each group of OD movements, and the lower plots show errors in the SA approximation. The six OD movements of the aggregated network (Figure 6) are shown in the centre and right hand plots. The internal urban movements are also captured from the same SA (shown in the left hand plots).

As the level of urban traffic fluctuates, the *commuter* and *through* OD costs change. Here the additive perturbation to each urban OD is ± 300 and hence the total change is up to $36\% = 1200/(1125 + 675 + 850 + 600)$. Figure 15 shows that the SA aggregation approach has captured most of the resulting OD cost variation, with errors of less than 1.5% for the ODs of interest.

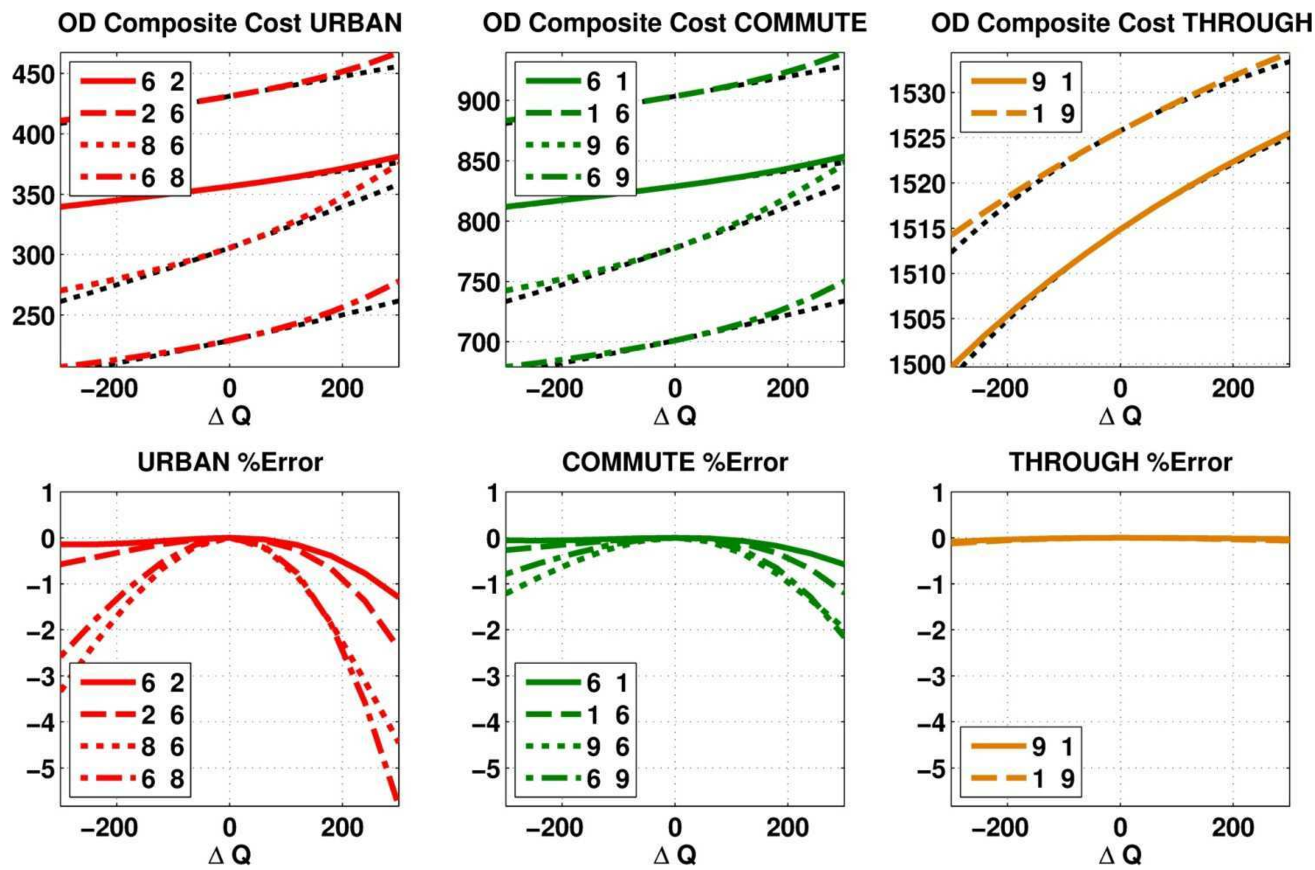

Figure 15: Reduced network with urban area aggregated

The aim here was to get good estimates of the *commuter* and *through* OD cost functions, nevertheless we can easily evaluate (from (16)) the *urban* OD costs. Here the errors are higher, but within 6% over this range.

6.4 Example 4: Larger Network Aggregation

We consider the Anaheim network (network parameters from www.bgu.ac.il/~bargera/tntp/) having 416 nodes, 914 links and 1406 OD movements. The base demand leaves this network rather uncongested and the accuracy of the SA approximation is more likely to be challenged when links are congested. To this end, we set the base demand to be 5x the original base demand matrix. We set the logit dispersion parameter for each OD independently, in order that the standard deviation of the distribution of perceived OD costs is approximately 10% of the mean OD cost at the (new) base demand level. The wide range of OD costs results in dispersion parameters ranging from 0.0043 to 20.7.

In this scenario we select a single OD movement, namely OD 5 →2, and consider how this OD cost changes as the level of network-wide congestion increases (Figure 16). The choice of OD is somewhat arbitrary[3], though we purposefully selected an OD with a relatively long shortest path comprised of high capacity links, representing a possible through-route for the network.

We follow a similar approach to the previous section, simultaneously adding a uniform demand perturbation to every other OD (all excepting OD 5 →2) in the network. For the OD movement of interest we therefore have a univariate OD composite cost function.

[3] While a map is provided on Bar-Gera's website the map node numbers do not match the data file(s) and hence it cannot be used to identify ODs.

We add demand to every OD, except for OD 5 →2, such that the total network-wide demand increases by up to 50% from its original level. Figure 16 shows the cost on OD 5 →2 as the network congestion level rises, by re-solving SUE in each demand scenario (solid orange line) and from the single base demand SUE computation and SA aggregate approximation (black dotted line). This OD is not particularly congested (see the low values for OD composite cost). In Figure 17 we also display the four highest cost OD movements in the network (ODs listed in the legend) and show how their OD composite costs vary, along with the SA aggregation approximation (black dotted lines) for each of them. The errors in the SA approximations remain below 1% throughout.

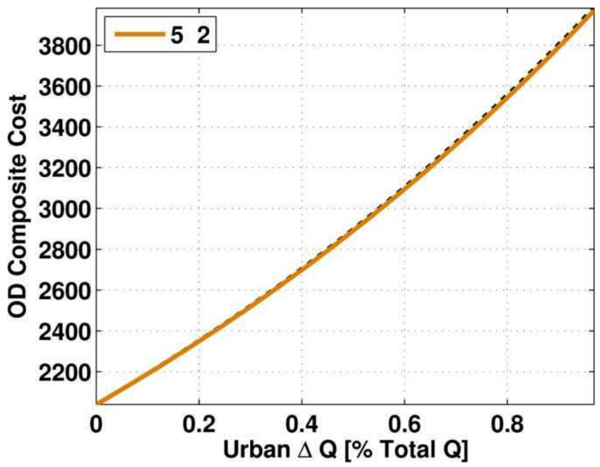

Figure 16: Anaheim network OD 5 → 2 Composite cost as network traffic increases (solid line). SA approximated composite cost (dashed black line)

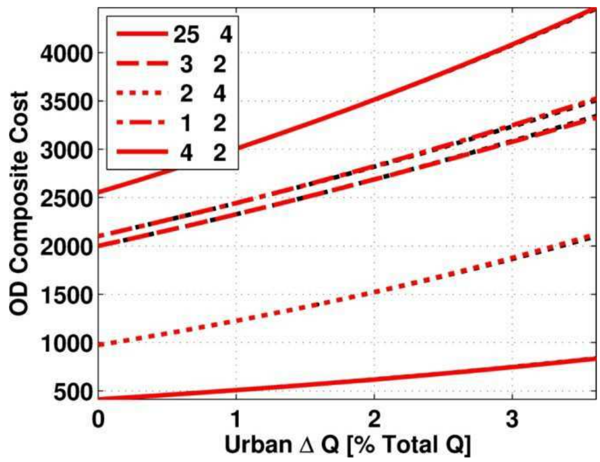

Figure 17: Largest Urban OD Composite costs as network traffic increases

As for the toy city above, we can also investigate the demand vulnerability matrix of this network. We display the percentage change in OD composite cost due to a 1% increase in one other OD; a matrix of demand elasticities. The diagonal terms represent increases in OD cost due to increase in demand on the same OD. All non-zero off-diagonal terms show interactions between different OD movements that are crucial to capture to understand network congestion under changing demand. The OD indices are ordered by destination and hence there is some spatial grouping which is reflected in the block diagonal banding. However, the significant interactions that occur far from the diagonal cannot be interpreted as spatially separate nor independent. As in Section 6.2, Figure 18 (on close inspection) allows demand vulnerabilities to be identified, such as those illustrated in Figure 19.

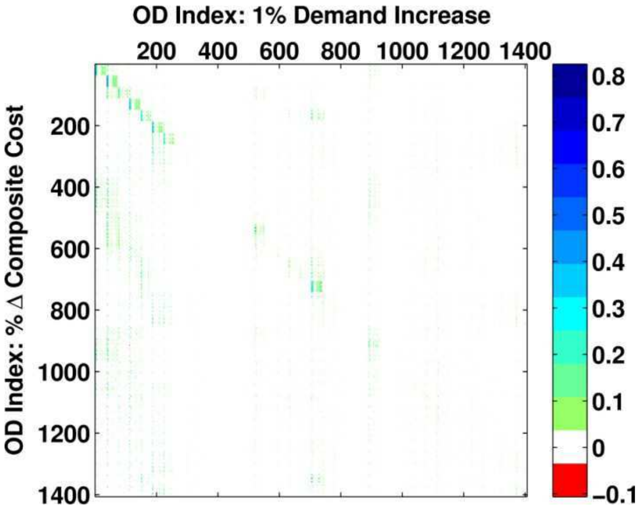

Figure 18: OD composite cost vulnerability to OD demand increase.

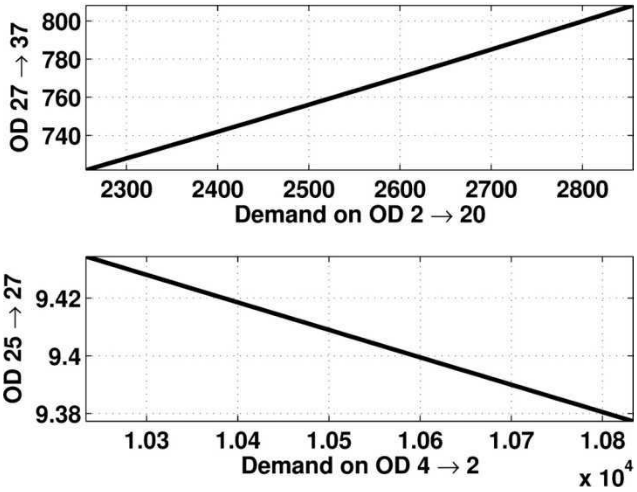

Figure 19: Composite cost as a function of demand for off-diagonal terms.

7. Conclusions

This paper has reported an approach to analyse demand vulnerability by means of analytic network aggregation. The first step in this process is to analyse the impact on mean perceived OD travel costs (formally, the 'composite costs') of changes to the OD travel demand matrix, with the process of re-equilibration at each stage implicitly embedded in this relationship. Sensitivity analysis is used for this step to avoid the need to re-solve equilibrium at many points, yielding an explicit, unique functional relationship between OD flows and OD composite costs ('explicit' meaning that an equilibrium problem need not be re-solved each time the OD flows are changed). This has removed the network/equilibrium component, but the resulting problem is not necessarily of lower dimension than the original. There are several ways in which this information may then be subsequently exploited in the overall aggregation process. One way, as illustrated, is to focus on a subset of OD movements and agglomerate others as a kind of overall 'traffic intensity', which reflects that as demands on these agglomerated movements grow they may delay the travel on our OD movements of interest. The resulting problem is an aggregation/simplification of the original problem, consisting of many fewer dimensions of 'flow' and 'travel cost', compared with the number of OD movements, links and paths in the original network. This approach is inherently multi-dimensional and allows the interactions of all ODs to be captured.

In this paper, we have focused on a particular application of the proposed aggregation methodology, to identify network demand vulnerabilities. However we believe this approach is sufficiently general that it may be extended to address other types of network aggregation problem. The issue of zonal aggregation and the problem of devising a methodology consistent with network aggregation remains a ripe area for further investigation.

APPENDIX: TOY CITY LINK COST FUNCTION PARAMETERS

$$t(x) = ff + B\left(\frac{x}{C}\right)^{\beta}$$

Table 3: Link Cost Function Parameters for Toy City Network

Link	From	To	ff	B	C	β
1	2	1	400	50	1800	3.5
2	1	2	400	50	1800	3.5
3	3	2	45	9.55	1800	4.5
4	8	2	600	10	2200	3.5
5	2	3	45	9.55	1800	4.5
6	4	3	108	108	1100	3
7	5	3	120	120	1100	3.1
8	3	4	108	108	1100	3
9	6	4	90	90	1100	3.2
10	7	4	270	57.27	1100	3.5
11	3	5	120	120	1100	3.1
12	6	5	96	96	1100	3.1
13	7	5	274.5	58.23	1100	3
14	4	6	90	90	1100	3.2
15	5	6	96	96	1100	3.1
16	7	6	72	72	1100	3.1
17	4	7	270	57.27	1100	3.5
18	5	7	274.5	58.23	1100	3
19	6	7	72	72	1100	3.1
20	8	7	45	9.55	1800	4.5
21	2	8	600	10	2200	3.5
22	7	8	45	9.55	1800	4.5
23	9	8	400	50	1800	3.5
24	8	9	400	50	1800	3.5

References

Ardekani S and Herman R (1987) Urban network-wide traffic variables and their relations. *Transportation Science*, 21(1), 1–16.

Batty M and Sikdar PK (1982a) Spatial Aggregation in Gravity Models: 1. An Information-Theoretic Framework. *Environment and Planning A*, 14(3), 377–405.

Batty M and Sikdar PK (1982b) Spatial Aggregation in Gravity Models: 2. One-dimensional population density models. *Environment and Planning A*, 14(4), 525–553.

Batty M and Sikdar PK (1982c) Spatial Aggregation in Gravity Models: 3. Two-dimensional trip distribution and location models. *Environment and Planning A*, 14(5), 629–658.

Batty M and Sikdar PK (1982d) Spatial Aggregation in Gravity Models: 4. Generalizations and Large-Scale Applications. *Environment and Planning A*, 14(6), 795–822.

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

Berdica, K. (2002) An introduction to road vulnerability: what has been done, is done and should be done., Transport Policy, 9, 117-127.

Bouthelier F and Daganzo CF (1979) Aggregation with Multinomial Probit and Estimation of Disaggregate Models with Aggregate Data - a New Methodological Approach. *Transportation Research Part B-Methodological*, 13(2), 133–146.

Bovy PHL and Jansen GRM (1983) Network Aggregation Effects upon Equilibrium Assignment Outcomes: An Empirical Investigation. *Transportation Science*, 17(3), 240–262.

Boyles SD (2012) Bush-based sensitivity analysis for approximating subnetwork diversion. *Transportation Research Part B: Methodological*, 46(1), 139–155.

Cantarella GE (1997) A general fixed-point approach to multimode multi-user equilibrium assignment with elastic demand. *Transportation Science*, 31(2), 107–128.

Chan Y (1976) Method to Simplify Network Representation in Transportation-Planning. *Transportation Research*, 10(3), 179–191.

Chang KT, Khatib Z and Ou YM (2002) Effects of zoning structure and network detail on traffic demand modeling. *Environment and Planning B-Planning & Design*, 29(1), 37–52.

Chen, A., Yang, H., Lo, H.K., Tang, W.H. (2002). Capacity reliability of a road network: an assessment methodology and numerical results. Transportation Research 36B (3), 225–252.

Chen, B.Y., Lam, W.H.K., Sumalee, A., Li, Q., Li, Z.-C. (2012). Vulnerability analysis for large-scale and congested road networks with demand uncertainty. Transportation Research A 46, 501-516.

Clark, S.D., Watling, D.P. (2002). Sensitivity analysis of the probit-based stochastic user equilibrium assignment model. *Transportation Research Part B* 36 (7), 617–635.

Connors RD, Sumalee A and Watling DP (2007) Sensitivity analysis of the variable demand probit stochastic user equilibrium with multiple user-classes. *Transportation Research Part B: Methodological*, 41(6), 593–615.

Dafermos S (1980) Continuum Modeling of Transportation Networks. *Transportation Research Part B-Methodological*, 14(3), 295–301.

Daganzo CF (1980a) An Equilibrium Algorithm for the Spatial Aggregation Problem of Traffic Assignment. *Transportation Research Part B-Methodological*, 14(3), 221–228.

Daganzo CF (1980b) Network Representation, Continuum Approximations and a Solution to the Spatial Aggregation Problem of Traffic Assignment. *Transportation Research Part B-Methodological*, 14(3), 229–239.

Daganzo CF (2007) Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B-Methodological*, 41(1), 49–62.

Daganzo CF and Geroliminis N (2008). An analytical approximation for the macroscopic fundamental diagram of urban traffic. *Transportation Research Part B: Methodological* 42 (9), 771–781.

Daniele P, Idone G and Maugeri A (2003) *The continuum model of transportation problem*. Daniele P, Giannessi F, and Maugeri A (eds).

Davis G (1994) Exact Local Solution of the Continuous Network Design Problem Via Stochastic User Equilibrium Assignment. *Transportation Research Part B-Methodological*, 28(1), 61–75.

D'Este, G.M. and Taylor, M.A.P (2001) Modelling network vulnerability at the level of national strategic transport network, Journal of East Asia Society for Transport Studies 4(2), 1-14.

Friesz T (1985) Transportation Network Equilibrium, Design and Aggregation - Key Developments and Research Opportunities. *Transportation Research Part a-Policy and Practice*, 19(5-6), 413–427.

Hearn DW (1984) Practical and theoretical aspects of aggregation problems in transportation planning methods. In: Florian M (ed.), *Transportation Planning Models*, Amsterdam, North-Holland, pp. 257–287.

Herman R and Prigogine I (1979) A Two-Fluid Approach to Town Traffic. *Science*, 204(4389), 148–151.

Hills PJ (2001) Supply curves for urban road networks - A comment. *Journal of Transport Economics and Policy*, 35, 343–348.

Ho HW, Sumalee A, Lam WHK & Szeto WY (2013). A continuum modelling approach for network vulnerability analysis at regional scale. Paper presented at 20th ISTTT and appearing in Procedia – Social and Behavioral Sciences 80, 846-859.

Jenelius E, Petersen T & Mattsson L-G (2006). Importance and exposure in road network vulnerability analysis. Transportation Research A 40, 537-560.

Knoop, V.L., Snelder, M., Van Zuylen H.J., Hoogendoorn, S.P. (2012). Link-level vulnerability indicators for real-world networks. Transportation Research A 46, 843-854.

Lam WHK, Shao H, Sumalee A (2008). Modeling impacts of adverse weather conditions on a road network with uncertainty in demand and supply. Transportation Research B 42(10), 890-910.

Lighthill M and Whitham G (1955) On Kinematic Waves .2. a Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences*, 229(1178), 317–345.

May AD, Shepherd SP and Bates JJ (2000) Supply curves for urban road networks. *Journal of Transport Economics and Policy*, 34, 261–290.

May AD, Shepherd SP and Bates JJ (2001) Supply curves for urban road networks - A rejoinder. *Journal of Transport Economics and Policy*, 35, 349–352.

Nakayama S, Watling DP (2014). Consistent formulation of network equilibrium with stochastic flows. Transportation Research B, in press.

Nicholson, A.J. and Dalziell, E. (2003) Risk evaluation and management: a road network reliability study. In: Iida and Bell (Eds), 'The Network Reliability of Transport', Pergamon-Elsevier, Oxford, 45-59.

Openshaw S (1977) Optimal zoning systems for spatial interaction models. *Environment and Planning A*, 9(2), 169–184.

Patriksson M (2004) Sensitivity analysis of traffic equilibria. *Transportation Science*, 38(3), 258–281.

Rogers DF, Plante RD, Wong RT, et al. (1991) Aggregation and Disaggregation Techniques and Methodology in Optimization. *Operations Research*, 39(4), 553–582.

Sbayti H, El-Fadel M and Kaysi I (2002) Effect of roadway network aggregation levels on modeling of traffic-induced emission inventories in Beirut. *Transportation Research Part D: Transport and Environment*, 7(3), 163–173.

Shao, H., Lam, W.H.K., Tam, M.L. (2006). A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand. Networks & Spatial Economics 6, 173-204.

Sheffi Y (1984) Aggregation and Equilibrium with Multinomial Logit-Models. *Applied Mathematical Modelling*, 8(2), 121–127.

Sheffi Y (1985) *Urban transportation networks*. Prentice-Hall Englewood Cliffs, NJ.

Smith MJ (1979) The existence, uniqueness and stability of traffic equilibria. *Transportation Research Part B: Methodological*, 13(4), 295–304.

Sullivan JL, Novak DC, Aultman-Hall L & Scott DM (2010). Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: A link-based capacity-reduction approach. Transportation Research A 44, 323-336.

Sumalee A (2004) Optimal Road User Charging Cordon Design: A Heuristic Optimization Approach. *Computer-Aided Civil and Infrastructure Engineering*, 19(5), 377–392.

Sumalee A & Watling DP (2008). Partition-based algorithm for estimating transportation network reliability with dependent link failures. Journal of Advanced Transportation 42(3), 213-238.

Sumalee A, Uchida K, Lam WHK (2011). Stochastic multi-modal transport network under demand uncertainties and adverse weather condition. Transportation Research C 19(2), 338-350.

Szeto, W.Y., O'Brien, L., O'Mahony, M. (2006). Risk-averse traffic assignment with elastic demands: NCP formulation and solution method for assessing performance reliability. Networks & Spatial Economics 6, 313-332.

Taylor, M.A.P., Somenahalli, V.C.S. and D'Este, G.M. (2006) Application of accessibility based methods for vulnerability analysis of strategic road networks, Networks and Spatial Economics, 6, 267-291.

Theil H (1957) Linear Aggregation in Input-Output Analysis. *Econometrica*, 25(1), 111.

Tsekeris T & Geroliminis N (2013) City size, network structure and traffic congestion, *Journal of Urban Economics* 76, 1–14.

Uchida K (2014). Estimating the value of travel time and of travel time reliability in road networks. Transportation Research B, in press. doi: 10.1016/j.trb.2014.01.002 .

Viegas JM, Martinez LM and Silva EA (2009) Effects of the modifiable areal unit problem on the delineation of traffic analysis zones. *Environment and Planning B-Planning & Design*, 36(4), 625–643.

Wardrop JG (1952) ROAD PAPER. SOME THEORETICAL ASPECTS OF ROAD TRAFFIC RESEARCH. *ICE Proceedings: Engineering Divisions*, 1(3), 325–362.

Watling, D.P. (2002). A second order stochastic network equilibrium model. II: Solution method and numerical experiments. Transportation Science 36, 167-183.

Watling D.P., Balijepalli, N.C. (2012). A method to assess demand growth vulnerability of travel times on road network links. Transportation Research A 46, 772-789.

Wong SC (1998) Multi-commodity traffic assignment by continuum approximation of network flow with variable demand. *Transportation Research Part B: Methodological*, 32(8), 567–581.

Wright I, Xiang Y, Waller L, et al. (2010) The Practical Benefits of the SATURN Origin-based Assignment Algorithm and Network Aggregation Techniques. In: Available from: http://trid.trb.org/view.aspx?id=1118347 (accessed 21 August 2013).

Xie F and Levinson D (2007) Measuring the structure of road networks. *Geographical Analysis*, 39(3), 336–356.

Xu Z and Sui DZ (2007) Small-world characteristics on transportation networks: a perspective from network autocorrelation. *Journal of Geographical Systems*, 9(2), 189–205.

Yang, H., Lo, H.K. and Tang,W.G. (2000) Travel time versus capacity reliability of a road network. In: Bell M.G.H and Cassir, C. (Eds) Reliability of Transport Networks. Research Studies, Baldock, Herts, 119-138.

Yang L & Qian D (2012). Vulnerability analysis of road networks. Journal of Transportation Systems Engineering and Information Technology 12(1), 105-110.

Zipkin PH (1980) Bounds for aggregating nodes in network problems. *Mathematical Programming*, 19(1), 155–177.