Exploring the possibility of combining discrete choice modelling and social networks analysis: an application to the analysis of weather-related uncertainty in long-distance travel behaviour

Discussion paper for presentation at the International Choice Modelling Conference, Leeds 4-6 July 2011

Alberto M Zanni and Tim J Ryley
Transport Studies Group, Department of Civil and Building Engineering
Loughborough University
a.m.zanni@lboro.ac.uk – t.j.ryley@lboro.ac.uk

Abstract

One deviation from the traditional theory of choice is that individuals, when facing a decision, are affected by the choices and behaviours of others. This discussion paper explores the way that this can be taken into account in discrete choice stated preference (SP) experiments. In particular, the possibility of integrating, to various extents, SP experiments and Social Networks Analysis (SNA) is discussed. Issues relating to social interactions in discrete choice are presented, as drawn from existing literature in economics and transportation. This is supplemented by work in progress on a study which attempts to analyse the way long-distance travellers react to uncertainty caused by extreme weather conditions. Finally, a number of theoretical, practical and empirical challenges are discussed, including the role of respondents reporting preferences of members of their networks, the information required on network members, and the treatment of endogeneity in the analysis.
1. Introduction

The traditional economic approach to analyse choice situations considers individual decision makers facing a set of alternatives, about which they possess perfect information, who, in total isolation, only considering their personal preferences, select the alternative bringing them the highest utility. Over the years, however, theoretical and applied economic studies, especially when integrating elements of psychology and sociology, have considerably discussed the validity of this axiom. First, it has been demonstrated that in real life settings choices are often carried out collectively rather than individually. This for example applies to a range of decisions that are taken by households collectively rather than by their single members. Second, it has been argued that individuals, when facing a decision, are affected by the choices and behaviour of others, either the entire society or particular others, and this has an important impact on their decision making process. All of these deviations from the traditional theory of choice have a number of practical applications when analysing choices in the context of transportation.

The aim of this discussion paper is to explore the way the second deviation from the traditional theory of choice, briefly introduced above, can be taken into account in discrete choice stated preference (SP) experiments and, in particular, assess the possibility of integrating, to various extents, SP experiments and Social Networks Analysis (SNA). Various theoretical and empirical issues are discussed in this paper. These issues are drawn both from the existing literature in economics and transportation and from an on-going work on a study which attempts to analyse the way long-distance travellers react to uncertainty caused by extreme weather conditions.

The remainder of this paper is organised as follows. Section 2 discusses the theoretical background. Section 3 briefly introduces SNA, while Section 4 reviews the existing literature. Section 5 introduces our application to uncertainty in long distance travel, while Section 6 discusses and concludes.

2. Background: why social dimensions in individual choice?

The analysis of choice in transportation contexts has developed over the years following economic choice theory, and important advances in the theoretical and empirical treatment of discrete choices have had their origin in the field of transportation research, for example the earlier studies by Ben-Akiva (1972), Bhat (1995), Hensher and Johnson (1981), Louviere (1988) and McFadden (1974), among many others.

As noted in the introduction, economic theory has traditionally looked at the choice mechanism as a simple process, albeit well structured, during which perfect information, consistent preferences, self-interest and expectations play an important role in determining choice behaviours and outcomes (McFadden, 2001). Both mainstream economic choice theory and transportation analysis have however recognised that such model of choice may give a limited picture of the array of determinants likely to have a role in affecting individuals’ choices. This has materialised in the development of different types of models capable of allowing for various interdependencies within the complex society in which individuals operate to be rigorously incorporated into economic theory (Zanella, 2007) and therefore provide a better explanation to complex aggregate social phenomena (Cont and Lowe, 2010; Soetevent, 2006).
In particular, the concept of *social interaction* has been introduced by Brock and Durlauf in their seminal paper (2001, p.235), as “the idea that the utility or payoff an individual receives from a given action depends directly on the choices of others in the individual’s reference group, as opposed to the sort of dependence which occurs through the intermediation of markets”. Interactions have been identified to be based on preferences, when the decisions over a choice set are dependent on the actions of the other agents; on expectations, when a decision maker considers the likely outcome of her decisions before making a choice and, for example, assess the outcome of similar people (peers) facing similar situations; and on constraints, when different users share a common resource, like in the case of congestion (Manski, 2000; Soetevent, 2006). Another important classification discussed in the economic literature considers *endogenous interactions* or *endogenous social effects*, when the behaviour of an agent is shaped by the way the others behave, *exogenous contextual interactions*, when the behaviour of an agent depends on common exogenous shocks or characteristics to which all group members are subject to (these also include simultaneity or non-random group selection and general characteristics like race, ethnicity and income), and *correlated effect*, when the members of a group behave similarly because they have similar, unobserved, characteristics, or face similar institutional environments (Ioannides and Zabel, 2008; Krauth, 2006; Manski, 1993, 2000; Soetevent, 2006).

In terms of their source, social interactions in decision making have been defined as dependent on two main dimensions, spatial and/or social (Akerlof, 1997; Hayakawa, 2000; Ioannides and Zabel, 2008). These two dimensions are not always separable and their influence depends not only on the nature of the effect but also on the examined economic variable (whether it is consumption, education, job search, unemployment, welfare decisions, etc). The spatial economics and econometrics literature has for example recognized the necessity of considering that a large number or economic decisions, especially those related to transportation and residence (although certainly not the only ones) display a spatial dependency or autocorrelation (Wang and Kockelman, 2009).

There a number of reasons why social, as well as spatial, influence should be considered when analysing transport situations. First of all, social reasons often generate the need of travelling. Activities in which individuals engage, and the consequent travelling, are in fact a direct consequence of the spatial locations of their social contacts and not only based on their personal characteristics (Carrasco and Miller, 2009; Farber and Paez, 2009; van der Berg et al., 2011). Secondly, exchanging information with other individuals in the social space has been identified as an important strategic tool for travellers, together with personal experience and information from transport operators (Avineri and Prashker, 2006; Denant-Boemon and Petiot, 2003), when facing uncertainty due to day-to-day variability in the performance of the transport systems (Bonsall, 2004). As argued by Schwanen (2008), travellers often react and cope with uncertainty not individually but as members of a social network. These networks are therefore an important source of information and decision support for individuals in the planning of activities and related trips, as they represent relatively low-cost choice heuristic solutions. Their support to decision-making can materialise in various way. Travellers may either simply conform to the behaviour of others (observed or unobserved) or directly ask for suggestions when choosing a departure time, a route, a mode or a vehicle. Neglecting the consideration of social interactions in the analysis of the way travellers generally behave, and perceive and react to uncertainty can therefore leave aside important aspects.
From a policy point of view, it is essential to identify the existence of a critical mass of citizens in a particular area which could trigger the use of public transportation over private cars (Dugundji and Gulyas, 2008; Goetzke, 2008), for example. The same applies to the case of vehicle choice, especially with respect to alternative technology vehicles, whose novelty is likely to create obstacles to their diffusion, unless potential buyers can observe a considerable number of individuals using them, refer to acquaintances for suggestions and testimonials (Axsen et al., 2009; Mau et al., 2008; Sartzetakis and Tsigaris, 2005).

We have seen above that the influence may come from the overall society, and therefore based on simple observation or belief, or from a more restricted group of individuals the decision makers have contacts with. These restricted groups are generally referred to as reference groups or social networks. Reference groups in particular are defined in the economic literature as the group of individuals with whom decision makers are likely to interact to a greater extent than others, and are therefore assumed to influence their opinion, behaviour and choices (Hayakawa, 2000). These groups has often been defined in the existing studies in accordance with general information that has rather reflected data limitation than research purposes (Soetevent and Kooreman, 2007), as a proper identification of the individuals likely to have a considerable impact on choices can be complex. In many cases, data and resource limitations have forced researchers to define reference groups using common sample characteristics as a proxy of reference and therefore limit their analysis to anonymous rather than named networks. For this reason, in recent years, economic choice theory, and transportation research, have begun to borrow the sociological concepts and methods of Social Network Analysis (SNA) (Bramouille et al., 2009; Calvo-Armengol, 2009; Carrasco et al., 2008; Carrasco and Miller, 2009; Ioannides, 2006; Sunitiyoso et al., 2011).

3. An introduction to Social Network Analysis (SNA)

Sociological theory defines social networks as the sum of personal networks, which represent the group of persons (alters) with whom a given individual (ego) considers having a link of any nature and has contacts with over a lifespan (Degene and Lebaux, 2005). Social networks have two main components: actors (persons, groups, organisations) who interact within each other, and relationships. The latter can be derived, for example, from control, dependence, competition, and information exchange (Carrasco and Miller, 2009). The main objective of Social Network Analysis (SNA) is to explore these links between people and organisations, their formation and their dynamics (Larsen et al., 2009). The ties forming social networks appear and disappear and have a considerable variability in their intensity over a life time, and choices made by their members in different situations have also an important effect on their structures and dynamics (Bidart and Degenne, 2005; Feld et al., 2007).

In practical terms, in sociological analysis, various survey techniques have been used to identify personal and social networks and assess their structure and dynamics. Among these techniques, the name generator appears to be one of the most popular tools. Other methodologies involve identifying social contacts by using personal sources (like social media contacts, email address books) or institutional sources (like memberships to clubs, mailing lists, etc.). In transport settings in particular, travel diaries have been used to identify social contacts (Axhausen, 2008).
The name generator technique identifies the social network members through in-depth interviewing techniques with questions such as: Who are the people with whom you discuss matters important to you? Who are the people you really enjoy socialising with? Who are the people you have the most contacts with? (Carrasco and Miller, 2009; Marin and Hampton, 2007). Interviewees reveal first a set of alter names and then information about their characteristics in order to assess the nature and magnitude of the relationship (Carrasco et al., 2008).

The identification of members and the consequent assessment of the size of the network and the nature of the relationships are only the first steps in the definition of the structure of a network. A more rigorous analysis requires the definition of the number of isolates (the members who are not connected to everyone but only with particular egos), the density of the network (which is the ratio between the existing number of ties among the members of the network and the maximum possible given the size of it) and the network sub-groupings.

When the purpose of SNA is to identify social activities (i.e. travel) that can be performed by the various members individually or in group it is also necessary to assess the potential activity level between the alters (Carrasco and Miller, 2009). The activities the individuals undertake in both their social and geographical spaces have an important impact on the probability of meeting another individual. Then, the probability of beginning a social interaction depends on the size of the agents’ current networks and their need for information. The agents’ utility depends on the similarities with the other agents and how the interactions with them satisfy their social and information needs. Trust and credibility play an important role as well (Arentze and Timmermans, 2008).

The discussion above shows that studies aiming at mapping social networks and assessing their influences on individual decision makers requires a complex and time consuming data collection effort. Networks are often considerably large and an assessment of their likely influence also need information on trust, credibility, and relative power of their members. Snowball sampling techniques have, for example been used to collect information of social networks (Kowald et al., 2009). Sociograms, graphically showing the connections of individuals in the social space have also been used (Everett, 2006; Hogan et al., 2007; Ioannides, 2006).

In this study, we are interested in social networks when they translate into social influence and therefore affect choice behaviour. In the following section we review the existing literature considering social influence in the particular case of discrete choices in transportation.

4. Social interactions in discrete choice: a brief review of the literature

In discrete choice analysis, to which we address our attention in this paper, the effect of social dimensions was first formalised by Brock and Durlauf for both the binomial (2001) and multinomial cases (2002). Extensions of that model can be found in the more recent economics papers by Krauth (2006), Zanella (2007) and Ioannides and Zabel (2008), for example. In general the basic model considers that an agent utility is formed by a private and social component. The private components normally depends on both the decision-maker’s and his neighbourhood’s characteristics. The social component describes both the strength of social utility and the decision makers’ expectations of the percentage of others in their neighbourhood selecting the same alternative in a choice set (Brock and Durlauf, 2001).
Individual consumers could therefore build expectations on the percentage of users of a particular transport mode (either private car or public transport) in their neighbourhood. This is the approach followed, for example, by Dugundji et al. (Dugundji and Walker, 2005; Dugundji and Gulyas, 2008) in their studies of revealed discrete modal choice data in Amsterdam. These authors explore the interactions between an individual decision maker and the aggregate behaviour of other agents in both the individual’s spatial and social networks, using SNA. Dungdji and Walker (2005) estimate social and spatial interactions in a mixed Generalised Extreme Value (GEV) model structure. They discuss five different empirical strategies to explore social and spatial network interdependencies in individual discrete choices. These strategies consider a feedback effect as for Brock and Durlauf (2001); unobserved group heterogeneity, by considering random error correlation among the members of a specific network as a panel effect (choices of agents in the same group are treated like the choices over different years by the same person in a typical panel data approach); and observed heterogeneity, by using alternative specific variables as well as alternative-specific variance. Dugundji and Gulyas (2008) concentrate instead on the simulation of choice dynamics over time using a multi-agent based model. Both papers (Dugundji and Walker, 2005; Dugundji and Gulyas, 2008) are based on the same dataset and consider spatial and social heterogeneity, as individuals, differently from Brock and Durlauf (2002) are assumed to belong to different reference groups. In details, their network interdependences are taken into account by considering the share of respondents in the same district, socio-economic group, and neighbourhood (based on postcode) selecting the same mode.

In such situations, it is reasonable to assume, when modelling travel choices, that there are unobservable effects likely to have an impact on the choices of different members of a particular reference group. This could be the case for example of the availability and accessibility of a certain transport mode in a particular neighbourhood, which will certainly have an impact of the mode choices of the people living in that neighbourhood. A typical case of endogenous effect will therefore occur (Dugundji and Gulyas, 2008). Endogeneity could in fact arise due to consumers self-selecting their neighbourhood (Zanella, 2007) as well as if multi-directionality of the influence is taken into consideration.

Dugundji and Walker (2005) tackle the endogeneity issue by allowing a constant correlation across alternative within a pre-determined spatial (and social) network, without allowing for this correlation to spread to alternative outside that geographical unit. Their study can therefore be classified among those using spatial filtering to analyse spatial (and social) relationships across decision makers (Wang and Kockelman, 2009). However, while their approach is considered to be effective in analysing social dependence and group-effects among decision makers, some authors have expressed concerned over the capability of similar models to productively analyse spatial dependence (Bhat and Sener, 2009; Smirnov, 2010). In a recent article, Walker et al (2011) address the issue of endogeneity in choice models considering social influence using the Berry, Levinsohn and Parker (BLP) two-stage method. The method, originally used to address similarity among decision makers in a specific market, involves the decomposition of the error in two different parts, in order to identify the portion which causes endogeneity. This enables the analysts to distinguish between the utility relevant to the individual, and that relevant to the peer group. The latter is then substituted by constants capturing the average effect of the peer group (defined both spatially, in accordance to respondents’ post code, and socially, based on income similarities) which are then estimated using linear regression (with the social effect – the fact that each
decision maker choice depend on the choice of others - as explanatory variable). In a linear setting, an instrumental variable approach is then used to correct for endogeneity. Their application on a dataset of mode choice in Amsterdam demonstrates the importance of correcting for endogeneity when analysing social influence in behavioural choice models. Goetzke and Rave (2011) also considers heterogeneity in their analysis of bicycle use in Germany.

When the spatial dimension is the only one which is directly considered (simply neighbouring decision makers), the percentage of other individuals selecting the same alternative (in this specific case public transport over private car) can also be taken into account by inserting in the utility specification a spatially autoregressive choice indicator (Adjemian et al., 2010; Goetzke, 2008). This model is conditional on the observed choice patterns of the neighbouring decision makers and, therefore, the spill-over process is modelled exogenously (Anselin, 1988; Goetzke, 2008). In this way, the potential issue of endogeneity is overcome with relatively simple assumptions. Inspiration from both spatial econometrics, as in the case above, and the theory of externality, has also produced a slightly different approach. This is the case of Paez et al. (Paez and Scott 2007; Paez et al, 2008) in their study of telecommuting practices and residential choices. The authors consider the decision maker’s utility as dependent on whether the same alternative was chosen in the past and the previous actions of the other decision makers the agent can observe. These approaches, where the probability is dependent on variables specific to the decision maker and the choices of others in a reference group, rather than utility, are sometimes classified as variants of the Autologistic model in spatial analysis (Smirnov, 2010). Taking into account the previous decisions by the same agent also formalises learning effects, which have found a strong empirical evidence in transportation research (Thogersen, 2006)\(^1\). Paez et al. (2008) also integrate elements of Social Networks Analysis by assigning a weight to the influence of others, that is assumed to be a proxy of the nature of the relationship between the others and the decision maker. In their application, the weighting variable specifically indicates whether another individual is part of the decision maker ‘significant others’ or not. The model of Paez et al. (2008) also allows for the consideration of potential congestion effects, important in a transportation context, likely to occur if everyone adopts the same practice or decides to use the same mode.

A similar approach to Paez et al (2007; 2008) is used by Kuwano et al. (2007) to analyse social dimensions in choice when looking at car purchase decisions. As noted in the introduction, considering social influence in the analysis of vehicle choice, in particular alternative technology vehicles, is certainly useful in order to identify possible barriers to their diffusion. Kuwano et al (2007) utility framework considers past and future utilities, the influence of habit persistence or variety seeking, and decision makers’ expectations over the choices of members of their reference groups. In particular, the influence from three types of social interactions on vehicle choice is explored: nation, neighbourhood, and homogeneous group (based on a classification of households by income level).

Stated preference experiments are used by Mau et al. (2008), to analyse preferences for new vehicle technology (hybrid and hydrogen fuel cell powered vehicles) in Canada. The authors carry out four different choice experiments in which respondents, prior to the completion of the choice cards, are told the hypothetical market share of the different alternative vehicles

\(^1\) In their 2002 paper, Brock and Durlauf also included the consideration of past society behaviour as an extension to their basic model.
they have to select from. Market share information is employed as a proxy of ‘neighbourhood effect’ and manipulated by the analysts, as a blocking variable in the experimental design, to understand their effect on consumer acceptability. In a similar paper, Axsen et al. (2009) combine revealed and stated preference data to analyse neighbourhood effect in the purchasing of alternative technology in Canada and California. Spatial interdependence at the neighbourhood level is also analysed by Adjemian et al. (2010) in their study of revealed car ownership in the San Francisco Bay. Alternative technology vehicles are also among the transport variables considered by Gaker et al. (2010). Before each choice card, students are given information about the number of their peers (other students participating to the experiments) that have selected the different cars as their favourite one, as well as the number of people who decided not to buy any.

5. An application to the analysis of uncertainty caused by extreme weather conditions

Above, we have discussed the theoretical reasons behind the necessity to extend the traditional model of choice within discrete choice decisions in transportation context. We have briefly reviewed a limited but rapidly growing number of studies that provide empirical evidence on the fact that transportation (as any other consumption or behavioural) decisions, even when taken individually, are influenced by the choices or behaviours of others, to whom the decision maker has a direct or indirect contact.

The theoretical and empirical frameworks briefly described above inspired the development of a study on the way long-distance travellers react to uncertainty caused by extreme weather conditions. A number of opportunities for research were therefore identified, and they are discussed in what follows.

First of all, in recent years in the UK, and elsewhere in the world, a number of natural events have caused considerable disruptions to both short and long distance travel plans. As extreme weather events are forecast to become more frequent as a consequence of climate change, there is a clear need of studies looking at how travellers and transport operators react and adapt to these events. While there are studies on the effect of short-term weather changes on travel demand, especially on commuting patterns (for example Cools et al., 2010; de Palma and Rochat, 1999), no studies, at the best of our knowledge, have done so by considering (long distance) travellers collectively or as members of social networks, and, in particular, their reaction to extreme weather events. As we noted earlier in this paper, referring to one’s social network when facing uncertainty while travelling is an important strategy at the disposal of travellers that is often neglected in the analysis of travel behaviour under uncertain conditions.

From a methodological point of view, as discussed above, few studies have fully integrated social dimensions in discrete choice settings, in particular in stated preference choice experiments. And no studies, to the best of our knowledge, have integrated SNA and these experiments. Although this integration does appear complex, as shown above, this integration

---

2 This research is funded by the FUTURENET (Future Resilient Transport Networks) Project, where a consortium among three universities (Birmingham, Loughborough, Nottingham), TRL, Network Rail and the Highways Agency, examines how to make the UK’s transport systems (with the transport corridor between Glasgow and London as the case study) resilient to climate change from both the engineering and social science points of view. The project is funded by the Engineering and Physical Research Council (EPSRC) as part of the “Adaptation and Resilience to a Changing Climate” programme.
has a strong theoretical basis in the economic theory of choice, as well as associated approaches in psychology and sociology.

**Objectives, methodology and survey development**

The conceptual framework of our study, in line with what was discussed above, considers that an individual decision maker’s social network has an impact first on their propensity to engage in long-distance travel and, importantly, it is a relatively low-cost decision heuristic in uncertain conditions. In particular, we are seeking to provide an answer to the following research questions:

- In what way is long-distance travel affected by extreme weather conditions?
- How do travellers normally react to failures in the provision of the transport service?
- Do travellers (both prior and during travelling) refer to their social network when taking travel decisions in uncertain conditions?
- Can discrete choice models effectively be extended to consider that individuals do not act in isolation but refer to others when making choices?

In order to attempt to answer these research questions and test our hypotheses, it was necessary to conduct a series of primary data collection efforts. We have therefore been developing an internet based survey instrument\(^3\) that contains both stated preference experiments and SNA elements, through two pilot tests (with the second one currently ongoing) and two workshops attended by a number of experts in both travel behaviour and SNA. The first pilot test was carried out in November/December 2010 on a sample of 116 respondents. A second test was carried out in April 2011 on 54 respondents, with the main survey planned for the second half of July 2011. The main survey will be distributed to a sample of 2,000 people in four different neighbourhoods in both Glasgow and London, randomly selected with the following quotas applied: one from each of a North-East-South-West quadrant, a balance of Inner/Outer neighbourhoods, a range of deprivation levels, and no adjoining neighbourhoods. The eight neighbourhoods that have been selected are the Glasgow Wards of North East, Linn, Garscadden/Scotstounhill, and Anderston/City and the London Boroughs of Barking & Dagenham, Barnet, Merton, and The City of Westminster.

Survey respondents provide information on their travelling habits, propensity to risk with respect to extreme weather events, susceptibility to social influence in transport decisions, propensity to use mobile technology to obtain real-time information on travel conditions, previous travel experiences on the London-Glasgow route, and previous experience of transport service failure due to natural events.

In order to assess respondents social networks, a simple name generator was used, as survey participants are asked to provide the list of persons “they have regular contact with, and/or who are the most important to them, and/or those they would want help to discuss personal matters, and/or those they can trust, and/or those they really enjoy socialising with”. For each of the contacts, respondents are then asked to indicate whether the particular person lives

---

\(^3\) There are a number of issues in using internet panels, in particular concerning the low representation of those segments of the population who do not normally use the internet. However, internet panels have a number of other advantages, as they offer an efficient utilisation of time and finance, by providing, for example, a number of additional information about the panellists that would be time consuming to collect.
with them, the type and length of relationship, and the type and frequency of contacts (by various means like face-by-face, phone, SMS, email, chat). Respondents are also asked to indicate which of their contacts they turn to for advice on travel decisions, and, in particular, who (and why) they would contact if they were experiencing an uncertain situation prior or while travelling.

Finally, the survey contains an interactive choice experiment focusing on mode choice under different extreme weather conditions. This experiment not only analyses mode choice but also explores respondents’ need for social interactions, either direct communication or simple imitation, in order to inform their choice patterns in uncertain conditions. We discuss more in details this section of the questionnaire in what follows.

The choice cards

The choice cards that respondents faced during the first test of the questionnaire depicted hypothetical travel situations first described in terms of the purpose of the trip, its importance, the persons (if any) travelling with you, and the weather conditions on the day of travelling (randomly chosen between ‘general fine weather’ and a number of extreme weather alert). Subsequently, four different methods of travel were compared based on five attributes. An example choice card, and preliminary questions, is illustrated below.
Table 1. Example of a choice card used in this study

The reason for you to undertake this trip to Glasgow is:

1. Visiting friends and relatives

Also imagine that the importance score of the trip (from 0 “not very important, I could have easily postponed the trip to another moment” to 10 “the trip was extremely important and there was no way I could postpone it to another moment) is

Five

Also imagine you are travelling:

1. On your own

And please imagine that the average weather over the trip on the day of travelling is:

4. Weather Alert3: Heavy snow and icy roads

<table>
<thead>
<tr>
<th>Method of travel</th>
<th>Air (1hr and 10 minutes)</th>
<th>Train (4 hours)</th>
<th>Car (5 hours and 10 minutes)</th>
<th>Coach (9 hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of service</td>
<td>10 per day</td>
<td>12 per day</td>
<td>Not applicable</td>
<td>8 per day</td>
</tr>
<tr>
<td>Cost – single ticket Air/Train/Coach or fuel</td>
<td>£100</td>
<td>£150</td>
<td>£125</td>
<td>£55</td>
</tr>
<tr>
<td>Motorway toll costs and city congestion charges</td>
<td>n/a</td>
<td>n/a</td>
<td>£25</td>
<td>n/a</td>
</tr>
<tr>
<td>Likely travel time delay</td>
<td>About 15 minutes</td>
<td>About 15 minutes</td>
<td>About 30 minutes</td>
<td>About 45 minutes</td>
</tr>
</tbody>
</table>

For travelling to Glasgow under ‘Heavy snow and icy roads’, please select your preferred method, if any, from the four shown above.

TICK ONE BOX ONLY

<table>
<thead>
<tr>
<th>Please select your preferred method</th>
<th>Air</th>
<th>Train</th>
<th>Car</th>
<th>Coach</th>
<th>Would not travel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

A number of practical and methodological options to integrate social dimensions in the analysis of individual choices were identified during the review of the existing literature and the development of this survey. They are discussed in what follows.

It is possible to carry out a traditional stated preference experiment and then assess the degree of social influence in the econometric analysis phase. This is normally achieved by detecting similarities between respondents in the same social or spatial context (similarly to Dugundji and Walker, 2005; Goetzke, 2008 for example) using various econometric techniques. Unnamed or anonymous social or spatial networks are therefore used in this case as respondents are compared with others similar to them based on a number of proxy social (generally income, working status) and/or spatial variables (residence in the same postcode area). Additional information about respondents’ ‘social exposure’, ‘likelihood to be influenced by others in their choices’ or ‘susceptibility to be influenced by opinion leaders’,

---

4 Please note this table does not exactly reproduce what was shown to respondents, as they had an online version of the survey.
for example, could also be collected in the questionnaire in order to further profile respondents.

Alternatively, choice cards can be modified for purpose. Depending on whether the interest of the analyst is on named or unnamed networks, a number of options seem possible. An additional attribute could be added to the stated choice exercise. This attribute could describe different situations depending on who has chosen the same alternatives. This could refer to a particular person in the respondents’ network (and in this case it would be necessary to properly identify the members of the network) or a generic indication of the number of people in the respondents network selecting a particular option. The information reported in the choice card could be either hypothetical or real. In this case it would therefore be necessary to interview the members of the respondents’ networks as well in order to ascertain their preferences. This could be done using a snowball sampling technique (as, for example, by Kowald et al., 2009).

If the interest was rather on unnamed networks, and generally on a larger source of social dimension than the respondents’ closest contacts, information about percentage of people in the respondents’ neighbourhood, region or country could be appended to choice cards. This is the approach followed by Mau et al (2008), where ‘market share’ of different types of vehicles are used as a blocking variable in stated preference experiments (although not directly as an attribute). In the case of Gaker et al (2010), an indication of what others, in this case the more restricted group of ‘other students in the lab carrying out the survey’ have done is given as a simple preliminary information to respondents.

The same information about the behaviour of others within the respondents’ social network could be included in the survey after the completion of the choice cards. This could explore the potential for “a change of mind” generated by the understanding of what other people would do in the same situation. This could be based on real information (the specific member of the network is also interviewed) or hypothetical (“what would you do if that particular person would choose differently in a similar situation”).

All of the options discussed above present both challenges and opportunities. First of all, most challenges arise due to the stated preference nature of our exercise. When analysing revealed preference data, it seems logical to assume that consumers have observed the behaviour of others, either close to them or the society; the same does not easily apply to the case of stated preference. This is even more the case when respondents face unfamiliar or very hypothetical situations. In this case it is reasonable to assume respondents may be influenced by what other ‘would do’ in similar situation, in accordance to belief for the members of their networks, or more general social rules for the larger society.

The options showed above are not necessarily mutually exclusive and could be integrated. Decisions over which method to apply in our study were not only based on methodological grounds, but also taking into consideration the available financial resources for conducting this survey. For example, in our view, the ideal option would have been to carry a full SNA, identifying each respondent’s social networks and then distributing the questionnaire to the relevant member (using a snowball sampling technique). This option, however, would have required a considerable effort in terms of time and financial resources. In addition, it would have been necessary to identify a number of respondents that would have faced the questionnaire without knowing what the other members of the network would do (the seeds of the network). Last but not least, there are also ethical considerations to be taken into
account when asking respondents to indicate other potential sample members, and supplying full names and addresses. Although these issues are particularly relevant in Social Network Analysis types of studies that handle more sensitive information, like health issues in order to analyse the possible pattern of diffusion of an illness or drug use (see for example Vervaeke et al., 2007), they had to be taken into consideration nonetheless.

The following questions were appended to the choice cards in test 1 for some of the respondents:

Table 2. Additional questions – version 1

<table>
<thead>
<tr>
<th>Imagine if Person X of the group of persons you normally ask for help in making travel decisions has to face the same choice as you. Do you think his/her choice would be the same as yours?</th>
</tr>
</thead>
<tbody>
<tr>
<td>TICK ONE BOX ONLY</td>
</tr>
<tr>
<td>1. Yes</td>
</tr>
<tr>
<td>2. No</td>
</tr>
<tr>
<td>3. I do not know</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How confident are you in knowing what they would choose?</th>
</tr>
</thead>
<tbody>
<tr>
<td>TICK ONE BOX ONLY</td>
</tr>
<tr>
<td>1. Very unconfident</td>
</tr>
<tr>
<td>2. Unconfident</td>
</tr>
<tr>
<td>3. Neither confident nor unconfident</td>
</tr>
<tr>
<td>4. Confident</td>
</tr>
<tr>
<td>5. Very confident</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Which alternative do you think he/she would have chosen given the same conditions and information?</th>
</tr>
</thead>
<tbody>
<tr>
<td>TICK ONE BOX ONLY</td>
</tr>
<tr>
<td>1. Air</td>
</tr>
<tr>
<td>2. Train</td>
</tr>
<tr>
<td>3. Car</td>
</tr>
<tr>
<td>4. Coach</td>
</tr>
<tr>
<td>5. Not travelling</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Would you now change your mind because of their opinion, and choose the same option as them?</th>
</tr>
</thead>
<tbody>
<tr>
<td>TICK ONE BOX ONLY</td>
</tr>
<tr>
<td>1. Yes</td>
</tr>
<tr>
<td>2. No</td>
</tr>
<tr>
<td>3. I do not know</td>
</tr>
</tbody>
</table>

In the question at the top of the table, ‘Person X’ was randomly selected by the survey software among the persons the respondents had indicated previously as being their closest contacts. Unfortunately, there was a mistake in the instructions and the person was drawn from a smaller set that normally included only one or two persons for each respondent (the closest ones). For this reason, most respondents reported that the particular person would have done exactly the same as them (in a sense giving evidence of ‘homophily’ or similarity between the members of a social network) (McPhersons et al., 2001), and that they were either confident or very confident in reporting the preferences of the given members of their networks. In only a few cases did respondents state their intention to change their mind.

---

5 This version should have been given to half of the respondents but a problem in the questionnaire software meant that version 1 was distributed to 25 respondents while version 2 to 88 of them.
having reflected upon what a particular member of their social networks would have done in a similar situation.

The first question in Table 1 bears some similarities with the Interactive Agency Choice Experiments (IACE) method, first developed by Brewer and Hensher (2000) to analyse collective choices and negotiations made by supervisors and employees over the adoption of telecommuting work practices. The method was subsequently extended to the analysis of sender/deliverer collective choice and negotiations (over cost, route, pick-up and delivery time) in freight transport settings (Hensher and Puckett, 2008; Hensher et al., 2007; Puckett et al., 2007) and, more recently, to the study of household members’ interactions in automobile purchase decisions (Hensher et al., 2008). IACE integrates elements of both discrete choice and game theory, and considers two or more decision-makers that engage in sequential choices in order to reach an agreement after a negotiation and bargaining process. Respondents are asked to select not only their favourite option, but also the one that they think their counterpart (in the agreement process) would select.

The following questions were appended to the choice cards for the remaining respondents.

Table 3. Additional questions – version 2.

<table>
<thead>
<tr>
<th>In a real life situation, would you have needed some additional help in taking the decision over the different travel situations?</th>
</tr>
</thead>
<tbody>
<tr>
<td>TICK ALL THAT APPLY</td>
</tr>
<tr>
<td>1. Yes – I would have needed to access additional information (e.g. transport operators, MET office)</td>
</tr>
<tr>
<td>2. Yes – I would have needed to ask my family/friends for help in deciding</td>
</tr>
<tr>
<td>3. No</td>
</tr>
</tbody>
</table>

Who would you have asked in particular?

TICK ONE BOX ONLY [here the options were the contacts respondents indicated previously]

Would you have particularly needed help in deciding...? |

TICK ALL THAT APPLY |
| 1. Whether to go or not |
| 2. Method of travel (e.g. by car, train, bus, air) |
| 3. The departure time |
| 5. The route |
| 6. Whether the weather conditions would be seriously disruptive or not |
| 7. Whether the cost would be appropriate for such a trip |
| 8. Other (PLEASE SPECIFY) |

Questions reported in Table 3 had two main purposes. The first was to assess whether, rather than referring to their social contacts, respondents would have needed additional information from transport operators to perform the choices they were asked to carry out. Secondly, we wanted to understand whether respondents’ need for additional information, and suggestions, had a role in other moments of the decision process than the final choice of the mode (or not to leave). For example, we assumed that respondents may have needed assistance in building their preference over certain attributes of the alternatives only, or to decide between two most favourite alternatives rather than among the four. We wanted, therefore, to explore whether a semi-compensatory model of choice (Martinez et al., 2009) would have been a better picture to analyse decisions in uncertain situations.
Initial findings from the first test survey

Here we report some simple descriptive statistics about the first test sample and some of the questions discussed above.

There were 116 respondents in the first test survey (56% in London, 44% in Glasgow). About 50% consider that their social life does depend on their capability to travel, and 50% consider that their way of travelling is not influenced by the people they know. In the uncertainty section, there were 79 affected trips (half affected by snow, about 10 by volcanic ash). Each respondent could name up to 10 social network contacts; the average was 4.9 contacts. In 40% of cases, the first contact is a partner or spouse.

Person 1 is the one most people (55%) would turn to for general travel advice, the same applies to the case of uncertainty while travelling. The reasons to contact a particular person generally were: experience either general or a specific route, organization skills, capability of providing emotional support, trust, good at finding info on the internet, people who would need to know we are late, simply the person I turn to for support in any situation (etc).

For the questions that were appended to the choice card (version 1 to 25, version 2 to 88) we obtained the following information:

- 52% of people said they would have needed to access additional information from the transport operators to better perform their choices (48% would have not needed)
- 25% of people said they would have needed to ask family and friends for help in deciding (75% would have not needed); 45% of these would have contacted Person 1
- Help would have been especially sought for deciding the method of travel
- As already noted above, the vast majority of respondents thought specific members of their network would have selected the same option. Most respondents were either confident or very confident in reporting their contacts’ preferences

Changes for the second test

As seen above, the questions which explored social dimension in individual choice did not perform effectively during the first test of the questionnaire. A number of changes were therefore applied, and they are discussed below.

The routing in the first question of version 1 was modified in order to increase the range of people that respondents could imagine in their place for a similar choice situation. The number of decision issues (in version 2) about which respondents could have needed help was increased. This was undertaken as it appeared that the help from members of the social networks could have been invoked in different stages of the decision process and perhaps not the main one (in a semi-compensatory choice framework, where it is reasonable to assume that referring to choice heuristics is done to process single attributes, rather than the totality of them, or single nodes within the process of choice)

An additional question reporting hypothetical share of choice in the respondents’ neighbourhood was inserted, using a variant of the ‘market share’ approach discussed above. This was done to further explore social influence and add a reference to a more generic unnamed network to the named network mainly used in the previous questions. Hypothetical
percentages for each of the choice options, including the option of not travelling, were inserted. Respondents were then asked whether seeing this information would have triggered an opinion change. Respondents were also given the possibility to explain their concerns over the percentages as not representative at all of their perceived neighbourhood general preferences. This question is illustrated in Table 4 below.

Table 4. Market share questions

| Please now consider that in your neighbourhood in London/Glasgow the following choices were recorded when facing the same choice card: |  |
|---|---|---|---|---|---|
| Air | Train | Car | Coach | No travel | Choice card n. |
| 25 | 15 | 25 | 10 | 25 | 5 |

Would your choice remain the same?

TICK ONE ONLY

Yes, I would still choose [option chosen in Q58 here]

No, I would choose [first option not chosen]

No, I would choose [second option not chosen]

No, I would choose [third option not chosen]

No, I would choose [fourth option not chosen]

I do not know

I do not believe these percentages are realistic of what people in my neighbourhood would do in a similar situation.

6. Discussions and conclusions

This paper reports work in progress and for this reason lacks more detailed and robust empirical results, as a more complete analysis has not yet been performed due to time constraints. Nevertheless, a number of theoretical, practical and empirical challenges and opportunities have arisen from the review of the theory and the available empirical literature, as well as the discussion over the different issues encountered during the preparation of our questionnaire. We discuss these issues in the remainder of this section.

The first important theoretical issue lays in the fact that we are asking respondents’ to report other people preferences. This seems to be an important alteration to the traditional model of choice, and although this has already been undertaken when analysts have employed the unitary model of household choice (see Vermeulen, 2002 for a review of household models in economics), when the preference of one only member of the household is deemed sufficient to represent the other members preferences in terms of consumption and labour supply, it is certainly a challenging issue both from the theoretical and empirical point of view. As discussed previously, the best approach would have been to interview the members’ of the respondents’ social network in order to provide them with real information over their peers preference facing the same choice. This was however not possible due to financial constraints.

In addition, an important question arises: can we trust respondents when they report the preference of members of their networks? The fact that these others are being indicated as part of their social networks should be enough for that. Although there are cases in the
literature where consumers did not seem to be able to report the right preferences of other consumers, even when they were very close to them (like between husband and wives) (see for example Beck et al., 2009; Beharry-Borg et al., 2009). Last but not least, in our approach, social dimensions are analysed post choice exercise, by assessing a possible change of mind by respondents, when information about choice patterns in their neighbourhood are revealed to them, or making them to reflect about the choices of the members of their network. There are, indeed, issues associated with this approach. For example, it is possible (and it is also part of our hypotheses) that respondents have already conformed to their peers or to a more general society when making their choice. In any case, it will possible to assess this, by identifying the presence of social and spatial correlation across choices.

It is anticipated that our econometric approach will develop as follows. Mode choice data will be first analysed using the econometric instruments typical of stated preference experiments. Social dimensions will then be explored using the methods used in the literature detailed above, for example by Dugundji et al (2005; 2008) who consider both spatial and social elements, and Goetzke (2008) who instead mainly considered spatial elements. We also believe that there may be a potential in employing Latent Class in order to explore social dimensions, especially if it is considered that the usual causality (I choose therefore I belong) could be reversed (I belong therefore I choose) in a manner similar to the application of Morey et al (2006).

However, the approaches mentioned above have been mainly used in a revealed preference setting. There may be therefore issues in applying them in a stated preference framework. Importantly, the analysis approach discussed above does not consider the post-choice cards questions in the analysis. These questions could be analysed independently, or could be part of a multiple discrete choices approach. At the moment we are considering the issues involved in this.

To conclude, the main econometric issue will certainly be the treatment of endogeneity. We have briefly seen above the way this issue has been treated in the literature. While endogeneity is well discussed in the economic literature (Soetevent, 2006 discusses the issue when analysis continuous choices mainly), it seems to us this issues is less discussed in discrete choice applications, and the recent paper by Walker et al (2011) in particular, it is certainly a welcome addition. Whether the approaches used in the literature so far will be applicable to our dataset remains, however, an unanswered question at this stage.

**Acknowledgements**

The authors wish to thanks the participants to the Social Network Analysis and Transport Workshops, held in Nottingham, 2-3 March 2010 and Manchester 11-12 January 2011, for the useful discussions that have inspired this paper. We also thank the Engineering and Physical Sciences Research Council (EPSRC) and the Economic and Social Research Council (ESRC) for their financial support of the ‘FUTURENET’ project.

---

Unfortunately a number of difficulties were encountered during the pilot tests of the questionnaire in programming the appropriate routines for the questions exploring social dimensions, and this has caused delay to the analysis of test data. We are currently carrying out simple tests to assess whether the questions did perform satisfactorily.
Reference list


