

Network Effects in Public Transit Use: Evidence from a Spatially Autoregressive Mode Choice Model for New York

Frank Goetzke

[Paper first received, December 2005; in final form, March 2007]

Abstract

This article empirically tests for positive network effects in transit use by applying a spatial autoregressive logit mode choice model with 1997/98 work trip data from New York City. Positive network effects exist when people prefer to use transit together with other people as a result of social spill-over. Although these network preferences should differ for each person, because of statistical restrictions in the model, individual network preferences cannot be obtained. However, it is possible to derive econometrically a measure of aggregate network preference. This paper can be seen in the wider context of other recent work focusing on the analysis of spatial dependencies and network effects in travel.

Introduction

In this paper, I empirically test for positive network effects in transit use by applying a spatial autoregressive logit mode choice model with 1997/98 work trip data from New York City. Positive network effects exist when people prefer to use transit *together* with other people as a result of social spill-over. Although these network preferences should differ for each person, because of statistical restrictions in the model, I cannot assess individual network preferences. However, I will be able to derive econometrically a measure of their *aggregate* network preference.

There is increasing interest in analysing spatial dependencies and network effects in travel behaviour. LeSage and Polasek (2005) examine commodity flow matrices by extending a gravity model, a tool widely used in the field of transport, to include spatial autoregression. Using a Monte Carlo simulation, Páez and Scott (2005, 2007) investigate the impact of social networks in discrete choice models.

It is not entirely new to include network effects within empirical choice models. Brock and Durlauf (2001, 2002) have been researching methods to account for social interaction and neighbourhood effects in

Frank Goetzke is in the School of Urban and Public Affairs, University of Louisville, 426 W. Bloom Street, Louisville, Kentucky 40208, USA. Fax: +1 502 852 4558. E-mail: f0goet01@louisville.edu.

both binary and multinomial choice models. However, they do not explicitly account for the spatial dimension in their discussion of the problem. On the other hand, very few spatial discrete choice models have been estimated because of their computational complexity, but none of them deals with social interaction and spatial spill-over effects in disaggregate transport behaviour (Anselin, 1988; Fleming, 2004).

While there exists a spatially autoregressive mode choice model using aggregate data (Boldoc *et al.*, 1995), a spatially autoregressive disaggregate mode choice model based on real-world, empirical data to examine network effects has never been implemented.

In the next section, I will lay out the theory of how econometrically to model network effects, followed by an overview of the data used and the issues surrounding the design of the weight matrix. Then, I develop the actual econometric models and evaluate the results before concluding with a discussion of the consequences for policy-makers.

Theory of Modelling Network Effects

In Goetzke (2006), the theory of network effects is introduced as it applies to transport mode choice decision-making. Interpreting network effects as a signal that a transport mode is safe and reliable, the main claim of the paper is that the utility of using a certain mode of travel increases with its mode share.¹ Therefore, the more people who use the mode, the more attractive this transport mode becomes for all other people. These social spill-over effects lead to positive demand-side network externalities, the exact opposite from what is known as congestion.²

However, Goetzke did not provide any empirical evidence that these network effects exist. In this paper, I develop a rigorous econometric framework to examine real-world data and quantify these network effects. The

starting-point for the analysis is the random utility conditional mode choice model, the traditionally used workhorse of transport demand modelling and forecasting (McFadden, 1974; Ben-Akiva and Lerman, 1985; Train, 2003). Because of the structure of the spatial weight matrix used in the model, I will restrict the analysis in this paper to a binary mode choice model. However, the general approach could be extended to a multinomial mode choice model, as discussed in the conclusion.

Let us consider the utility v_{ni} of N individuals with $n = 1, \dots, N$ using mode i , which is a function of A personal characteristics, x_{an} , with $a = 1, \dots, A$, and the B mode-specific characteristics, t_{bi} , with $b = 1, \dots, B$. Assuming that v_{ni} is a linear combination of x_{an} , t_{bi} , as well as a random error term ϵ_{ni} (stemming from unobserved utility determinants), an empirical utility model could be expressed as follows

$$\begin{aligned}
 v_{ni} &= \alpha_1 x_{1n} + \dots + \alpha_a x_{an} + \beta_1 t_{1i} \\
 &+ \dots + \beta_b t_{bi} + \epsilon_{ni} \\
 &= \mathbf{x}\mathbf{a} + \mathbf{t}_i\mathbf{b} + \epsilon_i
 \end{aligned}
 \tag{1}$$

where, \mathbf{x} is a $N \times A$ matrix of personal characteristics, \mathbf{t}_i is a $N \times B$ matrix of mode characteristics for mode i , and \mathbf{a} and \mathbf{b} are the vectors of corresponding regression parameters.

Data can only be collected on actual transport mode choice decisions, not on mode-specific utility levels. Therefore, equation (1) has to be transformed in such a way that v_{ni} becomes an unobserved, latent variable. Assuming that ϵ_{ni} is logistically distributed, a binary logit mode choice model can be derived, with $P_n(i)$ being the probability of person n to choose mode i over mode j

$$\begin{aligned}
 P_n(i) &= P(v_{ni} \geq v_{nj}) \\
 &= (\exp v_{ni}) / (\exp v_{ni} + \exp v_{nj}) \tag{2} \\
 &= 1 / [1 + \exp-(v_{ni} - v_{nj})]
 \end{aligned}$$

If the above random utility model is extended to account for network effects, it is necessary to add a spatially autoregressive mode choice term to Equation (1). Therefore, the model includes a spatial weight matrix, \mathbf{W} , of the dimension $N \times N$, as well as an $N \times 1$ vector \mathbf{m}_i of revealed mode choice decisions by people. Both terms together, $(\mathbf{W}\mathbf{m}_i)$, result in the spatially weighted average mode share of mode i for all the people neighbouring individual n . The scalar ρ is the regression coefficient for the spatial lag term, so that the full matrix form looks as follows.

$$\mathbf{v}_i = \mathbf{x}\mathbf{a} + \mathbf{t}_i\mathbf{b} + \mathbf{W}\mathbf{m}_i\rho + \mathbf{e}_i \quad (3)$$

The spatially autoregressive structure becomes more visible when considering that \mathbf{m}_i is actually a function of the mode-specific utility vector \mathbf{v}_i , $f(\mathbf{v}_i)$. By including $f(\mathbf{v}_i)$ in equation (3), the issue of spatial autoregression can be easily seen

$$\mathbf{v}_i = \mathbf{x}\mathbf{a} + \mathbf{t}_i\mathbf{b} + \mathbf{W}f(\mathbf{v}_i)\rho + \mathbf{e}_i \quad (4)$$

Assuming that the error term does not exhibit any spatial autocorrelation, equation (4) now represents what Anselin (2002) calls a conditional spatially autoregressive discrete choice model, where the spatial lag term is assumed to be exogenous. The model is conditional upon the *observed* neighbouring mode choices, which means that the spatial spill-over process is not modelled as an endogenous process. The advantage is that the estimation of this model type is straightforward to estimate.

The $1 \times N$ vector of probabilities P to use mode i becomes a function of: the $N \times A$ individual characteristics matrix \mathbf{x} ; the differences between the two $N \times B$ mode-specific characteristics matrices $\mathbf{t}_{i \neq 1}$ and $\mathbf{t}_{i=1}$, which are related to both available modes; the average mode share of all surrounding neighbours $(\mathbf{W}\mathbf{m}_i)$, representing network

effects; and all the corresponding vectors of regression coefficients \mathbf{a} , \mathbf{b} , as well as the regression coefficient scalar ρ . Because the spatial weight matrix refers to just one mode, in its basic form this model can only be applied to binary choice models. I will discuss in the conclusion an extension towards multinomial choice models.

Equation (4) essentially says that, with increasing neighbouring mode share in mode i , the mode-specific utility of a person to also choose this transport mode increases as well. Therefore, a positive value of ρ , which is significantly different from zero, would indicate the existence of network effects. Now, if the utility of using the mode is an indicator of the willingness to pay (WTP) and is assumed to be correlated with the WTP for the transport mode, the regression coefficient ρ can be interpreted as the value of the slope for the upward-sloping WTP curve with respect to mode share, due to network effects (see Figure 1). However, in reality, different people have individual levels of ρ that cannot be modelled as such, because the number of variables, including all the personal network preferences, would surpass the available degrees of freedom. Therefore, the network effects estimate ρ refers to the aggregate network effects.

Dataset and Weight Matrix

The data used for the spatially autoregressive logit mode choice model are based on the 1997/98 comprehensive regional household travel diary survey conducted for the best practice travel demand forecasting model by the New York Metropolitan Transport Council (NYMTC, 2004).³

The data were collected for the whole metropolitan area of 28 counties, which include parts of upstate New York, New Jersey and Connecticut, but I only took data from the five New York City boroughs of the Bronx,

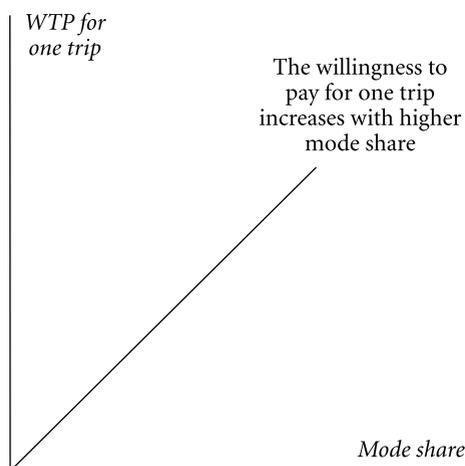


Figure 1. Positive network effects for an individual person in their mode choice decision.

Brooklyn, Manhattan, Queens and Staten Island. The main reasons for choosing these boroughs are

- The New York City area has a relatively high density of surveyed households, allowing a meaningful weight matrix to be developed, which is necessary for estimating the autoregressive model. Outside the NYC area, the distances between surveyed neighbours become rather large.
- Only in New York City is the number of transit riders high enough to get the variability in the dataset desired for econometric analysis. Also, NYC has an integrated transit system. Therefore, individuals there face roughly the same transit infrastructure.

For the mode choice model, only trips from home to work were used. The literature in classical transport modelling distinguishes between the three typical trip purposes: home-based work trips (HBW), non-home-based trips (NHB) and home-based other trips (HBO). Each trip purpose is modelled differently, since the mode choice determinants are expected to differ. Purely for practical reasons, I decided to restrict my analysis only to HBW

trips. I also included only transit and drive-alone trips (more than 90 per cent of all HBW trips) between different travel analysis zones (TAZ) within New York City which were longer than 1.5 miles or 2.4 km (to exclude the possibility of walking). In the end, I had 1652 person trips from all five boroughs of New York City, of which 32.7 per cent drove alone to work and the remaining 67.3 per cent used transit.

Using US census block group longitude and latitude information, Figure 2 displays the location of the surveyed households used to create the spatial weight matrix. As can be seen, the density of households is much greater in Manhattan compared with the remaining four boroughs. This can be attributed to the fact that the sample is not spatially random, since the sampling was not done with a spatially autoregressive model in mind.

Accounting for the heterogeneous household location density, a row-standardised spatial weight matrix of the dimension $N \times N$ was developed as follows. Each row contains the 40 closest neighbours of the corresponding observation n which are equally weighted. Adding more than 40 neighbours would have not given much more information, since the average mode share would not change significantly. As soon as a neighbour lives further away than 0.75 miles (1.2 km), however, he or she was not included, unless the number of neighbours would drop below 4. The weight matrix needs a minimum number of neighbours in order to be meaningful. If a neighbour is further away than 0.75 (1.2 km), or a 20-minute walk, the neighbourhood characteristics may change too much to reflect network effects. The sum of the weights in each row adds up to one, so that the spatial weight matrix effectively calculates the spatially moving transit mode split average.

This method of deriving a spatially moving transit mode split average represents a unifying decision rule, while it also allows account-

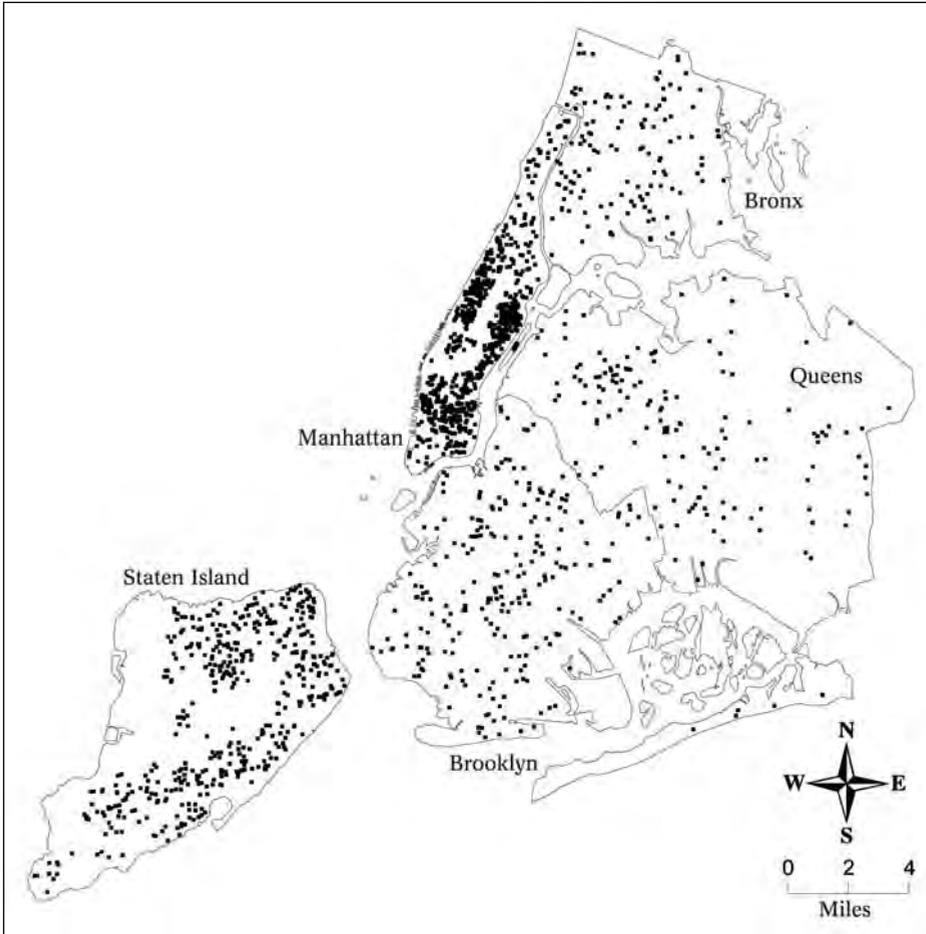


Figure 2. Geo-coded locations of all surveyed households in New York City.
Source: NYMTC (2004).

ing for the heterogeneity of the boroughs in terms of spatial household density. It turns out that the mean number of neighbours is 23, with a maximum number of 40 neighbours existing for about a quarter of the records. The mean distance of neighbours is 0.4 miles (0.64 km) and very few of them, a little more than 0.5 per cent, are further away than 0.75 miles (1.2 km).

Model and Results

The objective for developing the following models is to find out whether or not network

effects exist for transit use. The basic set-up is a McFadden-type mode choice model (McFadden, 1974) that includes the commonly used determinants of mode choice decision-making: total travel time, access to a car, income and gender. Age was not considered because of the large number of missing observations. Mainly to control for parking cost in Manhattan, a dummy was added as a proxy for all Manhattan destinations. Finally, the network effects for transit use were measured as the spatially autoregressive term of revealed mode choice decisions previously described, which essentially is equivalent to

the spatially weighted average transit mode share of all neighbours. The model, however, accounts only for *aggregate* network effects. The expected individual network effects cannot be captured econometrically, due to the restricted number of degrees of freedom (see earlier discussion).

Three different models were evaluated. The first model, called the basic model, omits network effects and represents the traditional approach of mode choice modelling. In the following two models, network effects are included. The model with network effects incorporates a variable for the aggregate network effects. In the model with network effects controlling for car access, which was run to do sensitivity analysis, the network effects were split into an aggregate component

and a component for people not having access to a car.

The results of all three models are exhibited in Table 1. Standard errors are shown in parentheses and statistical significance is indicated by asterisks. All estimated regression coefficients of the basic model have the expected signs. Except for the mode-specific constant and the high-income-group variable, they are also significantly different from zero at the 1 per cent level. While it is not a problem that the estimate of the regression coefficient for the high-income-group variable is not significantly different from zero, it may potentially be a problem for the mode-specific constant term, since it is expected that people prefer the automobile over transit.

Table 1. Estimation results for the New York City mode choice model

	<i>Basic model</i>	<i>Model with network effects</i>	<i>Model with network effects controlling for car access</i>
Transit mode constant (drive alone = 0, transit = 1)	-0.0788 (0.194)	-1.047*** (0.248)	-1.001*** (0.250)
Total travel time (0, 1)	-0.036*** (0.006)	-0.035*** (0.006)	-0.0356*** (0.006)
Average network effect for transit use (1)		1.887*** (0.293)	1.813*** (0.296)
Additional transit network effect for no car access (1)			2.174 (1.591)
No car access dummy (1)	4.258*** (0.398)	3.893*** (0.404)	2.463** (1.041)
Destination Manhattan (1)	1.991*** (0.201)	1.831*** (0.207)	1.819*** (0.207)
High-income dummy (1)	-0.242 (0.157)	-0.311 (0.161)	-0.309 (0.161)
Male dummy (1)	-0.702*** (0.159)	-0.735*** (0.163)	-0.730*** (0.163)
Sample size	1652	1652	1652
Log-likelihood	-534.03	-512.75	-511.80
Pseudo- R^2	0.528	0.546	0.546

significant at the 5 per cent level; *significant at the 1 per cent level.

The regression coefficient of the total travel time variable is, as expected, negative, because the indirect utility derived from using any transport mode decreases with more time spent on it. In this context, it should be mentioned that the total travel time regression coefficient of the McFadden-type mode choice model refers to the difference in mode-specific travel time, which in the case of transit is the increase (or decrease) in the time it takes to ride transit in comparison with driving a car.

Obviously, if a person does not have access to a car, he or she is more likely to use transit. This explains the positive sign for the no-access-to-a-car dummy coefficient estimate. Since parking is costly in Manhattan, people are more inclined to take transit there, which makes the regression coefficient positive. The automobile is a normal good; therefore, the high-income-group will avoid transit, resulting in a negative coefficient estimate. Finally, men are less willing to ride transit, leading to a negative regression coefficient.

The model with network effects improves upon the mode-specific constant (in comparison with the basic model), since the regression coefficient estimate, while keeping its negative sign, now becomes significantly different from zero. In addition, the new model reveals positive aggregate network effects for transit use, while all the other coefficient values remain very robust.⁴

It is interesting to interpret the constant term together with the network effects. With very low transit mode share, the indirect utility of using transit is smaller than for driving alone (see negative coefficient value for the constant). As soon as the transit mode share moves towards one, the difference of indirect utility between transit use and drive-alone trips decreases until transit use utility eventually surpasses the indirect utility of driving alone, since the network effects term is larger than the coefficient for the constant. This is exactly what is expected from

network effects. The willingness to pay for transit, which is related to the indirect utility, increases when more people take transit, and thus lowers the relative cost of using transit (vs the car).

In the model with network effects, total travel time is the major variable accounting for the supply side. For the transit mode, total travel time includes in-vehicle travel time, access and egress time, as well as wait time (which is a function of service frequency). For the drive-alone mode, total travel time is based on the automobile travel speed. Therefore, total travel time is an aggregate variable of the physical infrastructure, accounting for transit service frequency and highway congestion.

However, what if the network effects variable captures a portion of the supply side? It can be seen in Table 1 that the coefficient estimates for total travel time are essentially the same in the basic model and the model with network effects. In the case of network effects incorporating the infrastructure, it is expected that the coefficient values would be not as similar as they are. It could also be argued that the network effects variable accounts for some unmeasured supply level. For example, a person not having access to a car may live closer to a transit station and may be more likely to be surrounded by people also without car access. Thus, better transit availability could easily explain network effects.

The model with network effects controlling for car access addresses this issue by splitting the network effects into two components: the aggregate network effects and the network effects for people without access to a car. The most important finding is that the inclusion of the second network effect component does not greatly change the value of the regression coefficient for aggregate network effects. Furthermore, the estimate of the network effects regression coefficient for people without access to a car is not significantly different from zero. If the network effects variable for people with no access to a car

is analysed in conjunction with the no-access-to-a-car dummy, it can be seen that the network effects coefficient absorbs part of the dummy coefficient. The dummy coefficient estimate becomes less significant and decreases at about the same value of the additional network effects regression coefficient for people without access to a car, multiplied by the average transit mode share. All other coefficient estimates remain stable in value and significance. The attempt of the last model to control for potentially unobserved supply-side effects that stem from the physical infrastructure did not yield any improvements. The original model with network effects can be viewed as robust.

It still may be possible that the network effects variable captures some effects of the transit amenities such as transit safety, crime rates, station cleanliness or infrastructure quality. Furthermore, using transit in high numbers may even generate these amenities, when, for example, transit riders create a less crime-ridden environment, or demand a cleaner transit station. This is exactly what Goetzke (2006) discussed as the possible source of network effects, when he describes network effects as a signal for safe and reliable transit. Therefore, unlike the typical mode choice model, all these formerly omitted transit characteristics can now conveniently be seen as summarised in the network effects variable. In this context, the network effects variable gives way to a nice interpretation: it essentially becomes an indication of people's perception of how well transit works.

It can be concluded that, in the presence of network effects, individual mode choice decision-making does not only depend on personal traits and mode-specific characteristics, but also on the mode share of the person's neighbourhood. This means, everything else being equal, and counter to conventional wisdom, that poor people are less likely to use transit in areas with low transit share than in areas with high transit share. On the

other hand, more people who are wealthy take transit in areas with high transit share, when compared to areas with low transit share.

Policy Consequences

The fact that transit use exhibits network effects may lead to far-reaching policy consequences. It is typical for both transport planners as well as policy decision-makers to depend on travel demand forecasting models for evaluating new transit projects. The centrepiece of every travel demand forecasting model is the mode choice model, which traditionally does not include a network effects variable to account for spatial autoregression. This non-inclusion of network effects in the model formulation causes an omitted variable bias which becomes visible in the mode-specific constant term, as shown in Goetzke (2003). The result is, therefore, a systematic forecasting error where transit ridership in suburbs with low transit mode share is overestimated and transit ridership in the central city with high transit mode share is underestimated.

In the traditional mode choice model approach, the mode-specific constant term for transit use captures the difference in indirect utility between transit use and the alternative mode (in this case: drive-alone trips). If this mode-specific constant term is not allowed to change in space, a person in the suburbs with low transit ridership will be forced to have the same relative indirect utility for transit use (relative to driving alone) as the person living in the central city with high transit ridership. However, in reality, network effects are responsible for the fact that a person living in a transit-unfriendly suburb derives less utility from choosing transit over driving the car, than the person living in the transit-friendly city.

It should now be easy to see that, as long as the relative utility for using transit in the suburbs is lower than average and, in the

central city, higher than average, transit ridership will be overestimated in the suburbs and underestimated in the central city. Figure 3 demonstrates graphically how not including the network effects variable into the model leads to an omitted variable bias in the regression.

This new insight might at least partially explain why new rail starts in the past decade have had problems with inflated ridership forecasts compared with the observed ridership after opening (Pickrell, 1989; Kain, 1992). Since the federally funded rail projects studied by Pickrell mostly serve the commuter market from the suburbs to the CBD, a systematic bias in the mode choice model caused by unaccounted for spatial autoregression stemming from a network effect could be responsible for overestimating ridership. Not accounting for network effects may have also dwarfed the forecasts of Tren Urbano in Puerto Rico, which in 2005 is less than a third of the originally expected ridership (Green, 2005).

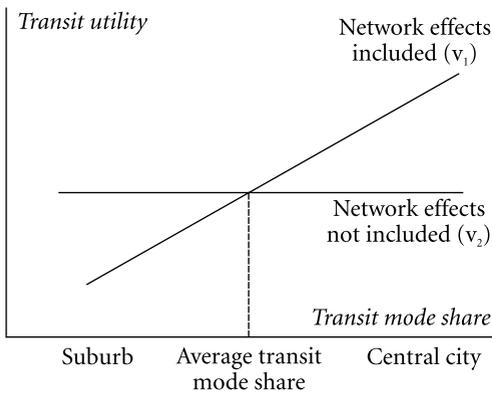


Figure 3. In the presence of network effects, transit utility increases with higher transit mode share. However, if network effects are not included in the model, transit utility is forced to be constant, which leads to an omitted variable bias. Therefore, transit ridership will be systematically underestimated in the suburbs ($v_1 < v_2$) and overestimated in the central city ($v_1 > v_2$).

Conclusion

I have shown that network effects play a role in mode choice decision-making. However, the theory only allows for spatially autoregressive binary mode choice models because of the spatial weight matrix character (Wm^*), with m^* taking on the value of either zero or one, but never a larger number. Given what was discussed in this paper about spatially autoregressive binary mode choice models, it is not difficult to extend the general concept to multinomial mode choice models. The spatial lag term representing the network effects is just the average mode share of all the spatially weighted neighbours. By replacing the social interaction with spatial dependency, this concept can easily be included in the model structure.

An extension of the mode choice model would be not only to include network effects on the origin side of the trip, but also on the destination side of the trip. Intuitively this makes sense, since people choose their mode based on both origin characteristics as well as on destination characteristics. LeSage and Pace (2005) found an elegant solution to model econometrically this type of spatial problem.

While the estimation of a spatially autoregressive mode choice model is feasible as long as the data are collected on the basis of spatial sampling, the real challenge is the inclusion of network effects in mode choice forecasting for new infrastructure investment projects. In traditional mode choice models, the mode share is exclusively determined by social characteristics and trip costs—both assumed to be exogenous. Network effects, however, lead to a portion of the trip cost (WTP for using the mode with others) itself becoming endogenous. At the same time, we do not know what the individual (as opposed to aggregate) preference is for using the transport mode with other people.

This problem is somewhat similar to the issue of congestion in mode choice models, where a portion of the travel cost, the congestion cost, becomes endogenous, just like the network effects. To account for the endogeneity of congestion cost, a simulation approach is typically used, where the model results are iteratively looped until convergence is reached.

In the context of network effects, the simulation approach assumes that socioeconomic characteristics are exogenous and will not change with the implementation of new transit facilities, which essentially means that improved public transit does not trigger anybody to move. Such an assumption is unrealistic, since the social composition of a neighbourhood is, of course, endogenous with respect to transport infrastructure. This problem even exists for traditional mode choice models, but may be alleviated by including a land use model. However, network effects do magnify the neighbourhood mobility problem. People who move into the area because of the new transport choices, such as transit, will then use the new mode now available. In addition, new residents will also prompt already-existing residents to switch to the newly available transport mode. Therefore, a mode choice model that includes network effects always has to integrate a land use model when deriving forecasts.

In the context of the four-step transport model, I am convinced that not just the mode choice step is spatially autoregressive, but also the trip generation model and the trip assignment model. LeSage and Polasek (2005) discuss the possibility of spatial autoregression in the trip distribution model. Naturally, people in the central city make more, but shorter, trips than their suburban counterparts. They also tend to select their destinations—i.e. shopping, based on the destinations of their peers. Especially in congested cities, people might even choose their routes based on what they (in concert

with others) perceive to be the fastest way. However, these route choice network effects can also cause increased congestion. After all, the actual decision-making in difficult traffic conditions will always lack full information with the result that assignment flows may be in a continuous disequilibrium, or in an equilibrium, but with network effects.

Notes

1. Just as a full restaurant is a sign of good food and satisfied customers.
2. While it is in this context possible also to analyse negative congestion externality, it is not the focus of this study.
3. This analysis could potentially be carried out for any city with significant transit ridership and sufficient geospatial data. New York City seemed especially well suited for this task because both the transit ridership and geospatial data requirements were met and, in addition, the New York Metropolitan Transport Council allows public access to its household survey and travel demand forecasting model data. The fact that the dataset is almost 10 years old should not have any impact on the results, because of the general nature of this research.
4. A negative network effects coefficient value would have been a sign of an aggregate personal aversion towards transit crowding which is typically known as congestion externality.

Acknowledgements

The author is grateful to Brian Cushing, Stratford Douglas, David Martinelli and Santiago Pinto for the support and help they provided him in writing this paper. The author would also like to thank three anonymous referees for their useful remarks, as well as the participants of several conference and seminar presentations for their discussion comments. The paper greatly benefited from their ideas and suggestions. Finally, the author would like to express his appreciation to the New York Metropolitan Transport Council for providing him with the necessary data for his analysis, as well as their technical support which was invaluable. All errors are of course the author's.

References

- Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.
- Anselin, L. (2002) *Spatial regression analysis short course*. Unpublished manuscript, University of Michigan, Ann Arbor, MI.
- Ben-Akiva, M. and Lerman, S. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA: MIT Press.
- Boldoc, D., Laferriere, R. and Santarossa, G. (1995) Spatial autoregressive error components in travel flow models: an application to aggregate mode choice, in: L. Anselin and R. J. G. M. Florax (Eds) *New Directions in Spatial Econometrics*, pp. 96–108. Berlin: Springer Verlag.
- Brock, W. and Durlauf, S. (2001) Discrete choice with social interactions, *Review of Economic Studies*, 68(2), pp. 235–260.
- Brock, W. and Durlauf, S. (2002) A multinomial-choice model of neighborhood effects, *American Economic Review*, 92(2), pp. 298–303.
- Fleming, M. (2004) Techniques for estimating spatially dependent discrete choice models, in: L. Anselin, R. Florax and S. Rey (Eds) *Advances in Spatial Econometrics*, pp. 145–168. Berlin: Springer Verlag.
- Goetzke, F. (2003) *Are travel demand forecasting models biased because of uncorrected spatial autocorrelation?* Paper presented at the *North American Meeting of the Regional Science Association International*, Philadelphia, PA, November.
- Goetzke, F. (2006) *Mode choice and social networks: the economics of walking, bicycling and public transit use*. Unpublished dissertation essay, West Virginia University.
- Green, A. (2005) A hesitant Puerto Rico tries commuting by train, *The New York Times*, 19 November.
- Kain, J. F. (1992) The use of strawmen in the economic evaluation of rail transport projects, *American Economic Review*, 82(2), pp. 487–493.
- LeSage, J. P. and Pace, R. K. (2005) *Spatial econometric modeling of origin–destination flows*. Paper presented at the *52nd North American Meeting of the Regional Science Association International*, La Vegas, November.
- LeSage, J. P. and Polasek, W. (2005) *Incorporating transportation network structure in spatial econometric models of commodity flows*. Unpublished paper, University of Toledo.
- McFadden, D. (1974) Conditional logit analysis of qualitative choice behavior, in: P. Zarembka (Ed.) *Frontiers in Econometrics*, pp. 105–142. New York: Academic Press.
- NYMTC (New York Metropolitan Transportation Council) (2004) *New York best practice model (BPM) for regional travel demand forecasting* (<http://www.nymtc.org/project/BPM/model/bpmuserdoc.pdf>).
- Páez, A. and Scott, D. M. (2005) *A discrete choice approach to modeling social influence on individual decision making*. Revised version of a presentation at the *51st North American Meeting of the Regional Science Association International*, Seattle, WA.
- Páez, A. and Scott, D. M. (2007) Social influence on travel behavior: a simulation example of the decision to telecommute, *Environment and Planning A*, 39(3), pp. 647–665.
- Pickrell, D. H. (1989) *Urban rail transit projects: forecasts vs. actual ridership and costs*. US Department of Transportation, Transportation System Center, Cambridge, MA.
- Train, K. (2003) *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press.