Exponential or power distance-decay for commuting? An alternative specification

Paper to be presented at the 45th Congress of the European Regional Science Association, Vrije Universiteit, Amsterdam, 23-27 August 2005

Jacob J. de Vries, Peter Nijkamp, Piet Rietveld

Vrije Universiteit, Amsterdam, and Tinbergen Institute, Amsterdam.

Address corresponding author: Department of Spatial Economics, Vrije Universiteit, De Boelelaan 1105, NL 1081 HV Amsterdam, The Netherlands. E-mail address: jvries@feweb.vu.nl. Web site: http://staff.feweb.vu.nl/jjvries/

Abstract: In this paper we determine the effect of transport cost on commuting flows, on the basis of an analysis of home-to-work journeys between municipalities in Denmark. Special attention is given to a proper estimation method and the form of the distance-decay function. It appears that neither an exponential nor a power distance-decay function fits the data well. The specification of log trips as a (downwards) logistic function of log cost results in a better fit. We find that the cost elasticity of commuting reaches a value of –4 for distances around 24 km, while it is close to 0 for both very short and very long distances. Finally, we demonstrate that the choice of functional form for distance-decay can make an important difference for predictions concerning the effect of infrastructure improvements on commuting flows.

Key words: Spatial Interaction, Distance-decay Function, Commuting, Denmark, Heteroscedasticity.

JEL classifications: C21, J61, R15, R23.

Acknowledgements. The authors would like to thank Anne Kaag Andersen for making the data available and related help, and Jos van Ommeren for constructive comments on an earlier version.

1. Introduction

Commuting has been the subject of increasing public attention and intensive research in recent decades (see Nijkamp and Rouwendal 2004; Rouwendal and Nijkamp 2004). Rising wages have enabled people to spend more money on transport and to choose a residential location further from their work. This increase in travel has prompted improvements in the transportation system, which in turn have further stimulated commuting. In combination with rising car ownership, the longer commuting distances have caused an unprecedented increase in traffic in and around urban areas. The increasing role of commuting has far-reaching consequences for public policy concerning traffic, infrastructure and spatial planning, as well as housing and labor markets. Theoretical and applied research on commuting issues can help to clarify the relevant effects.

An important policy issue is the evaluation of proposed new infrastructure. The opening of roads or railways will reduce generalized transport cost and stimulate commuting. The resulting changes in commuting patterns will also affect regional housing and labor markets. This is relevant for cost-benefit analysis and for the regional redistribution of welfare. To predict the size of these effects, it is important to know how commuting flows depend on transport cost. Further, commuting flows connect labor and housing markets. Knowledge of the intensity and spatial extent of this interaction is relevant for spatial planning. This knowledge can, for instance, be used to determine the accessibility of a location to centers of population and employment.

The analysis of the sizes of commuting flows between cities or regions, and their effects on housing and labor markets, is part of the much wider area of spatial interaction modeling. Spatial Interaction Models are applied not only to commuting but also to migration, international trade, shopping behavior, and other topics related to origin-destination flows. An important issue in such models is the specification of the functional form of distance or travel cost. In the literature, this distance-decay function is usually assumed to be an exponential or a power function.

> "In practice, the debate over the form of the cost function in spatial interaction models has evolved to a reasonably widespread consensus that the exponential function is more appropriate for analysing short distance interactions such as those that take place within an urban area. The power function, conversely, is generally held to be more appropriate for analysing longer distance interactions such as migration flows" (Fotheringham and O'Kelly 1989, pp. 12-13).

We will propose a more flexible specification of the decay function. The research presented in this paper is an application of Spatial Interaction Modeling, and the methodological issues are not limited to commuting.

The purpose of this paper is to determine the effect of transport cost on commuting flows between regions. We use data on commuting flows and travel cost between the 275 municipalities in Denmark in 1995. For Denmark, a set of complete and disaggregated data is available. The distance-decay equation which we estimate in this paper is part of several types of Spatial Interaction Models. It describes trips as being dependent on balancing factors for origins and destinations, and as a function of generalized cost. This is a partial analysis, and not an equilibrium approach. The relationships between accessibility to jobs and labor force, on the one hand, and the location of population and employment, on the other, are outside the scope of this paper. For predictions, it is important to take feedback effects into account, but econometric inference can be based on a single equation.

There are some econometric complications in the estimation of the trips equation of a Spatial Interaction Model. First, the equation contains not only a distance-decay function but also unobserved balancing factors. Secondly, zero flows are difficult to handle. Thirdly, the variance of a flow will be related to its expectation, so there is heteroscedasticity. Solutions for these problems have been proposed in the literature. In this paper we apply a nonlinear weighted least-squares estimation (NLWLS) method, that is mainly the nonlinear version of Procedure 1 described in Sen and Smith (1995, pp. 485-486). However, we improve the treatment of heteroscedasticity in two ways: we base the weights on a combination of Poisson error and lognormal specification error, and we use weighted averages to estimate the balancing factors. As the positive flows vary from one to thousands, heteroscedasticity can be considerable.

We do not impose a distance-decay function[1] from theory, but infer the functional form from the data. It appears that Danish commuting flows cannot be described by an exponential distance-decay function. A power specification gives better results, but performs poorly for very small and very large distances. To investigate this, we introduce a piecewise power function. The estimated elasticities are (in absolute value) small for small distances, large for intermediate distances, and then they decrease gradually as the distance increases. This ad-hoc specification can be approximated using a logistic function. In our final

[1] As costs and distances are highly similar, we often use the term “distance-decay function”, while it is actually a function of cost.

specification, log trips is a downwards logistic function of log generalized cost. The bending point of the logistic function, where commuting is most elastic, lies near 24 km. That the cost elasticity of commuting varies over distance is important for the evaluation of infrastructure projects. Major new links have been constructed in Denmark, such as the Great Belt crossing. The use of a power-decay function, which imposes constant elasticity, might lead to overestimation of the effect of new long-distance connections, and underestimation of the effect of shorter-distance infrastructure.

The paper is structured as follows. Section 2 describes the model and the data, and Section 3 the estimation method. In Section 4 we estimate various distance-decay functions. In Section 5 we discuss the results. Finally, Section 6 concludes. Several derivations are placed in Appendices A and B.

2. Model and data

Spatial Interaction Models describe flows between a set of origins and a set of destinations. For a general introduction to Spatial Interaction Models, we refer to Fotheringham and O'Kelly (1989). Most Spatial Interaction Models can be represented in a form that contains the following equations (Batten and Boyce 1986; De Vries et al. 2001):

$$T_{ij} = A_i B_j O_i D_j F_{ij} \tag{1}$$

$$O_i = \sum_j T_{ij} \tag{2}$$

$$D_j = \sum_i T_{ij} \tag{3}$$

where T_{ij} is the commuting flow[2] ('trips') from municipality i to municipality j; A_i and B_j are unobserved balancing factors, which make it possible to satisfy (2) and (3); O_i is the total number of working people living in (origin) municipality i, so that it includes not only the outward commuting flow but also the people living and working in that place; D_j is the total number of workers in (destination) municipality j; and F_{ij} is a decreasing function of the generalized travel cost from municipality i to municipality j, G_{ij}.

In the literature, it is generally agreed that an exponential specification

$$F_{ij} = \gamma_0 e^{\gamma_1 G_{ij}} \tag{4}$$

[2] If $i = j$, this flow is the number of people both working and living in that region.

is in accordance with a utility framework (Cochrane 1975). However, in applications often a power specification

$$F_{ij} = \gamma_0 G_{ij}^{\gamma_1} \tag{5}$$

appears to have a better fit. Choukroun (1975) provides a possible theoretical justification for a power specification. Deviations from an exponential form towards a power form can be caused by heterogeneity of user preferences, use of various travel modes, and lower trip frequency for long-distance commuters. The resulting functional form depends on the mathematical assumptions concerning the heterogeneity. Other types of decay functions are not excluded in advance.

We try to answer the question concerning the functional form of the distance-decay function from an empirical starting-point. So we specify F_{ij} very generally as:

$$\ln F_{ij} = f(G_{ij}) + u_{ij} \tag{6}$$

The function f, which represents[3] the distance-decay, will be chosen in the course of the inference, in Section 4. The construction of generalized cost G_{ij} will be discussed at the end of the present section. Assumptions on the disturbance term u_{ij} will be stated in Section 3. The estimation of the distance-decay function will be based on (1) to (3) in combination with (6), a set of equations which can be part of several types of Spatial Interaction Models.

In this paper we do not specify a complete model. A complete model can be obtained by adding equations for O_i and D_j. The balancing factors A_i and B_j can be interpreted as accessibility indicators (Alonso 1978; Hua 1980). If A_i is low, residence municipality i has good access to jobs. If B_j is low, work municipality j has good access to the labor force. The difference between the various types of Spatial Interaction Models lies in assumptions concerning the relation between origin and destination totals and accessibility. In the genuine doubly-constrained model, O_i and D_j are assumed to be exogenous. If O_i is proportional to accessibility to jobs (A_i^{-1}) and D_j is proportional to accessibility to the labor force (B_j^{-1}), the unconstrained gravity model (Carey 1858; Stewart 1941) results. Singly-constrained models are a combination of these (Wilson 1971). In Alonso's Theory of Movements (Alonso 1978; Hamerslag 1980; Bikker 1987; Bikker and De Vos 1992), the relation between O_i and

[3] Note that f is not the distance-decay function itself, but its logarithm. For most functional forms this is a convenient representation, e.g. if the decay function is exponential, f is linear.

A_i and the relation between D_j and B_j are specified in additional equations, containing parameters to be estimated.

The estimation of the distance-decay function will be based on data on commuting between municipalities in Denmark in 1995. Denmark consists of the Jutland peninsula, a few sizeable islands, and many small islands. In 1995 Denmark had 5.2 million inhabitants.[4] Population distribution between regions is rather uneven. Large parts of Jutland are sparsely populated. The share of the Copenhagen metropolitan area in total population is over 25%.

The Great Belt (strait) divides the country into two parts, with approximately half of the population on each side. West of the Great Belt 2.4 million people live in Jutland and 0.5 million on Funen. East of the Great Belt, the population of Zealand is 2.2 million (including 1.3 million in Greater Copenhagen), and 115,000 people live on Lolland-Falster. In 1995 the fixed (road and rail) link over the Great Belt was not yet finished. Various ferry routes provided the connection over the Great Belt. Within the Jutland/Funen region and within the Zealand/Lolland/Falstar region, bridges connect the major islands. The island of Bornholm (population only 45,000) is farther east, in the Baltic Sea, and is connected to the rest of Denmark by ferries.

We use the following data:[5]

- T_{ij}: a 275 by 275 table of the active population in Denmark in 1995 by municipality of residence and municipality of work. This table results from the population census. Although we will call these persons 'commuters', we have to keep in mind that especially for longer distances people probably do not make the trip every day.
- O_i and D_j: origin and destination totals, computed from T_{ij} using (2) and (3).
- C_{ij}: a 275 by 275 table of travel cost. This is the financial cost of a commuting trip by car, in Danish crowns (DKK). This cost is constructed as the land distance (as the crow flies)

[4] Population data from Statistics Denmark.

[5] Data on the commuting between municipalities in Denmark in 1995, travel costs between the municipalities in 1995, as well as a list of the municipalities and their official statistical numbers, were made available by Anne Kaag Andersen of the Danish Institute of Local Government Studies AKF, http://www.akf.dk/. More specific definitions of the background data including travel cost are given in Andersen (1999). We obtained the division of the municipalities over the counties from http://www.danmark.dk/. In addition, we used several maps to locate outliers.

times a kilometer price (set to 1 DKK/km), plus the rate of a possible ferry (for a car and driver), minus the tax deduction for commuting cost[6] (Andersen 1999).

- The administrative division of Denmark. We will construct dummy variables for specific combinations of origin and destination counties, to represent geographical effects (islands) and urbanization effects (congestion). Figure 1 shows the counties of Denmark (with their Danish names).

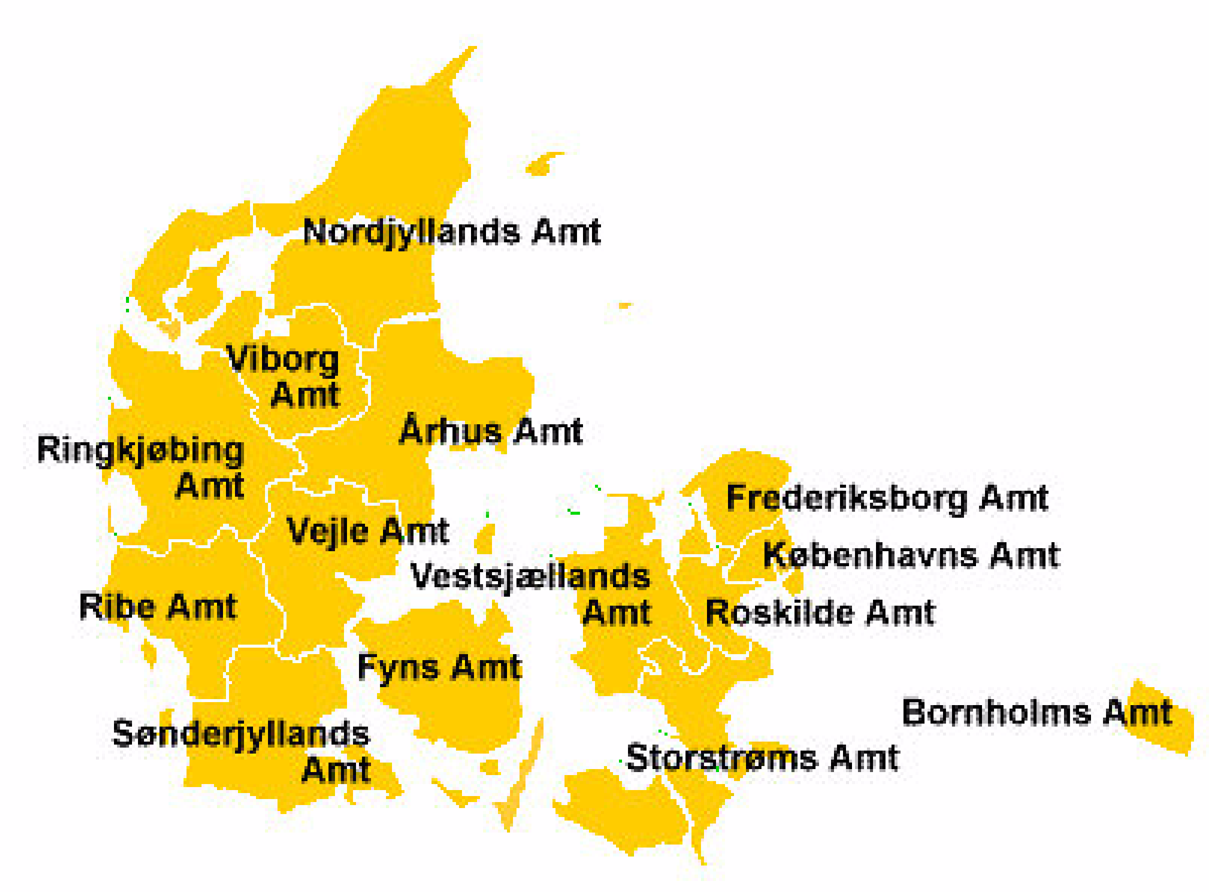

Figure 1. Map of Denmark, with division into counties. The Great Belt is the strait between Vestsjællands Amt and Fyns Amt.

[6] One DKK is approximately 13 eurocents. The fact that the kilometer rate is set to the round number of 1 DKK implies that the travel cost numerically equals the distance, as long as no ferries are involved and we ignore the tax deduction. This will be useful for the interpretation of the estimated distance-decay functions. If the actual kilometer price were to deviate from this, the interpretation in kilometers would still hold, but a correction to ferry rates would be required.

Table 1. Definitions of dummy variables [a]

Dummy variable	is 1 for and only for (0 else)
`GreatBelt`	trips from Københavns Amt, Frederiksborg Amt, Roskilde Amt, Vestsjællands Amt, Storstrøms Amt to Fyns Amt, Sønderjyllands Amt, Ribe Amt, Vejle Amt, Ringkjøbing Amt, Århus Amt, Viborg Amt, Nordjyllands Amt or the other way around
`Bornholm`	trips from/to Bornholms Amt (but not within Bornholms Amt)
`Copenhagen`	trips within Københavns Amt [b]
`Isefjord`	trips from Frederiksborg Amt to Vestsjællands Amt or the other way around

[a] In the definition of dummy variables Københavns Amt (county Copenhagen) is taken to include the municipalities Copenhagen and Frederiksberg, which are formally not part of the county Copenhagen but enjoy their own county status.
[b] This covers the major part of the metropolitan area.

As stated above, ferry rates are included in the cost measure.[7] However, the cost data does not include time cost. Usually, distance, travel cost, and travel time are highly correlated, but in the case of ferries the ratio might differ from that on the roads. Therefore, we include dummy variables for the crossing of the Great Belt and travel from/to Bornholm. We also include a dummy variable for the area around Copenhagen, to compensate for the relatively high travel time in congested areas. Finally, we include a dummy for travel around the Isefjord, which is the bay in the north of Zealand.[8] Table 1 provides the exact definitions of

[7] Note that the cost is based on the rate for bringing a car on the ferry. If a substantial part of the commuters use the ferry as foot passengers or by train (in both cases at lower cost), the cost is overestimated in the data. We do not have data on mode choice. Generally, most of the commuting is by car, but on larger distances the share of the train might be higher.

[8] As both sides of the Isefjord are on the same island, the cost in the data is computed over the distance as the crow flies, while in reality a longer trip around the Isefjord or a more expensive ferry passage is required.

the dummy variables. The cost data C_{ij} are combined with the dummy variables[9] in order to obtain generalized cost G_{ij}.

$$G_{ij} = C_{ij} + \delta_1 * \text{GreatBelt}_{ij} + \delta_2 * \text{Bornholm}_{ij} + \delta_3 * \text{Copenhagen}_{ij} * C_{ij} + \delta_4 * \text{Isefjord}_{ij} / C_{ij} \quad (7)$$

We omit the observations on intrazonal flows in all graphs and estimations presented in this paper.[10] In our analysis we first included them, using an approximation for the distance within a municipality, and a dummy variable for intrazonal flows. This led to estimation problems for several specifications of the decay function. To understand that, it is relevant that we model dummy variables as corrections to generalized cost, not as separate effects on trips. Intrazonal flows appear to be larger than the model would predict for zero generalized cost. So the problem is not in the measurement of travel cost within a municipality, but in the fact that commuting behavior within a municipality apparently differs from that between municipalities. Home-based workers are the probable cause of the difference. Therefore, we exclude the intrazonal flows from the further analysis.

The equation which we are going to estimate results from the substitution of (7) into (6) into (1). Some elements are not yet specified: the distribution of the disturbance term will be discussed in Section 3; and in Section 4 we will investigate various functional forms for the distance-decay function.

3. Estimation method

The estimation method that we use in this paper combines known elements from the literature with a few innovations. We apply the least-squares method of Cesario (1974), but add to that the use of weights, based on a combination of specification error and Poisson error. Our specification of the decay function and generalized cost makes the estimation nonlinear. For the treatment of small and zero flows, we use a method developed by Sen and Soot (1981). Our estimation method is similar to “Procedure 1” of Sen and Smith (1995, pp. 485-486), but differs in three respects. Firstly, it is nonlinear, to allow for more complicated decay functions. Secondly, the weights differ, as we think that specification error can not be ignored. Thirdly, the weights are also used when deviations from the average are taken, as

[9] The way the dummy variables affect generalized cost differs. The correction for the use of a ferry is a fixed amount. The correction for congestion is proportional to the cost. The correction for the Isefjord is large for low cost, and decreases as cost increases.

[10] The intrazonal flows are, however, still included in the origin and destination totals.

follows from the first-order conditions for weighted least-squares (WLS) (De Vries et al. 2002). In the remainder of this section, we describe the estimation method in more detail.

The estimation is based on (1). If we take logs and substitute (6) for $\ln F_{ij}$, we have:

$$\ln T_{ij} = \ln A_i + \ln B_j + \ln O_i + \ln D_j + f(G_{ij}) + u_{ij} \tag{8}$$

This equation contains the endogenous variables A_i, B_j, O_i and D_j, but there is no endogeneity problem, as these variables do not have an associated parameter. O_i and D_j are observed, and A_i and B_j will be estimated. For prediction it is important to use a complete model. For estimation purposes, however, we can concentrate on a subset of equations. (2) and (3) are just definitions. We can estimate (1), irrespective of how equations for O_i and D_j are specified. So the methods and results presented in this paper are valid for a wide range of Spatial Interaction Models. In Alonso's Theory of Movements, the estimated A_i and B_j serve as input for a further estimation stage, where the relationship between the balancing factors and the inflows and outflows is investigated (Ledent 1980; Poot 1986; De Vries et al. 2002).

In (8) we use logarithms of the trips. But there are two problems associated with that. The logarithmic transformation of the dependent variable will cause a bias, as the expectation of the logarithm is not equal to the logarithm of the expectation (Sen and Soot 1981, p. 167). And many flows are zero, so that in these cases the logarithm does not exist. To solve these problems, we apply the suggestion of Sen and Soot (1981, p. 167) to add ½ to each flow before taking logs. Sen and Smith (1995, p. 480) assume that the trips have a Poisson distribution, so $\operatorname{var} T_{ij} = ET_{ij}$. From a Taylor-approximation around ET_{ij}, it can then be derived (using the equality of variance and expectation) that:

$$E\ln(T_{ij} + \tfrac{1}{2}) \approx \ln ET_{ij} \tag{9}$$

PROOF:

$$E\ln(T_{ij} + \tfrac{1}{2}) \approx E\left[\ln ET_{ij} + \frac{1}{ET_{ij}}\left(T_{ij} + \tfrac{1}{2} - ET_{ij}\right) - \frac{1}{2}\frac{1}{\left(ET_{ij}\right)^2}\left(T_{ij} + \tfrac{1}{2} - ET_{ij}\right)^2\right]$$

$$= \ln ET_{ij} + \tfrac{1}{2}\frac{1}{ET_{ij}} - \frac{E\left\{\left(T_{ij} - ET_{ij}\right)^2 + \left(T_{ij} - ET_{ij}\right) + \tfrac{1}{4}\right\}}{2\left(ET_{ij}\right)^2} = \ln ET_{ij} + \frac{1}{2ET_{ij}} - \frac{\operatorname{var} T_{ij}}{2\left(ET_{ij}\right)^2} - \frac{1}{8\left(ET_{ij}\right)^2} = \ln ET_{ij} - \frac{1}{8\left(ET_{ij}\right)^2}$$

If ET_{ij} is not too small, the latter term can be neglected.

END OF PROOF.

So the use of $\ln(T_{ij} + \tfrac{1}{2})$ as the dependent variable will strongly reduce the bias (Sen and Soot 1981; Sen and Smith 1995, p. 480). It also provides a solution to the problem of zero flows.

Sen and Soot (1981, p. 168) and Sen and Smith (1995, p. 484) suggest applying WLS, as heteroscedasticity may be expected. They assume that the flows have a Poisson distribution. In that case:

$$\operatorname{var}\ln(T_{ij}+\tfrac{1}{2}) \approx 1/ET_{ij} \tag{10}$$

and weights can be derived from that. However, a disadvantage of the assumption of a Poisson distribution is that it does not take specification error into account (De Vries et al. 2002). There are factors affecting the commuting pattern that are not included in the model. If this is ignored, large flows will get much too heavy a weight. An alternative approach is to assume that u_{ij} has constant variance σ_u^2. This will work well to treat specification error, but ignores the fact that small flows have a relatively larger standard deviation. Therefore, we combine both variances,[11] and assume:

$$\operatorname{var}(u_{ij}) = \sigma_u^2 + 1/ET_{ij} = \sigma_u^2\left(1+\frac{1}{\sigma_u^2 ET_{ij}}\right) \tag{11}$$

We estimate σ_u^2 in the regression, and weigh the observations by:

$$w_{ij} = \left(1+\frac{1}{\sigma_u^2 ET_{ij}}\right)^{-\frac{1}{2}} \tag{12}$$

This implies that zero flows get a very small weight if the expected flow is close to zero, while the weight approaches 1 for large expected flows. The weights depend on both σ_u^2 and ET_{ij}, which have to be estimated, so we iterate between the WLS procedure and the calculation of the weights. The iteration has the additional advantage that we can use expected instead of observed values in the weights, thus taking away a source of bias. We exclude the intrazonal flows by setting their weights to zero.

For most choices of decay function f, the equation resulting from substitution of (7) into (8) is nonlinear. So we apply nonlinear least-squares (Stoer and Bulirsch 1980, pp. 209-210). This is an iterative procedure. A regression based on a linear approximation around the most recent parameter estimates results in updates for these estimates. The iteration for the weights can be integrated in the iteration of the nonlinear least-squares procedure.

[11] The introduction of specification error has no consequences for the procedure of adding ½ to each observation before taking logs. The data generating process can be considered to consist of two steps. A value drawn from a lognormal distribution is used as the expected value in a Poisson distribution. Taking logs of a Poisson variable requires correction by adding ½. Taking logs of a lognormal variable results in a normally distributed variable.

The unobserved balancing factors A_i and B_j in (8) need to be estimated. Basically, we use the method proposed by Cesario (1974, p. 252). The idea is that (8) can be estimated by least-squares, including dummy variables for all origins and destinations. However, with 550 dummy variables, a standard regression procedure will take a considerable amount of computer time and memory (Sen and Smith 1995, p. 476). Therefore, we decompose (8) into the overall average, an i-component, a j-component, and the remainder. The latter component is the deviation from the average over i and the average over j, so the balancing factors cancel. The parameters of the distance-decay function can then be estimated by regression (Cesario 1974, p. 252; Sen and Smith 1995, p. 477). However, the use of WLS has implications for the decomposition, so our method is somewhat different. Appendix A describes the decomposition in more detail, including the role of weights. If the equation is nonlinear, this decomposition needs to be applied to the linear approximation in each iteration step.

To summarize, we estimate:

$$\ln(T_{ij} + \tfrac{1}{2}) = \ln A_i + \ln B_j + \ln O_i + \ln D_j + f(G_{ij}) + u_{ij} \qquad (13)$$

where G_{ij} is given by (7), by NLWLS, using weights given in (12). This procedure[12] results in estimates for: the parameters in the distance-decay function; the dummy parameters that are included in the generalized cost; and balancing factors for all origins and destinations. In the next section we will apply this method to various specifications of the decay function.

4. Distance-decay function

In this section we analyze the relation between trips, origins, destinations, and travel cost. We investigate various functional forms for the distance-decay function. The data we use are on a fairly low level of aggregation, so the data set is large. Of the over 75,000 data points on municipality-to-municipality commuting flows, 64% are 0. So the issue of how to handle zero flows is important here. As was discussed in the previous section, the zero flows are included in the analysis, but the weight they get is small if the expected flow is close to zero. Also the positive flows are usually small. 13% of the observations are equal to 1, 14% lie between 2 and 10, and only 9% are larger than 10. Flows within municipalities are omitted

[12] The computations for this paper were done using the package Ox (Doornik 1998). For more information on Ox, see http://www.nuff.ox.ac.uk/Users/Doornik/. The program and the text part of the output are available on request from the authors or on their web site (see first page of this paper).

Table 2. Number of trips per cost category

Cost category	Number of commuters	Share in total	Cumulative share	Average commuting flow[a]
Local	1,560,616	59.15%	59.15%	5,674.97
< 10	208,685	7.91%	67.05%	709.81
10 – 15	316,072	11.98%	79.03%	391.18
15 – 20	242,697	9.20%	88.23%	169.24
20 – 30	167,972	6.37%	94.60%	48.09
30 – 50	61,425	2.33%	96.92%	11.81
50 – 100	41,440	1.57%	98.49%	3.30
100 – 200	11,797	0.45%	98.94%	0.86
200 – 300	6,854	0.26%	99.20%	0.68
300 – 400	16,183	0.61%	99.81%	0.81
> 400	4,885	0.19%	100.00%	0.64
Total	**2,638,626**	**100.0%**		**34.89**

[a] Average commuting flow = the number of commuters divided by the number of origin-destination combinations in that cost category.

from the analysis. The largest flows between different municipalities are those between Copenhagen and nearby municipalities: 20,846 commuters from Frederiksberg to Copenhagen and 12,808 the other way around, and 12,949 commuters from Gentofte to Copenhagen.

Table 2 shows how the commuting is distributed over cost categories. More than half of the working population works in the municipality of residence. 5% have a commuting cost of over 30 DKK (equivalent to 30 km as the crow flies). 1% have a commuting cost of over 200 DKK. Those cases usually include a ferry passage. The last column of Table 2 gives the average size of the commuting flows, i.e. the number of commuters divided by the number of origin-destination combinations in that cost category. The average flows decrease rapidly with cost, up to a cost of 100 DKK. Note that the figures in this table are affected not only by the role of transport cost in commuting but also by the spatial configuration of cities.

To get an idea about the functional form of the distance-decay function, we made some plots based on a rearrangement of (13):

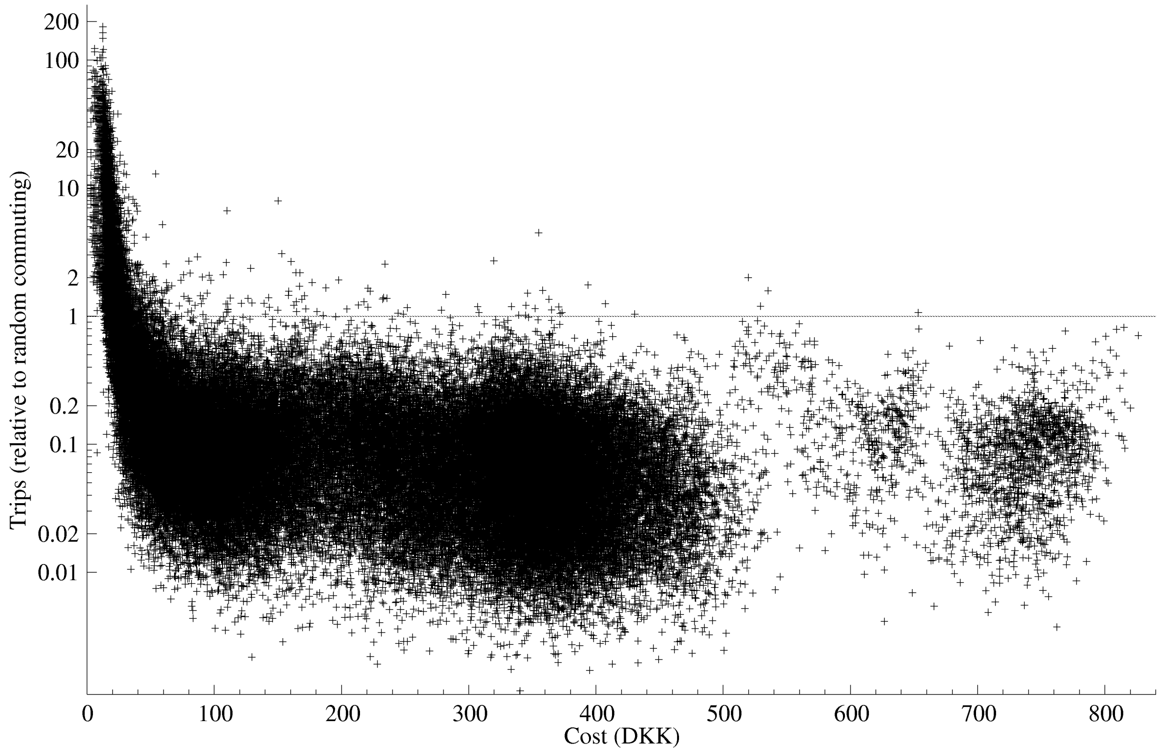

Figure 2. The logarithm of the commuting flow (relative to random commuting) as a function of travel cost (flows within municipalities omitted). If the decay function were exponential, the points would lie around a straight line.

$$\ln(T_{ij} + \tfrac{1}{2}) - \ln O_i - \ln D_j = \ln A_i + \ln B_j + f(G_{ij}) + u_{ij} \tag{14}$$

We want to decompose the observations on the left-hand side of (14) into balancing factors, a function of cost and geographical dummies, and a disturbance term. The balancing factors represent the accessibility of the municipalities, which is related to their position relative to other municipalities. If we ignore differences in accessibility for the moment, a plot based on (14) can indicate the functional form of the decay function f. Figure 2 shows the left-hand side of (14) against the cost.[13] If the distance-decay function were exponential, the points would be around a straight line. This is clearly not the case. First, there is a sharp decrease,

[13] We use observed cost here, not generalized cost. Intrazonal flows are omitted. To improve readability, we scaled the variable on the vertical axis, $(T_{ij} + \tfrac{1}{2})/(O_i D_j)$, by multiplying it by the total number of workers in Denmark, $\sum_i O_i$ (which of course equals $\sum_j D_j$). In this way, we compare the size of the commuting flow with the theoretical case of random commuting, where distance is irrelevant. Setting F_{ij} equal to $1/\sum_i O_i$ in (1), and all balancing factors to 1, will produce this case (Thorsen et al. 1999).

and then the graph is roughly horizontal. In Figure 3 we made the plot using the log of cost. This is more like a constant decrease. This corresponds to a power distance-decay function.

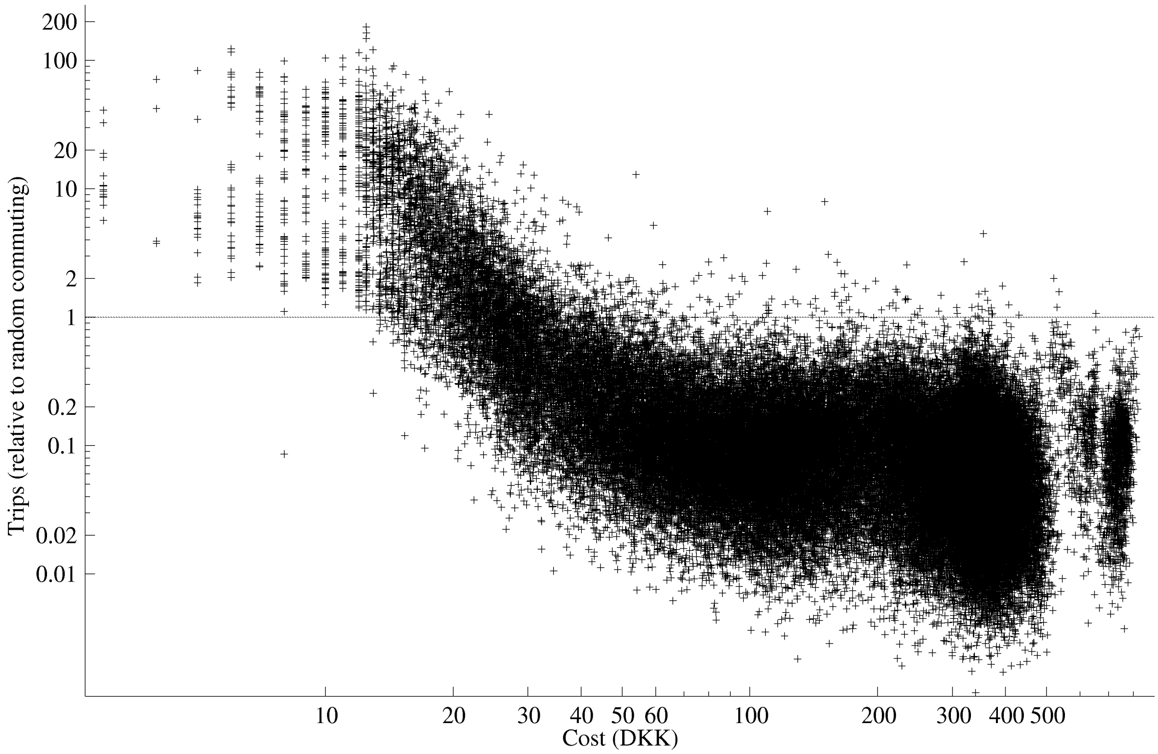

Figure 3. The logarithm of the commuting flow (relative to random commuting) as a function of the logarithm of travel cost (flows within municipalities omitted). If the decay function were a power function, the points would lie around a straight line.

For completeness, we first estimate the exponential decay function:

$$\ln(T_{ij} + \tfrac{1}{2}) = \gamma_0 + \ln A_i + \ln B_j + \ln O_i + \ln D_j + \gamma_1 G_{ij} + u_{ij} \quad (15)$$

where generalized cost G_{ij} is a function that includes several dummies, see (7). Table 3 gives the result of the estimation. The standard deviation of the residual is 0.645 (so there is a 64.5% error, which is rather large.). All parameter estimates are extremely significant.[14] The estimated cost coefficient implies that the number of commuters is reduced by 1.8% if the cost rises by 1 DKK (or kilometer). The estimate for the Copenhagen dummy has not only the

[14] The standard errors for the parameter estimates are computed in the usual way. This implies that they are valid under the assumption of independent, identically distributed disturbances. This assumption is violated in the present application. Differences in quality of the transportation network will affect multiple flows. In particular, the disturbance for the flow from i to j will be correlated with the disturbance for the flow from j to i. So the standard errors are underestimated. As all estimates are extremely significant, this is no great problem. (This holds for all estimates in Table 3.)

wrong sign but also an unacceptable value: it indicates that transport cost is 561% lower around Copenhagen, and so is negative. Figure 4a, which shows the residuals as a function of log cost, also indicates that exponential decay is not the correct functional form.

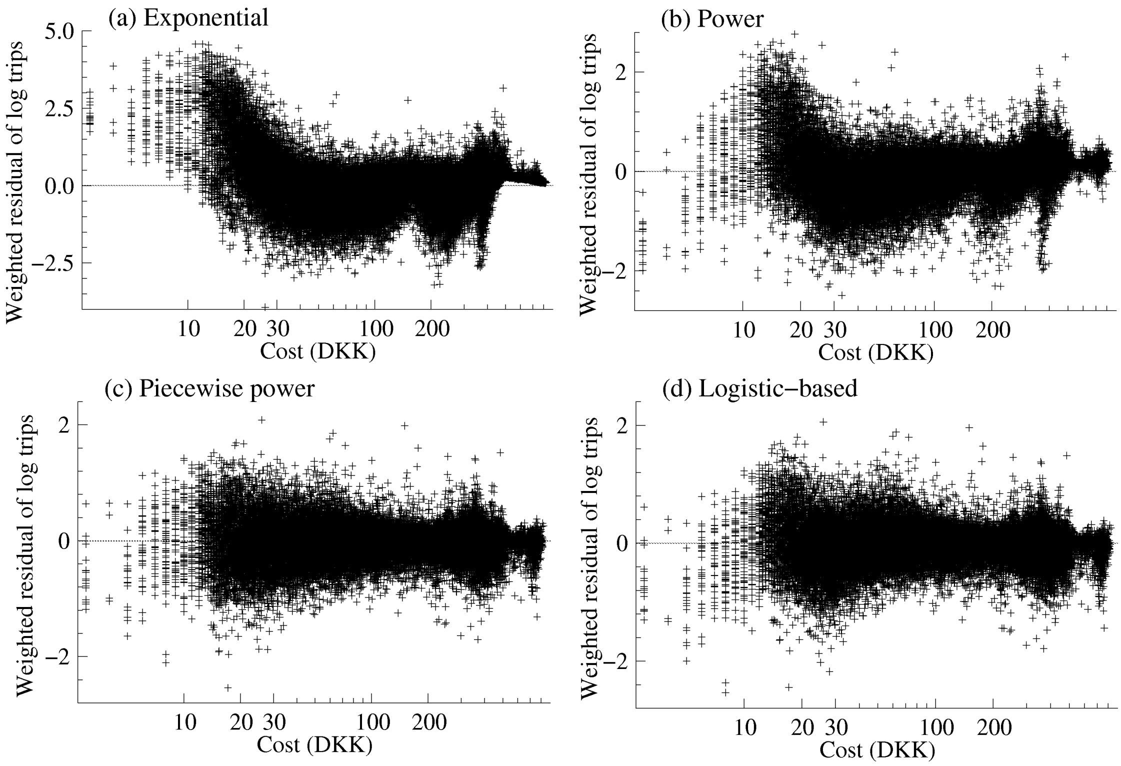

Figure 4. Weighted residuals of various decay functions (as functions of the logarithm of travel cost)

From Figures 2 and 3 we inferred that a power specification might do better, and for the moment we will continue with that specification. So we estimated a power-decay function by NLWLS:

$$\ln(T_{ij} + \tfrac{1}{2}) = \gamma_0 + \ln A_i + \ln B_j + \ln O_i + \ln D_j + \gamma_1 \ln G_{ij} + u_{ij} \qquad (16)$$

The results are also presented in Table 3. We see that the power function performs much better than the exponential decay function. R^2 is much higher, and the residual standard deviation is reduced to 35.5%. The cost parameter in this power specification is close to the classical value of -2. From Figure 4b it appears that the specification (16) is still not the correct functional form. The residuals show an upward tendency up to a cost of around 15; a downward tendency for intermediate cost; and again an upward tendency for cost over 40. A straight line and some dummy variables can not describe the pattern in Figure 3.

To investigate the functional form, we allowed some kinks in the decay function. We replaced (16) by a specification that is on intervals linear in $\ln G_{ij}$, and is continuous, but

Table 3. Estimation results for various decay functions

	Exponential	Power	Piecewise	Logistic-based
R^2	0.704	0.900	0.953	0.950
Standard deviation of residuals	0.645	0.355	0.228	0.237
Constant	-15.109 ****	-7.970 ****	22.461 ****	-18.417 ****
	(0.008)	(0.018)	(0.120)	(0.008)
Cost	-0.0178 ****			
	(0.0001)			
Log Gen.C		-2.070 ****	-1.194 ****	
		(0.005)	(0.023)	
> 15			-4.204 ****	
			(0.017)	
> 30			-2.356 ****	
			(0.025)	
> 50			-1.713 ****	
			(0.024)	
> 100			-0.815 ****	
			(0.041)	
> 150			-0.381 ****	
			(0.020)	
Loglogist				7.926 ****
				(0.034)
Bending point				24.158 ****
				(0.108)
Steepness				1.823 ****
				(0.011)
GreatBelt	-147.3 ****	-126.21 ****	-87.0 ****	-76.39 ****
	(1.7)	(0.64)	(2.6)	(4.45)
Bornholm	-180.2 ****	-209.89 ****	-258. ****	-274.30 ****
	(4.7)	(3.84)	(4.4)	(3.65)
Copenhagen	-5.61 ****	0.113 ****	0.296 ****	0.286 ****
	(0.17)	(0.014)	(0.007)	(0.008)
Isefjord	2812.8 ****	650.7 ****	308.9 ****	287.5 ****
	(67.9)	(34.3)	(20.5)	(15.3)

Standard errors in parentheses.

* 10% significant, ** 5% significant, *** 1% significant, **** 0.1% significant.

changes in slope at various points. We set the kink points at generalized cost of 15, 30, 50, 100 and 150. (These values can roughly be interpreted as distances in kilometers.)

$$\ln(T_{ij}+\tfrac{1}{2}) = \gamma_0 + \ln A_i + \ln B_j + \ln O_i + \ln D_j +$$
$$\gamma_1 \ln \min(G_{ij},15) + \gamma_2 \ln \max(15, \min(G_{ij},30)) +$$
$$\gamma_3 \ln \max(30, \min(G_{ij},50)) + \gamma_4 \ln \max(50, \min(G_{ij},100)) +$$
$$\gamma_5 \ln \max(100, \min(G_{ij},150)) + \gamma_6 \ln \max(150, G_{ij}) + u_{ij} \quad (17)$$

The estimation results for this piecewise decay function are also given in Table 3.[15] The standard error of the residuals is further reduced, to 0.228. Figure 4c shows that the problems with the functional form have been solved, although there is still some heteroscedasticity left. The decay function has a clear pattern. For cost < 15, it is much flatter than the power function. Between 15 and 30, it is very steep, with an elasticity that is in absolute value greater than 4. For cost > 30, the function becomes gradually flatter. The development of the coefficients is rather regular. Except for the first coefficient, the pattern is that, when the cost doubles, the slope of the decay function approximately halves.

The piecewise decay function in (17) can be approximated using a (downwards) logistic function of $\ln G_{ij}$:

$$f(G_{ij}) = \gamma_0 + \gamma_1 \frac{1}{1+\gamma_2^{-\gamma_3} e^{\gamma_3 \ln G_{ij}}} = \gamma_0 + \gamma_1 \frac{1}{1+\left(\frac{G_{ij}}{\gamma_2}\right)^{\gamma_3}} \quad (18)$$

This is demonstrated in Appendix B. The parameter γ_2 determines the location of the bending point (in the same units as generalized cost). The parameter γ_3 governs the steepness of the logistic function. A logistic function has the properties that we noted. It starts rather flat, then becomes steeper, and then gradually flatter again. So the next specification we estimated is:

$$\ln(T_{ij}+\tfrac{1}{2}) = \gamma_0 + \ln A_i + \ln B_j + \ln O_i + \ln D_j + \gamma_1 \frac{1}{1+\left(G_{ij}/\gamma_2\right)^{\gamma_3}} + u_{ij} \quad (19)$$

The parameter estimates are listed in the last column of Table 3, and the residuals are depicted in Figure 4d. For the logistic-based decay function the standard error of the residuals is 0.237, slightly higher than for the piecewise specification. However, the logistic-based specification uses three parameters less, and does not depend on a choice of kink points. We conclude that log trips can be adequately described as a logistic function of log generalized cost, with origin and destination dummies added.

[15] The iteration for the piecewise decay function did not converge fully, but cycled in a small parameter range.

5. Results and discussion

In this section we will discuss the results of the estimation. In Figure 5 we plotted the four decay functions[16] for cost up to 50 (which can roughly be interpreted as kilometers). The exponential decay function is so close to the horizontal axis that it is hardly visible.[17] The logistic-based decay function starts at a finite value for zero cost and decreases smoothly. Both the power function and the piecewise power function start at infinity for zero cost. The value of the power function is considerably below that of the logistic-based for cost around 10. Table 4 lists the values of the decay functions for various values of generalized cost. Figure 6 shows the decay functions with logarithmic axes. This graph shows the different characteristics of the four decay functions, and can be compared to the point cloud in Figure 3.

The bending point of the logistic-based distance-decay function lies at 24 DKK. This is in accordance with the results for the piecewise power distance-decay function, which has the largest elasticity for the cost range 15 to 30 DKK. These numbers can more or less be interpreted as kilometer distances, although for the Copenhagen region a correction is required due to the dummy variable. Estimates for the dummy parameters are highly similar in the logistic-based and the piecewise specification. The estimated parameters for the ferry dummies are both negative, reflecting that the generalized cost is overestimated in the data. Ferries may be slow compared with car traffic, but they are even more expensive. Inasfar as travel cost functions as a proxy for travel time, it is overestimated for relations using a ferry. These two parameters can be interpreted as amounts in DKK. The estimated coefficient for the Bornholm dummy is large. Commuting from and to Bornholm is not usually done by car, so the high ferry cost is not very relevant here. Around Copenhagen, travel cost per kilometer is nearly 30% higher. The Isefjord dummy parameter has no direct interpretation.

The elasticity of –2.07 estimated in the power specification gives a too simple representation of reality. For the logistic-based decay function, the elasticity varies between 0

---

[16] Figures 5 and 6 and Table 4 compare the four estimated decay functions. Again, we scaled F_{ij} by multiplying it by the total number of workers in Denmark, $\sum_i O_i$ (which of course equals $\sum_j D_j$). In this way, we compare the values of the decay function with the theoretical case of random commuting, where distance is irrelevant. (See footnote 13.)

[17] Exponential decay functions can be useful for short distances. However, in this case, the exponential decay function was estimated using data for a wide range of distances. The result is that the estimated function is rather flat, and too low for short distances. For costs between 30 and 200, it has the correct level (cf. Figure 6).

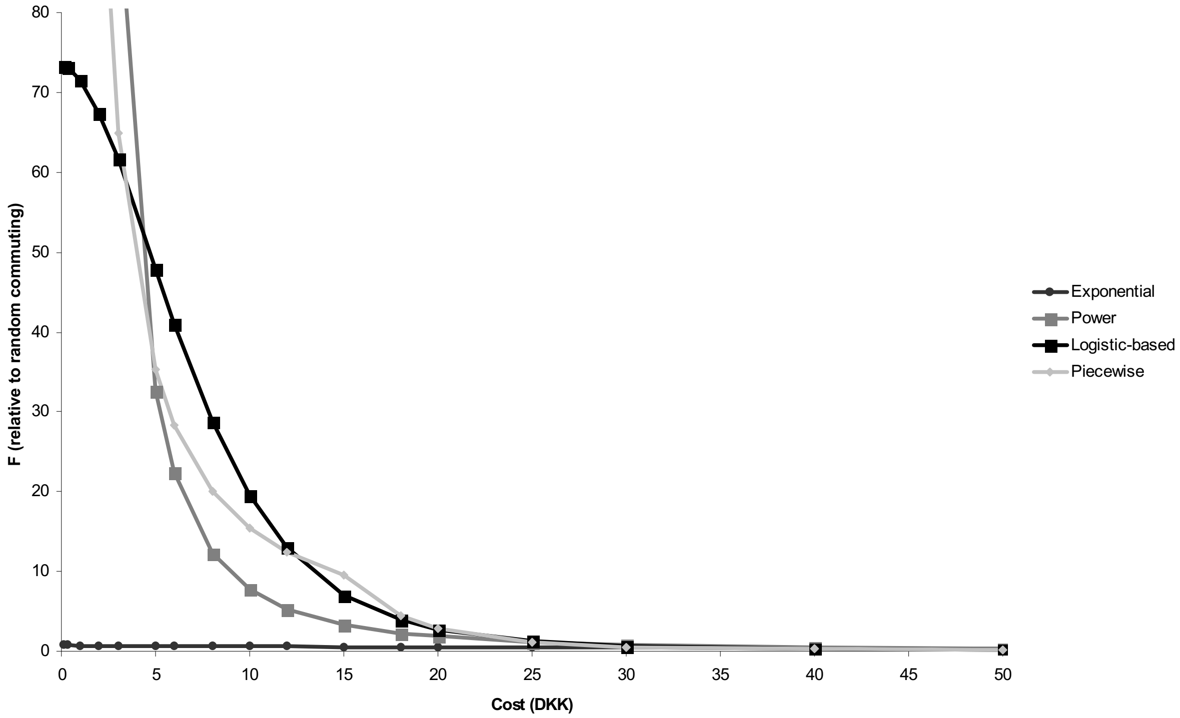

Figure 5. Various decay functions.

and $-\gamma_1\gamma_3/4$= -3.61 (see Table 5 and Appendix B). The following picture of commuting behavior arises. As long as distances are under, say, 10 km, commuters are rather indifferent. In the intermediate range 10 to 60 km, distance is very important for commuting decisions. Long-distance commuters, traveling more than 60 km, do not care very much about the distance.[18]

Several other researchers also found that the sensitivity of commuters with respect to travel cost (or time or distance) is stronger for intermediate distances than for short and long distances. To model this, an S-shaped decay function is required. Hilbers and Verroen (1993) suggest using a log-logistic decay function:

$$F(G_{ij}) = \frac{\gamma_0}{1 + e^{\gamma_1 + \gamma_2 \ln G_{ij}}} \tag{20}$$

[18] The fact that the logistic-based decay function becomes flatter for large distances, and does not approach zero if distance becomes infinite, is partly caused by the type of data we use. The flows are not actual trips, but combinations of residence and work locations. For long distances people will travel less frequently from home to work. For the longest distances, data errors might not be negligible.

Table 4. Value of decay functions (relative to random commuting), for various values of generalized cost

Cost	Exponential	Power	Piecewise	Logistic-based
0	0.72	infinity	infinity	73.29
5	0.66	32.60	35.26	47.95
10	0.60	7.76	15.41	19.53
20	0.50	1.85	2.83	2.74
25	0.46	1.16	1.11	1.23
40	0.35	0.44	0.26	0.25
60	0.25	0.19	0.11	0.09
120	0.08	0.05	0.04	0.04
400	0.00	0.00	0.02	0.03

The difference with (18) is that here on the left-hand-side is the decay function itself, and not its logarithm.[19] This log-logistic decay function is also used by Geurs and Ritsema van Eck (2003) for Dutch commuting. An S-shaped decay function is also used by Johansson et al. (2002a, 2002b) for an empirical analysis of commuting in Sweden. They use different time-sensitivity parameters for local, regional, and extra-regional interaction.

Thorsen et al. (1999) use a logistic function to model distance-deterrence, and provide a theoretical justification for such an S-shaped curve. Their basic idea is that short distances give random commuting flows, whereas long distances are governed by a minimum cost principle (p. 76). They express the trips matrix as a convex combination of these extremes. The weight of the minimum-cost solution should be small for small distances, increase with distance, and approach 1 if distance goes to infinity. A logistic function satisfies this.[20] In

[19] In the literature, (20) is known as log-logistic decay function. To avoid confusion, we can not use the same name for (18). Therefore we use the term "logistic-based decay function".

[20] Despite the similarities there are nevertheless important differences between our model specification and that of Thorsen et al. (1999). We take logs of cost before applying the logistic function. Actually, this satisfies the conditions of Thorsen et al. (1999) better than their own choice, as it results in a weight of 0 for zero cost. Furthermore, their formulation is in trips, ours in log trips. And they use another type of spatial interaction model, where the deterrence function has a different role. The point of similarity is that an S-shaped curve should be used to model distance-decay, and that some transformation of the logistic function is a convenient choice.

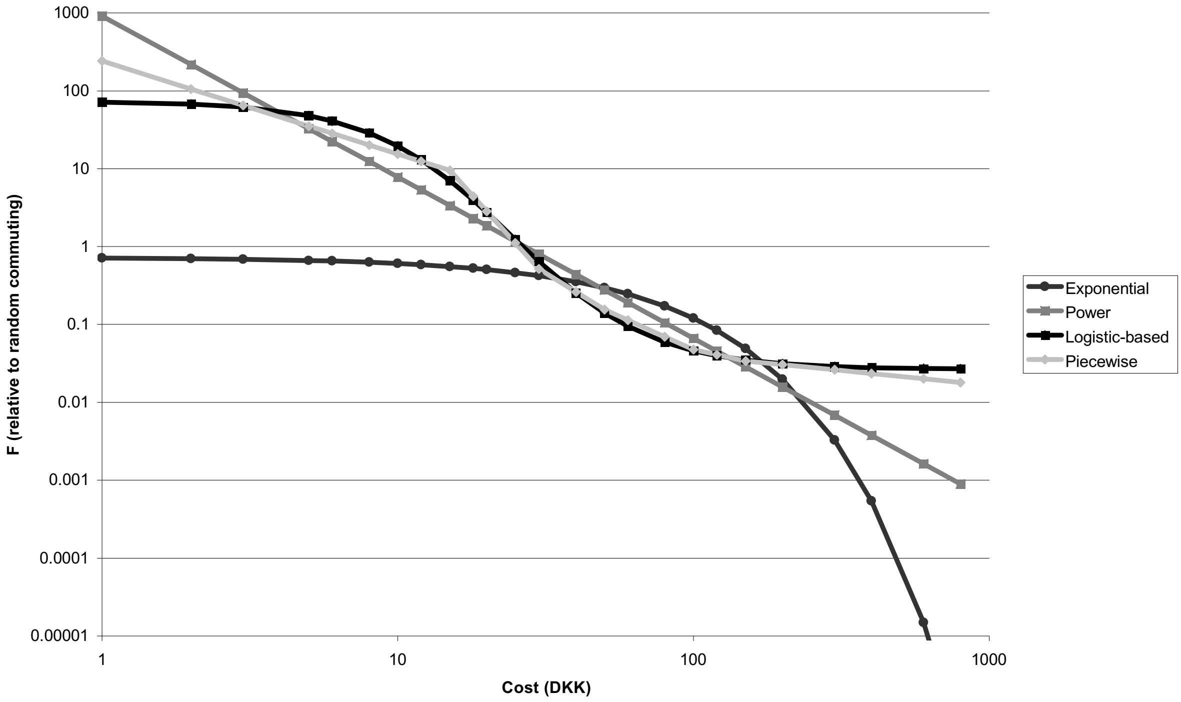

Figure 6. Various decay functions in a log-log setting

Section 2 we mentioned that various forms of heterogeneity could explain deviations from an exponential towards a power form of the decay function. This reasoning might be extended to derive a logistic-based form.

In Appendix B we derive several properties of the logistic-based decay function (18). For short distances, the logistic-based decay function can be approached by an exponential decay function. For intermediate distances, it can be approached by a power decay function. For long distances, the elasticity in the logistic-based decay function approaches zero. The different behavior of the decay function for various cost ranges is relevant for estimation. The various functional forms found in empirical studies can be interpreted as approximations of certain parts of an S-shaped curve. Also, the estimated elasticity in a power specification will depend on the cost range in the data. The treatment of heteroscedasticity affects the weight of various parts of the decay function, as low costs are associated with large flows and high costs with small flows. In the absence of weighting, the many small flows, especially zero flows, will dominate the result. If the weights are based on a Poisson distribution, a few large flows on short distance will dominate. To avoid such problems, we took into account both specification error and the Poisson process.

Table 5. Elasticity of decay functions with respect to generalized cost, for various values of generalized cost

Cost	Exponential	Power	Piecewise	Logistic based
0	0	-2.07	-1.19	0
5	-0.09	-2.07	-1.19	-0.73
10	-0.18	-2.07	-1.19	-2.01
20	-0.36	-2.07	-4.20	-3.51
25	-0.45	-2.07	-4.20	-3.61
40	-0.72	-2.07	-2.36	-2.95
60	-1.08	-2.07	-1.71	-1.94
120	-2.16	-2.07	-0.82	-0.70
400	-7.20	-2.07	-0.38	-0.09

The small elasticities for larger distances in the logistic-based decay function contrast with the power decay function (constant elasticities), and with the exponential decay function (strongly increasing elasticities). This difference is important for the evaluation of major infrastructure projects, such as the fixed link over the Great Belt. Even after the opening of the new link, distances and costs for crossing the Great Belt are relatively high compared with common commuting distances and costs. According to the logistic-based decay function, elasticities are relatively small where distances are long. In these circumstances, the use of a power decay function might lead to overestimation of the effect of the opening of the fixed link. This is illustrated by the following example.[21] Suppose that there is a project that reduces the travel cost between two cities from 120 to 60. (Something like the fixed link over the Great Belt.) As can be seen from Table 4, the power and the logistic-based decay functions have similar values for the old situation, with cost 120. However, for cost 60, the predicted value of the power decay function is twice that of the logistic-based decay function. Hence, the use of a model based on a power decay function can lead to a considerable overestimation of the effect of the new infrastructure on commuting.

[21] In the examples in this section, we only consider the direct effects. The changes in commuting flows will lead to changes on the labor and housing markets, and substitution between commuting relations. These indirect effects depend on the spatial configuration of population and employment. To evaluate these, a complete model is required.

Now suppose we have another project, on a regional level, which reduces travel cost from 25 to 20. Again the power decay function and the logistic-based decay function have similar values for the starting situation. But now the power function predicts an increase of commuting by 60%, while the logistic-based specification predicts an increase by 120%. So, for intermediate distances, the power function will underestimate the rise in commuting. As a third example, we take an overall reduction in generalized cost of 10%.[22] Then the power function predicts an overall increase of commuting by 20% for the direct effect.[23] According to the logistic-based decay function, the increase of commuting primarily occurs on intermediate distances (say between 10 and 60 km), and can be up to 40% for distances between 20 and 30 km.

6. Summary and conclusion

In this paper we have analyzed the effect of transport cost on the size of the commuting flows between the 275 municipalities in Denmark in 1995. We estimated the distance-decay function by NLWLS in logs. The balancing factors were estimated as dummy parameters (Cesario 1974). For computational reasons, we decomposed the variables into averages and deviations from the average (Cesario 1974). To correct for the bias caused by taking logs, we added ½ to each flow (Sen and Soot 1981). Most of the commuting flows are small and there are many zero flows, which are difficult to handle. To correct for heteroscedasticity, we used weights based on a combination of specification error and Poisson error. This specification of weights, which requires iteration, is, to our knowledge, new. We also used the weights in the calculation of averages for the decomposition (De Vries et al. 2002). We omitted the intrazonal flows, because they did not fit the model. We investigated various functional forms and finally chose a logistic-based form for the distance-decay function.

Commuting between municipalities in Denmark can not be described by an exponential distance-decay function, while a power distance-decay function satisfies only on an intermediate cost range. But a description where log trips is a downwards logistic function

[22] Overall changes in generalized travel cost can occur in the course of the years by increasing income (so that monetary cost will be reduced relative to income), changes in fuel taxes or road pricing, and changes in average speed (so that time cost changes).

[23] Of course, with an overall change, indirect effects are essential, and the final increase will be much smaller. Under a (singly or doubly) constrained model with a power decay function, nothing will change in the trips matrix by an overall multiplicative cost change (Evans 1970; Fotheringham and O'Kelly 1989, pp. 11-12).

of log generalized cost fits the data well. The elasticity of commuting with respect to generalized transport cost is (in absolute value) small for small distances, reaches a value of up to 4 for distances around 24 km, and approaches zero for large distances. The logistic-based distance-decay function can be approximated by an exponential decay function for short distances, and by a power decay function for intermediate distances. This is in accordance with the observation of Fotheringham and O'Kelly (1989) cited in Section 1. The logistic-based decay function can unify these approaches in a single model. However, for long distances, the logistic-based decay function differs considerably from those approximations. The elasticity of the size of the commuting flow with respect to generalized cost gradually decreases (in absolute value) if the distance increases.

The specification of the distance-decay function can have important consequences for the evaluation of new infrastructure. We found that commuting is very elastic with respect to costs on intermediate distances, and relatively inelastic for larger distances. The use of a power decay function, which imposes constant elasticity, ignores this difference. This will lead to overestimation of the effect of new long-distance connections on labor markets. On the other hand, the use of a power decay function will lead to underestimation of the effect of regional improvements in infrastructure. For realistic predictions, it is important to take into account that the cost elasticity of commuting varies over distance.

Appendix A. The decomposition

In this Appendix we present the formulas for the decomposition which is used in the estimation method (Section 3). To simplify the notation we use lowercase for logs of variables, and write f_{ij} for $f(G_{ij})$. We can then rewrite (8) as:

$$t_{ij} = a_i + b_j + o_i + d_j + f_{ij} + u_{ij} \tag{A1}$$

where f_{ij} is a submodel which will generally contain multiple explanatory variables and parameters, and can be nonlinear. We estimate (A1) by weighted least-squares (WLS), so we multiply the equation for each observation by weight w_{ij}. The WLS-estimators can be described by first-order conditions. For example, the first-order condition for $\hat{a}_i$ (the parameter of a dummy variable) is:

$$\sum_j (w_{ij}.1).(w_{ij}.\hat{u}_{ij}) = 0 \tag{A2}$$

The estimator can be computed from the weighted averages over j. The weights used in the computation of the weighted averages are the squares of the weights used in the regression. Similarly, $\widehat{b}_j$ can be computed from the weighted averages over i. The decomposition is a computationally-efficient way to estimate the dummy parameters.

The decomposition is as follows, where a dot indicates that a weighted average (with weights w_{ij}^{2}) is taken over the subscript it replaces.:

$$t_{\bullet\bullet} = a_{\bullet} + b_{\bullet} + o_{\bullet} + d_{\bullet} + f_{\bullet\bullet} + u_{\bullet\bullet} \tag{A3}$$

$$t_{i\bullet} - t_{\bullet\bullet} = a_i - a_{\bullet} + o_i - o_{\bullet} + f_{i\bullet} - f_{\bullet\bullet} + u_{i\bullet} - u_{\bullet\bullet} \tag{A4}$$

$$t_{\bullet j} - t_{\bullet\bullet} = b_j - b_{\bullet} + d_j - d_{\bullet} + f_{\bullet j} - f_{\bullet\bullet} + u_{\bullet j} - u_{\bullet\bullet} \tag{A5}$$

$$t_{ij} - t_{i\bullet} - t_{\bullet j} + t_{\bullet\bullet} = f_{ij} - f_{i\bullet} - f_{\bullet j} + f_{\bullet\bullet} + u_{ij} - u_{i\bullet} - u_{\bullet j} + u_{\bullet\bullet} \tag{A6}$$

(A1) is the sum of (A3) to (A6). We use this decomposition to estimate the parameters in the decay function f, as well as the balancing factors a_i and b_j. From the first-order conditions for the WLS estimator of (A1), it follows that the estimator for the parameters in f (other than the constant) is the WLS estimator based on (A6), where the disturbance term can be treated as independent. Using the estimated f, $a_i - a_{\bullet}$ can be estimated from (A4), $b_j - b_{\bullet}$ from (A5), and the constant in f from (A3). The balancing factors are only identified in deviations from their weighted averages, as f contains a constant term. Therefore, we set $a_{\bullet} = 0$ and $b_{\bullet} = 0$ as a normalization.

The use of such a decomposition was proposed by Cesario (1974, p. 252), but, as he used OLS and not WLS, his decomposition is unweighted. It is also described by Sen and Smith (1995, p. 477). However, as we use WLS, all the averages used for the decomposition also need to be weighted (De Vries et al. 2002).

Appendix B. The logistic function

The logistic-based distance-decay function

$$f(G_{ij}) = \gamma_0 + \gamma_1 \frac{1}{1 + \left(G_{ij}/\gamma_2\right)^{\gamma_3}} \tag{B1}$$

has some interesting properties. If generalized cost approaches zero, $G_{ij} \downarrow 0$, then

$$f(G_{ij}) \approx \gamma_0 + \gamma_1 \left[1 - \left(G_{ij}/\gamma_2\right)^{\gamma_3}\right] = \gamma_0 + \gamma_1 - \gamma_1 \left(G_{ij}/\gamma_2\right)^{\gamma_3} \tag{B2}$$

Now if the steepness parameter $\gamma_3 = 1$, we have:

$$f(G_{ij}) \approx \gamma_0 + \gamma_1 - \frac{\gamma_1}{\gamma_2} G_{ij} \tag{B3}$$

This is an exponential distance-decay function (compare (15), with a parameter transformation). So, if the steepness parameter equals 1, the logistic-based decay function approaches an exponential decay function for low cost. (As the estimate for the steepness parameter is 1.8, this result does not hold for the fitted logistic-based decay function.) Around the bending point the logistic function can be approximated by a linear function (in

$\ln G_{ij}$), so for intermediate distances the logistic-based decay function can be approximated by a power decay function. If generalized cost becomes very large, $G_{ij} \rightarrow \infty$, then:

$$f(G_{ij}) \approx \gamma_0 + \gamma_1 \left(G_{ij} / \gamma_2 \right)^{-\gamma_3} \tag{B4}$$

and the elasticity

$$\frac{df(G_{ij})}{d \ln G_{ij}} \approx -\gamma_3 \gamma_1 \left(G_{ij} / \gamma_2 \right)^{-\gamma_3} \tag{B5}$$

So, if generalized cost rises by a certain factor, the elasticity decreases by a given power of that factor. If the steepness parameter equals 1, these factors are the same. This is the property that we noted in Section 4 in the estimated elasticities for the piecewise power function: the slope halves for equal increases in $\ln G_{ij}$.

The elasticity of trips with respect to generalized cost is:

$$\frac{df(G_{ij})}{d \ln G_{ij}} = -\gamma_1 \gamma_3 \frac{1}{1 + \left(G_{ij} / \gamma_2 \right)^{\gamma_3}} \left[1 - \frac{1}{1 + \left(G_{ij} / \gamma_2 \right)^{\gamma_3}} \right] \tag{B6}$$

At $G_{ij} = 0$, the elasticity equals zero. The largest elasticity (in absolute value) is attained for $G_{ij} = \gamma_2$, and equals $-\frac{\gamma_1 \gamma_3}{4}$. For $G_{ij} \rightarrow \infty$, the elasticity approaches zero.

References

Alonso, W. (1978). “A Theory of Movements.” In: *Human Settlement Systems: International Perspectives on Structure, Change and Public Policy*, ed. N. M. Hansen, Ballinger Publishing Company, Cambridge, Massachusetts, pp. 197-211.

Andersen, A. K. (1999). *Location and Commuting*, unpublished PhD-thesis, University of Copenhagen.

Batten, D. F. and Boyce, D. E. (1986). “Spatial Interaction, Transportation, and Interregional Commodity Flow Models.” In: *Handbook of Regional and Urban Economics*, ed. P. Nijkamp, North-Holland, Amsterdam, pp. 357-406.

Bikker, J. A. and De Vos, A. F. (1992). “A regional supply and demand model for inpatient hospital care.” *Environment and Planning A*, *24*(8), 1097-1116.

Bikker, J. A. (1987). “An International Trade Flow Model with Substitution: An Extension of the Gravity Model.” *Kyklos*, *40*(Fasc. 3), 315-337.

Carey, H. C. (1858). *Principles of social science*, Lippincott, Philadelphia.

Cesario, F. J. (1974). “The interpretation and calculation of gravity model zone-to-zone adjustment factors.” *Environment and Planning A*, *6*, 247-257.

Choukroun, J.-M. (1975). “A general framework for the development of gravity-type trip distribution models.” *Regional Science and Urban Economics*, *5*, 177-202.

Cochrane, R. A. (1975). “A possible economic basis for the gravity model.” *Journal of Transport Economics and Policy*, *9*(1), 34-49.

De Vries, J. J., Nijkamp, P., and Rietveld, P. (2001). "Alonso's Theory of Movements: Developments in Spatial Interaction Modeling." *Journal of Geographical Systems*, *3*(3), 233-256.

De Vries, J. J., Nijkamp, P., and Rietveld, P. (2002). "Estimation of Alonso's Theory of Movements by Means of Instrumental Variables." *Networks and Spatial Economics*, *2*(2), 107-126.

Doornik, J. A. (1998). *Ox, an object-oriented matrix programming language*, Timberlake Consultants, West Wickham, Kent.

Evans, A. W. (1970). "Some properties of trip distribution models." *Transportation Research*, *4*, 19-36.

Fotheringham, A. S. and O'Kelly, M. E. (1989). *Spatial Interaction Models: Formulations and Applications*, Kluwer, Dordrecht.

Geurs, K. T. and Ritsema van Eck, J. R. (2003). "Evaluation of accessibility impacts of land-use scenarios: the implications of job competition, land-use, and infrastructure developments for the Netherlands." *Environment and Planning B: Planning and Design*, *30*, 69-87.

Hamerslag, R. (1980). "Spatial development, developments in traffic and transportation, and changes in the transportation system." In: *Changes in the field of transport studies*, eds. J. B. Polak and J. B. van der Kamp, Martinus Nijhoff, The Hague, pp. 60-80.

Hilbers, H. D. and Verroen, E. J. (1993). "Het beoordelen van de bereikbaarheid van lokaties" ("The evaluation of the accessibility of locations"; in Dutch). Rep. No. INRO-VVG 1993-09, TNO, Delft.

Hua, C.-I. (1980). "An exploration of the nature and rationale of a systemic model." *Environment & Planning A*, *12*(6), 713-726.

Johansson, B., Klaesson, J., and Olsson, M. (2002a). "On the Non-Linearity of the Willingness to Commute." Paper presented at the 42nd ERSA Congress, Dortmund.

Johansson, B., Klaesson, J., and Olsson, M. (2002b). "Time distances and labor market integration." *Papers in Regional Science*, *81*(3), 305-327.

Ledent, J. (1980). "Calibrating Alonso's general theory of movement: the case of interprovincial migration flows in Canada." *Sistemi Urbani*, *2*, 327-358.

Nijkamp, P., and Rouwendal, J., eds. (2004). Special issue on Advances in Commuting Studies. *Growth and Change 35(*3).

Poot, J. (1986). "A system approach to modelling the inter-urban exchange of workers in New Zealand." *Scottish Journal of Political Economy*, *33*(3), 249-274.

Rouwendal, J. and Nijkamp, P. (2004). "Living in Two Worlds: A Review of Home-to-Work Decisions." *Growth and Change*, *35*(3), 287-303.

Sen, A. and Smith, T. E. (1995). *Gravity Models of Spatial Interaction Behavior*, Springer, Berlin.

Sen, A. and Soot, S. (1981). "Selected procedures for calibrating the generalized gravity model." *Papers of the Regional Science Association*, *48*, 165-176.

Stewart, J. Q. (1941). "An Inverse Distance Variation for Certain Social Influences." *Science*, *93*, 89-90.

Stoer, J. and Bulirsch, R. (1980). *Introduction to Numerical Analysis*, Springer, New York.

Thorsen, I., Uboe, J., and Naevdal, G. (1999). "A network approach to commuting." *Journal of Regional Science*, *39*(1), 73-101.

Wilson, A. G. (1971). "A family of spatial interaction models, and associated developments." *Environment and Planning*, *3*, 1-32.