Commuting and panel spatial interaction models: evidence of variation of the distance-effect over time and space

Wouter Vermeulen*), **)
Jos Van Ommeren**)
Jaap de Vries**)

16-07-2003
WORK IN PROGRESS
PLEASE DO NOT QUOTE

*) : CPB, Netherlands Bureau for Economic Policy Analysis
**) : Free University, FEWEB, Amsterdam, De Boelelaan, 1108 HV Amsterdam
Email: w.vermeulen@cpb.nl, jommeren@feweb.vu.nl

Abstract: We apply spatial interaction models using panel data to explain commuting behaviour in the Netherlands. Our main conclusion is that the distance-decay effect is not constant over time and that changes in this effect are region specific. In more densely populated regions the change in the distance-decay parameter is small suggesting that regional increases in congestion have a large negative effect on the increases in average commuting distance. The panel spatial interaction model we derive is well-suited for testing significance of the centrality index (an often used variable in spatial interaction models). Although evidence is found for competition effects in a pooled cross section framework, controlling for time invariant unobserved heterogeneity renders this relation spurious.

1. Introduction

Spatial interaction models, a certain type of gravity models, are popular tools to predict commuting flows between regions (Fotheringham and O' Kelly, 1989). The focus of these models is on the distance-deterrence parameter, which measures, loosely speaking, the effect of the distance between two regions on the size of the commuting flow between these regions (conditional on the characteristics of the region, for example, the number of jobs). Previous studies have estimated the distance-deterrence effect based on cross-section data on commuting flow for a specific short period (usually one year, see Fotheringham and O' Kelly, 1989, for an overview). These studies usually acknowledge that it is open to debate to what extent the estimates can be generalised to other periods. This ambiguity is problematic as spatial interaction models are frequently used to evaluate the effect of new infrastructure projects on future commuting flows for different scenarios. To predict commuting flows in the future would be relatively straightforward if it can be assumed that the distance-deterrence effect is constant over time in the absence of infrastructure improvements. It is implausible however that the distance-deterrence effect is constant over time, because the relative costs associated with the commuting distance are thought to fall over time. The main reason is that as average income grows over time, the costs of commuting relative to wages fall, implying that employees will choose to travel by faster, but more expensive, modes, which increases the average distance travelled (even when the average commuting time remains constant).[1] An increase in the average distance travelled implies an increase in congestion, which may weaken the original effect. Because congestion tends to be a local phenomenon, it is generally expected that the time-variation in the distance-decay parameter is locally specific.

Recently, Thorsen and Gitlesen (1998) have empirically evaluated alternative model specifications to predict commuting flows. Their main conclusion is that spatial interaction models are sensitive to the chosen specification and potentially misspecified due to measurement errors in the distance function. Estimates of the distance-deterrence parameter appear not to be independent of the chosen model specification.[2] In the current paper, we will estimate the time-variation in the distance deterrence effect on commuting flows using panel data. By employing panel data, we are able to address both the specification issue and the problems associated with measurement errors. Surprisingly, the use of panel data in the current context is novel.[3] Panel data estimation turns out to be extremely straightforward.[4]

The benefits of using panel data have been extensively discussed (Hsiao, 1985; Baltagi, 2002). We will see that in the context of spatial interaction modelling, the main advantage is that one may control for origin-destination specific heterogeneity. Common sense suggests that any variable that measures the economic distance between regions fails to capture the heterogeneity of the economic distances. For example, when economic distance is measured by the geographical distance between the centres of regions, then this measure not only ignores the heterogeneity due to variation in infrastructure, but fundamentally ignores the variation in the specific spatial form of both regions including the distribution of jobs and residences within the regions. As is well known, omission of heterogeneity leads to bias in the resulting estimates if the omitted

[1] Another reason may be that the population density increases which may increase the costs per commuting distance due to increased congestion.
[2] Similarly, in the empirical literature on migration and competing destinations, which is based on spatial interaction models, it is generally reported that the estimates of the distance deterrence effects depends upon the chosen functional specifications (in particular, the inclusion of the competing destination parameter).
[3] Panel data applications of gravity models are common in the international trade literature (Brun et al., 2002).
[4] While interpretation of the results is less ambiguous than estimates based on cross-section data.

variable correlates with the explanatory variables. Panel data estimation controls fully for time-invariant heterogeneity.[5]

2. Panel data and spatial interaction models

2.1 Spatial interaction models

A common application of spatial interaction models in the field of commuting and infrastructure evaluation is the following doubly constrained gravity model (Fotheringham and O'Kelly, 1989), which will be the focus of our paper:

$$P_{ij} = A_i O_i B_j D_j F(d_{ij}) u_{ij}, \tag{1}$$

where P_{ij} denotes the number of commuters between region i and j, O_i denotes the size of the labour force in region i (origin), D_j denotes the number of employed workers in region j (destination) and $F(d_{ij})$ denotes the distance-decay, where F ($F > 0$) is assumed to be a decreasing function of the distance d_{ij} between the regions i and j. A_i and B_j are 'balancing factors', which guarantee that the origin and distance totals are constrained, so $\sum_i P_{ij} = D_j$ and $\sum_j P_{ij} = O_i$.[6] Finally, u_{ij} denotes the random error with u_{ij} independent and identically distributed. In empirical applications, $F(d_{ij})$ is usually specified as $\exp(\alpha d_{ij})$ or d_{ij}^{β} ($\alpha, \beta < 0$). In the following, we will assume that $F(d_{ij}) = d_{ij}^{\beta}$, but all the results can easily be adapted presuming different functional forms of F. So:

$$P_{ij} = A_i O_i B_j D_j d_{ij}^{\beta} u_{ij}. \tag{2}$$

In the empirical literature, the first aim is to estimate β, the distance-decay parameter, which determines how the number of commutes depend on commuting distance. The main underlying assumption of this model is that P_{ij} depends on factors related to region i

[5] In the context of commuting flows, the disadvantages of panel data are minimal, because the usual problems of panel data are related to non-response, attrition and self-selectivity (Kasprzyk et al., 1989) are absent. The main restriction is that the method of collecting data over time remains the same.

[6] Thorsen and Gitlesen (1998) extend the above model by including an effect of labour market characteristics of P_{ii}. Our estimation approach is insensitive to such an extension.

(O_i and A_i), factors related to region j (D_j and B_j) and depends on factors which are related to both region i and region j *only through the commuting distance d_{ij}*. Although such an assumption may be correct for some applications, it is plausible that other factors then d_{ij}, let's call them c_{ij}, influence P_{ij}. One example in the literature is that c_{ij} is a centrality index, which measures the competition from other regions or a contiguity variable, which measures if regions i and j are contiguous (Fotheringham and O'Kelly, 1989, chapter 3). Hence, a more general formulation of the spatial interaction model is:

$$P_{ij} = A_i O_i B_j D_j d_{ij}{}^{\beta} c_{ij}{}^{\theta} u_{ij}, \qquad i,j = 1..N. \qquad (3)$$

The main empirical problem is that estimates of β depend on the correct specification of c_{ij}, which is often problematic. This issue can be avoided by means of panel data.

A more general spatial interaction model is Alonso's Theory of Movements. (Alonso, 1978; Fotheringham and O'Kelly, 1989). In this model the origin and destination totals are not constrained, but dependent on the balancing factors. For the commuting application this means that employment in each region is affected by accessibility to the labor force, and active population is affected by accessibility to jobs. Estimation of Alonso's Theory of Movements falls apart into two stages (De Vries et al., 2002) The estimation of the distance-deterrence function is exactly the same as in the doubly constrained model. Estimation of the effect of accessibility on location is more complicated. As in this paper, we are only concerned with the effect of distance, the results are also valid for Alonso's Theory of Movements. The same holds true for special cases of this model, such as singly constrained models. Estimation of the distance-decay parameter is the same for all these models.

Sen and Soot (1981) propose three methods to estimate β. The first method involves maximum likelihood, the other two methods involve a linearisation of (3) such that β can be estimated in a less cumbersome way. For simplicity of exposition, we will ignore the function c_{ij}. The first of these linearisation methods implies that (3) is written as:

$$\ln(P_{ij}) + \ln(P_{ji}) - \ln(P_{ii}) - \ln(P_{jj}) = \beta(\ln(d_{ij}) + \ln(d_{ji}) - \ln(d_{ii}) - \ln(d_{jj})) + \varepsilon_{ij} + \varepsilon_{ji} - \varepsilon_{ii} - \varepsilon_{jj}$$

where ε_{ij} is independent and identically distributed IID(0, σ^2).

2.2 Spatial interaction models and panel data

A more general formulation of the model, which allows for variation over time in the commuting flows is:

$$P_{ijt} = A_{it} O_{it} B_{jt} D_{jt}^{\varphi} d_{ij}{}^{\beta_{ijt}} c_{ijt}{}^{\theta} c_{ij0}{}^{\theta_0} u_{ijt}\,, \qquad \text{i,j = 1,..N} \quad \text{j = 1,..T,} \tag{4}$$

where we acknowledge that β_{ijt} may vary over time and may be origin and destination specific. The latter may be important because given the distance, the economic costs may be origin and distance specific, for example due to local differences in infrastructure. Moreover, we recognise that the factor c_{ij} can be decomposed into a time varying factor c_{ijt} and a time-constant factor c_{ij0}. The latter factor mainly includes variables that are related to *observed* spatial particularities (e.g. contiguity). Further, we allow the effects of O_{it} and D_{jt} on the commuting flows to depend on parameters α and φ. In the empirical application, the research will focus on the change in β_{ijt}, whereas α, φ, θ, and θ_0 will be a nuisance parameters of less interest. We emphasise that the current specification of the model is extremely general. For feasible estimation, we will put restrictions on the functional form of β_{ijt}. We will first assume that β_{ijt} obtains the following particular functional form:

$$\beta_{ijt} = \beta_{ij} + \beta_t\,, \tag{5}$$

implying, that the change over time in β_{ijt} is not origin/distance specific and therefore the same for all commuting flows:

$$\beta_{ijt} - \beta_{ijt-1} = \beta_t - \beta_{t-1} = \Delta\beta_t\,. \tag{6}$$

In the empirical estimation, we will estimate $\Delta\beta_t$, the change in the distance-deterrence parameter. The above specification presumes that the change in the distance-deterrence function is not region specific, which may be unrealistic, because it does not allow for local changes in the distance-deterrence effect, e.g. due to new infrastructure or increased congestion. To allow for region-specific effects, the following less restrictive functional form may be more appropriate:

$$\beta_{ijt} = \beta_{ij} + \beta_{it} . \qquad (7)$$

This specification is more general than (5). Equation (7) presumes that the change in β_{it} is origin specific. In this case:

$$\beta_{ijt} - \beta_{ijt-1} = \beta_{it} - \beta_{it-1} = \Delta\beta_{it} . \qquad (8)$$

In the empirical analysis we will estimate $\Delta\beta_t$ (based on (4) and (5)) and $\Delta\beta_{it}$ (based on (4) and (8)). We will test whether $\Delta\beta_{it} = \Delta\beta_t$.

Without loss of generality, we can structure u_{ijt} in the following way:

$$u_{ijt} = v_{ij} . v_{it} . v_{jt} . v_{ijt} , \qquad (9)$$

where v_{ij}, v_{it} and v_{jt} are *unobserved* variables, and v_{ijt} is an *unobserved random* variable which is independent and identically distributed. The explanatory variables in (4) are assumed to be independent of v_{ijt}. An example of v_{ij} is the time-invarying *unobserved* measurement error due to spatial particularities (e.g. the spatial forms of regions i and j, the presence of natural barriers between i and j) and the *unobserved* measurement error in the costs associated with distance (e.g. the presence of specific types of infrastructure). Note that v_{it} reflects an unobserved time-varying deviation in the flows originating from region i, for example due to infrastructure improvements in region i. The variable v_{jt} has a similar interpretation.

2.3 Fixed or random effects?

In the panel data literature, there is a large literature on the assumptions of the type of unobserved variables v_{ij}, v_{it} and v_{jt} (Baltagi, 2001). These variables could either be assumed to be random or assumed to fixed parameters to be estimated. In the context of commuting flows, it makes sense to assume that v_{ij}, v_{it} and v_{jt} are fixed, because interference is based on a *specific* set of flows between regions (which cannot be interpreted as a random drawing from a large population of flows). One advantage of the fixed effect assumption is that the explanatory variables are allowed to be correlated to unobserved fixed variables. A disadvantage is that effects of time invariant variables (distance) are not identified.

2.4 Estimation

After taking the logarithm of both sides of (4), one can in principle estimate the model by means of ordinary least squares (OLS) to get estimates of $\Delta\beta_t$, α, φ, θ, θ_0, v_{ij}, v_{it} and v_{jt}. However, if N or T is large, estimation will involve too many individual dummy variables (v_{ij}, v_{it} and v_{jt} already involve $N^2 + 2NT$ dummies; in our application this would mean 2400 dummies), and the matrix to be invented by OLS is usually too large. We propose here a specific solution which encompasses estimation methods applied in cross-section spatial interaction models (Sen and Soot, 1981) and panel data models (Hsiao, 1985). This method is to write (4), using odds ratios as:

$$\frac{(P_{ijt}/P_{iit})(P_{jit}/P_{jjt})}{(P_{ijt-1}/P_{iit-1})(P_{jit-1}/P_{jjt-1})} = (d_{ij}{}^{\beta_{ijt}-\beta_{ijt-1}} / d_{ii}{}^{\beta_{iit}-\beta_{iit-1}})(d_{ji}{}^{\beta_{jit}-\beta_{jit-1}} / d_{jj}{}^{\beta_{jjt}-\beta_{jjt-1}}).$$

$$\left(\frac{\dfrac{c_{ijt}}{c_{iit}}\dfrac{c_{jit}}{c_{jjt}}}{\dfrac{c_{ijt-1}}{c_{iit-1}}\dfrac{c_{jit-1}}{c_{jjt-1}}}\right)^{\theta} \frac{\dfrac{v_{ijt}}{v_{iit}}\dfrac{v_{jit}}{v_{jjt}}}{\dfrac{v_{ijt-1}}{v_{iit-1}}\dfrac{v_{jit-1}}{v_{jjt-1}}} \qquad (10)$$

Maybe rather surprisingly, equation (10) demonstrates that the change in the commuting flows between i and j (relative to the internal commuting flow for i) relative to the return flow from j to i (relative to the internal flow for j) does *not* depend on nuisance parameters (v_{ij}, v_{it}, v_{jt}), the origin and distance size effects and does not depend on any observed (or unobserved) spatial particularity.

Hence, defining Δ as a change over time such that $\Delta x_t = x_t - x_{t-1}$ taking logarithms, and making use of (6), it appears that:

$$\Delta\ln(P_{ijt}) - \Delta\ln(P_{iit}) + \Delta\ln(P_{jit}) - \Delta\ln(P_{jjt}) = \Delta\beta_t[\ln(d_{ij}) - \ln(d_{ii}) + \ln(d_{ji}) - \ln(d_{jj})] + \theta[\Delta\ln(c_{ijt}) - \Delta\ln(c_{iit}) + \Delta\ln(c_{jit}) - \Delta\ln(c_{jjt})] + \text{random error.}^{7} \quad (11)$$

Readers familiar with the panel data literature will realise that although estimates of ΔBt obtained based on (10) are consistent, one can easily obtain more efficient estimators. Equation (10) has been based on the change in the commuting flow between two periods, but it can easily be seen that it is more efficient to focus on the change in the commuting flow compared to the *average commuting flow* over the whole period (since the variation in the average flow is less than the variation in the flow from one year).

One can see that the time-invariant variables that are associated with time-invariant coefficients are not identified and do not affect $\Delta\beta_t$. The time-varying factor c_{ijt} can be measured in several ways but it is common to specify c_{ijt} as c_{jt} (or c_{it}) see Fotheringham, 1983; 1986; Fotheringham et al., 2001; Pellegrini and Fotheringham, 1999; Ishikawa, 1987). For example, it may refer to the average education of the labour force in a region. In this case, using equation (11) simplifies into:

$$\Delta\ln(P_{ijt}) - \Delta\ln(P_{iit}) + \Delta\ln(P_{jit}) - \Delta\ln(P_{jjt}) = \Delta\beta_t[\ln(d_{ij}) - \ln(d_{ii}) + \ln(d_{ji}) - \ln(d_{jj})] + \text{random error.} \quad (12)$$

So estimation of $\Delta\beta_t$ is not affected by the c_{jt} (or c_{it}). Based on equation (12), $\Delta\beta_t$ can be estimated by means of OLS. In a similar way, $\Delta\beta_{it}$ can be estimated.

[7] Presuming that $F_t(d_{ij}) = \exp(\alpha_t d_{ij})$ and the factor c_{ijt} enters also exponentially as $\exp^{\theta c_{ijt}}$, it appears that we obtain the same equation as above, the only difference being that the first term on the right side is replaced by $\Delta\alpha_t[d_{ij} - d_{ii} + d_{ji} - d_{jj}]$ and the second by $\theta[\Delta c_{ijt} - \Delta c_{iit} + \Delta c_{jit} - \Delta c_{jjt}]$.

Alternatively, if the choice set of destinations is not constant over regions (Fotheringham and O'Kelly, 1989; Thorsen and Gitlesen, 1998): c_{ijt} may be specified as follows:

$$c_{ijt} = \sum_k D_{kt} d_{ik}^{\gamma}, \qquad \text{where } k \neq i \text{ and } k \neq j, \qquad \gamma < 0$$

3. Commuting in the Netherlands 1992 – 2001

3.1 Description of the data

The commuting flow data we use come from ten sequential labour force surveys (1992 to 2001), which contain each about one percent of Dutch households. The locations of residence and workplace of each employee are both known. We have calculated regional commuting flows for 40 (COROP) regions. Each region contains, on average, about 160,000 employees. In 1992, 83 % of the employees live and work in the same region. In 2001, 78 % of the employees live and work in the same region. So, in the Netherlands during the nineties, the population of employees which work in the region of residence has decreased substantially.

The measurement of distance is usually a sensitive issue, as spatial interaction models are sensitive to the measurement error in the distance function (Thorsen and Gitlesen, 1998). One of the main advantages of panel data analysis is that the consistency of the estimates is *not* affected by time-invariant measurement error (as demonstrated in (10) because v_{ij} is not identified).

In the current application, we have used the average commuting distance by car in 1995, which overestimates the average commuting distance for most commuting flows.[8] Although the measurement error is systematic, it is time invariant, and will therefore *not* affect the consistency of the estimates.

As a preliminary exercise to estimating the panel data model based on (12), we have estimated a spatial interaction model employing a cross-section analysis based on equation (3).

So, we have estimated ten times the distance-decay parameter (and not the change in this parameter as in the panel data analysis), which requires us to specify the spatial

[8] We would like to thank AVV Transport Research Center for providing these data.

particularities of the region. In this analysis, we included dummies for adjacent regions and for commuting flows to the region of residence. Further, we used weights as proposed by Sen and Soot (1981). Then 10 distance parameter estimates are plotted in Figure 1. The full results can be received upon request from the authors.

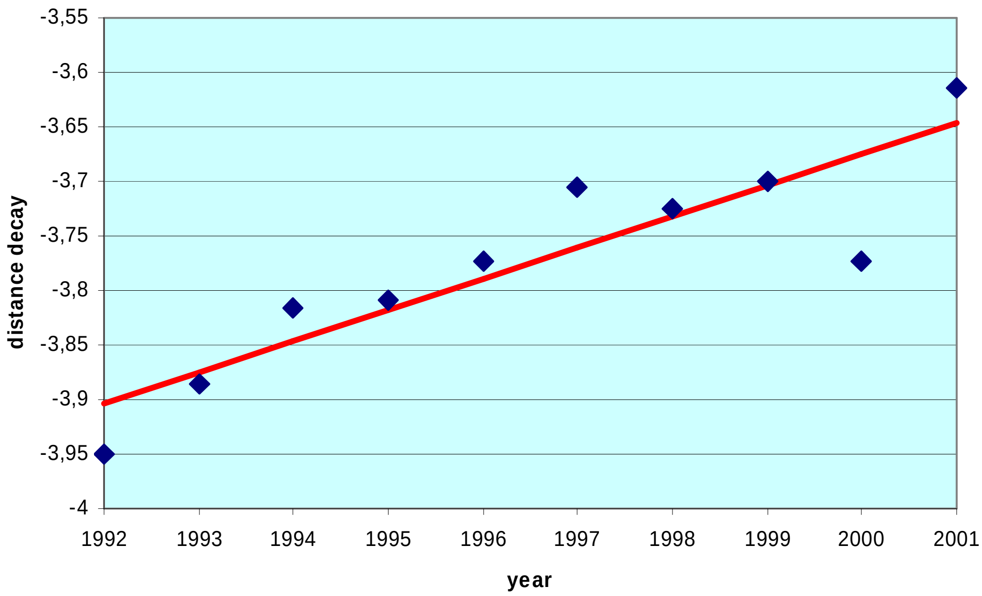

Figure 1: distance decay over time

An ordinary least squares regression on these estimates yields that the distance decay parameter increased with 0.029 each year.

3.2 Estimation results

In this section we estimate a trend in the distance decay parameter on Dutch commuting data for the period 1992 – 2001. To this aim, we derived various econometric models in section 2. The basic model is given in equation (2) and we use the first of the two linearizations given.[9] [10] We apply OLS, fixed effects and random effects estimators. Weights as proposed by Sen and Soot (1981) are used (averaged over time), reflecting the fact that large flows are measured more accurately. Results are shown in table 1 (standard errors between brackets).

	pooled OLS	panel estimators	
		FE	RE
distance decay	-3.908 (0.032)	-	-3.908 (0.056)
trend in dist. decay	0.0295 (0.0058)	0.0295 (0.0024)	0.0295 (0.0024)
centrality index	-0.133 (0.064)	0.031 (1.070)	-0.128 (0.182)

Table 1: estimation results

In the first specification, where no individual effects are allowed for, effects of distance and a trend in this effect appear to be highly significant. The hypothesis that the distance decay parameter does not vary over time is thus rejected against a positive trend. We find a negative coefficient for the centrality index, which is significant at the five percent level. This might indicate competition or congestion effects.

The second and third specification control for unobserved heterogeneity by allowing for individual effects. In the fixed effects specification, the distance decay parameter is not identified since this estimator is based on variation over time and not over individuals (space). The trend in distance decay estimate does not change compared to the OLS specification, but it is more efficient. However, the centrality index now turns out to be

[9] Before taking logs, we have added 1 to all flows.

insignificant. Apparently, the relation found in the OLS specification was spurious, and due to unobserved heterogeneity.

The random effects model assumes that unobserved heterogeneity is independent from explanatory variables, in our case distance. This seems a reasonable assumption. Estimates are then obtained from an optimal combination of time series and cross section information, so that the effects of time invariant variables like distance are also identified. Again we find the same coefficients for distance decay and trend as in the OLS specification, but the random effects estimator is more efficient. Just like in the fixed effects specification, the centrality index is insignificant.

3.3 Regional variation

We finally consider region specific distance decay coefficients and trends in the random effects model. Hypotheses that regional differences in these variables are statistically insignificant are strongly rejected. Figure 1 shows regional distance decay parameters in a map of The Netherlands. Commuting distances are relatively large in the west of the country, where population and economic activity are concentrated, and in the province of Groningen. In figure 2 we present a map of regional trends in distance decay. Again, regional differences are substantial. The increase in average commuting distances is smallest in the west of the country. In the introduction to this paper we have argued that average commuting distance should increase over time. Since roads in this region are often congested, a marginal increase in distance would come at a higher price.

---

[10] For relyability of the data, we consider commuting flows over a distance smaller than 100 km only. This leaves us with 494 of the 1600 possible flows.

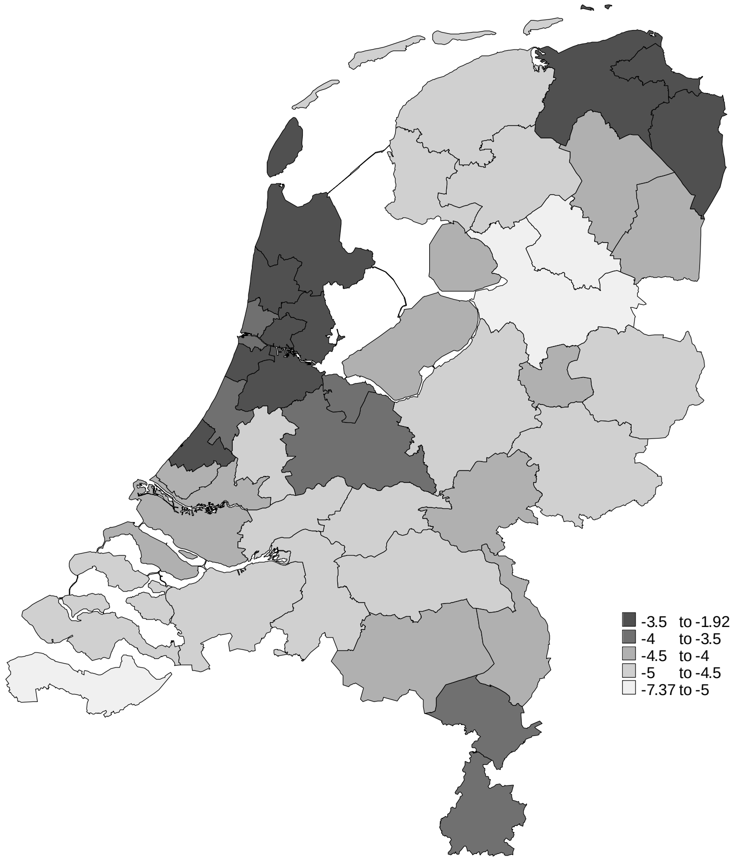

Figure 1: regional distance decay

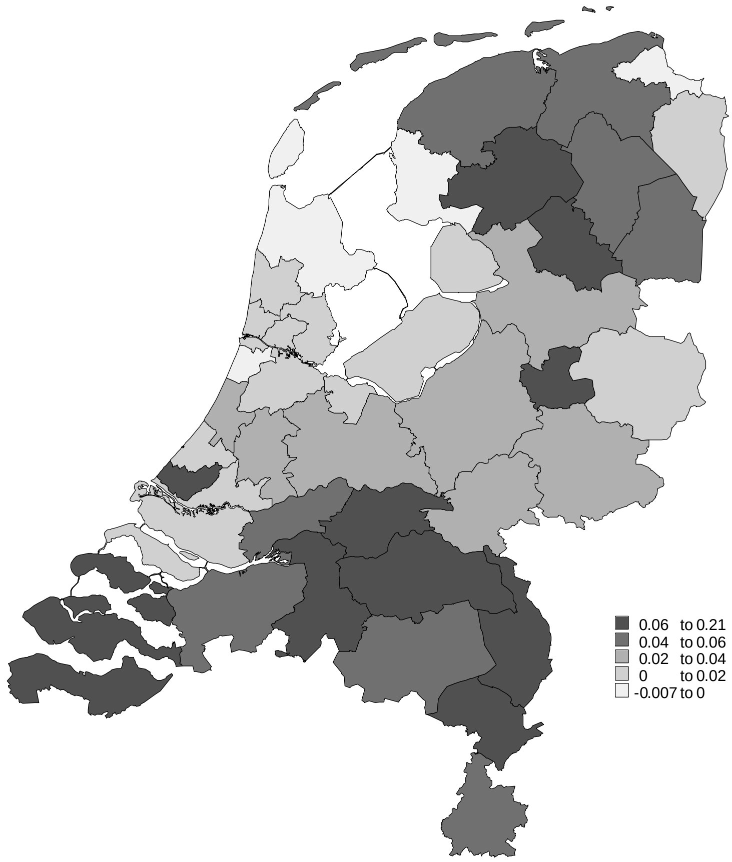

Figure 2: regional trends in distance decay

It appears that growth in the distance decay parameter was smaller in the west and centre of the country than it was in the north and south. A potential explanation would be that traffic congestion is considerable in the former regions. In the introduction to this paper we have argued that average commuting distance should increase over time. In congested areas a marginal increase of the commuting distance comes at a higher price.

4 Conclusions

This paper has proposed a spatial interaction model framework for estimating interregional commuting panel data. A central question was whether the distance decay parameter is constant over time. This question is of major importance for the analysis of infrastructure projects. A main finding of our empirical research is that a significant trend in the distance decay parameter exists, people indeed commute over increasing distances.

This finding is consistent with several micro analyses (eg. Rouwendal and Rietveld, 1994), but our results are established using data on aggregate flows. Also, we show that trends in the distance deterrence parameter vary over regions.

A major advantage of using panel data is the correction for possible omitted variable biases. In a regional context, biases could stem from measurement errors in the distance matrix or spatial particularities within or between regions. Since distance and most of these particularities can be considered time invariant, they do not affect a fixed effects estimator. The panel spatial interaction model we propose is thus very suitable for testing the impact of a centrality index. The relation be found in cross section analyses turns out to be spurious when we correct for unobserved (inter)regional heterogeneity.

References

Alonso, W. (1978). "A theory of movements." Human settlement systems: International perspectives on structure, change and public policy, N. M. Hansen, ed., Ballinger Publishing Company, Cambridge, Massachusetts, 197-211.

Baltagi, B.H. (2001), *Econometric analysis of panel data*, second edition, John Wiley, Chichester

Brun, J.F., C. Carrere, P. Guillaumont and J. de Melo (2002), Has distance died? Evidence from a panel gravity model, CEPR discussion paper 3500, London, UK

Desta, E. and B. Pigozzi (1991), Further experiments with spatial structure measures in gravity models, *Tijdschrift voor Economische en Sociale Geografie*, 82, 220-226

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.

Fotheringham, A.S. (1983), A new set of spatial interaction models: the theory of competing destinations, *Environment and Planning*, A, 15, 15-36

Fotheringham, A.S. (1984), Spatial flows and spatial patterns, *Environment and Planning*, A, 16, 529-543

Fotheringham, A.S. (1986), Modelling hierarchical destination choice, *Environment and Planning*, A, 18, 401-418

Fotheringham, A.S. and M.E. O'Kelly (1989), Spatial interaction models: formulations and applications, *Studies in Operational Regional Science*, Kluwer, Dordrecht

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

Fotheringham, A.S., T. Nakaya, K. Yano, S. Openshaw and Y. Ishikawa (2001), Hierarchical destination choice and spatial interaction modelling: a simulation experiments, *Environment and Planning* A, 33, 5, 901-920

Hsiao, C. (1985), Analysis of panel data, Econometric Society, Monographs, 11

Ishikawa, Y (1987), An empirical study of the competing destinations model using Japanese interaction data, *Environment and Planning*, A, 19, 1359-1373

Kasprzyk, D., G.J. Duncan, G. Kalton and M.P. Singh (1989), *Panel surveys*, John Wiley, New York

Kau, J.B. and C.F. Sirmans (1979), A recursive model of the spatial allocation of immigrants, *Journal of Regional Science*, 19, 47-56

Lo, L. (1991), Substitutability, spatial structure and spatial interaction, *Geographical Analysis*, 23, 132-146

Pellegrini, P.A. and Fotheringham, A.S. (1999), Intermetropolitan migration and hierarchical destination choice: a disagregate analysis analysis from the US PUMS, *Environment and Planning*, A, 31, 1093-1118

Sa, C., R. Florax and P. Rietveld (2002), Determinants of university choice: a gravity approach, mimeo

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

Thorsen, I. and J.P. Gitlesen (1998), Empirical evaluation of alternative model specifications to predict commuting flows, *Journal of Regional Science*, 38, 2, 273-292