SERIE RESEARCHMEMORANDA

VRIJE UNIVERSITEIT
EKONOMISCHE FAKULTEIT
AMSTERDAM

Sensitivity of Information for the Aggregation Level of Spatial Data

Peter Nijkamp
Piet Rietveld
Annemarie Rima.

Researchmemorandum 1983-15
oktober, 1983.

Free University
Amsterdam, April 1983

Peter Nijkamp
Piet Rietveld
Annemarie Rima

1. INTRODUCTION

An information system (monitoring system, computer graphic system, regional model, etc) contains a set of organized data so as to improve the quality of decision-making. Structuring of data involves a transformation of original specific data into categorized data encompassing general patterns and trends. If original data is regarded as a numerical attribute to a certain phenomenon, one may use the following descriptive symbol :

$$x_{i,r,t} ,$$

where $x_{i,r,t}$ represents a numerical value (on a certain measurement scale) assigned to phenomenon i in region r in period t.

In the framework of the present study, we will confine ourselves to the spatial aspects of this symbol. In regard to this, some important questions emerge: Is there an appropriate way of consistently aggregating spatial micro data to macro data? How are model results affected by the choice of a specific spatial scale? These questions will be dealt with in the present paper, with special emphasis on model results emerging from different spatial scales (Baumann et al., 1983).

Prior to any statistical and econometric regional analysis, a basic decision to be made when designing regional information systems is the choice of the appropriate level of spatial detail. Two main viewpoints play a role in this choice: the cost of collecting such data and the usefulness of the information which can be derived from the data. These viewpoints are in general conflicting. The costs involved in collecting data at the level of small basic areal units are usually extremely high. Also, rules of confidentiality may give rise to socio-political problems in the case of small basic units. On the other hand, the number of potential purposes for which information systems can be used may increase considerably when highly detailed spatial information can be produced.

Obviously, it is faily impossible to design simple rules of thumb for specifying the range of spatial detail within which a best compromise between these conflicting viewpoints has to be found. In this paper,

a more modest aim will be pursued. One particular aspect of the choice problem will be dealt with: the extent to which the choice of a certain level of spatial detail affects the quality of the information which can be produced for analytical and forecasting purposes.
Research in this field has given rise to remarkable results (see also Alker, 1969, Hordijk, 1979, Openshaw and Taylor, 1981, Lohmoeller et al., 1983).
It appears that statistical associations for aggregated populations may differ in magnitude and even in sign from those of individual populations members. The well-known ecological fallacy arises when correlations observed between variables at an aggregate level are used as substitutes for individual correlations. Clearly, when data are only available at a high level of spatial aggregation, it is impossible to test the sensitivity of statistical associations for lower levels of aggregation.

This paper has been organized as follows. Section 2 deals with different types of aggregation relevant for spatial information systems. In section 3, the use of disaggregate versus aggregate data is discussed in the context of (multi)regional models. Section 4 is devoted to an empirical illustration in the field of regional labour markets in which three spatial levels have been distinguished.

2. TYPES OF AGGREGATION

As indicated in the first section, aggregation may be carried out according to various dimensions: across individuals, economic sectors, spatial units, time, etc. Aggregation leads to a condensation of information. It inevitably entails a loss of detail (see also Orcutt et al., 1968), but - if property carried out - may improve the understanding of phenomena by focussing on important general features of data. Data as such is an insufficient basis, however, for determining the relevance and importance of a feature: the purposes for which the information is needed also have to play a role. This means that, preferably, aggregation procedures should be adapted to the purposes of a study.

As the present paper deals with regional information systems, we will pay special attention to the aggregation of spatial units (section 2.1). as already mentioned in the introduction, aggregation may considerably affect the size and sign of statistical associations between variables. Therefore section 2.2 will be devoted to an examination of the effects of aggregation on the specification of relationships between variables.

2.1 Aggregation of Spatial Units

In many countries regional information systems are based on a rather generally accepted hierarchy of regions consisting of several levels: municipality, county, province (state) and nation. For the large majority of purposes, users confine themselves also to this regionalization. This means that users usually only express their information needs in terms of the desired average size of regions, but not in terms of the composition of regions.

One should be be aware that the composition of regions may be based on entirely different principles (see e.g. Harvey, 1969). One way is to stick to existing administrative regions. Another way is to construct regions on the basis of a homogeneity principle, in order to achieve a high intraregional similarity according to economic structure, type of landscape, degree of urbanization, etc. A third way is to use the functionality principle which means that the composition of regions is determined on the basis of intensity of spatial linkages.

A regionalization principle does not completely determine the composition of regions. Each principle has to be supplemented with a clustering method before a regionalization can be achieved. As shown in Fischer (1982), there are many clustering methods which will in general give rise to different aggregation of basic areal units.

2.2 Aggregation of Relationships

As long as descriptive purposes are dominant in the use of information systems, aggregation of (spatial) units will not give rise to many problems. When analytical and forecasting purposes come to play an important role, this does no longer hold true, however. Then the question arises how relationships between variables are affected by aggregation of individual units (Van Daal, 1980, Akdeniz and Milliken, 1975).

In our presentation of the problems involved in aggregating relationships we will follow Theil (1954). Consider I individual units (for instance households or basic areal units). Suppose that for each unit i the variable y_i depends linearly on the predetermined variables x_{ki} (k= 1,..., K)

Thus, one arrives at the following micro-relations:

$$y_{it} = \alpha_i + \sum_{k=1}^{K} \beta_{ki} \; x_{kit} + u_{it} \quad (t=1, \ldots, T) \tag{1}$$

where α_i and β_{ki} are the micro-parameters and u_{it} is a disturbance term, being independent of x_{kit} ($\forall$ k,i,t) and having zero mean. Macro-variables can be defined as follows:

$$y_t = \sum_{i=1}^{I} y_{it} \; , \quad x_{kt} = \sum_{i=1}^{I} x_{kit} \quad (k=1,\ldots, K \; , \; t=1, \ldots, T) \tag{2}$$

(or as some weighted average of the micro-variables).

Aggregation of relationship (1) across individuals gives rise to:

$$y_t = \sum_{i=1}^{I} \alpha_i + \sum_{i=1}^{I} \sum_{k=1}^{K} \beta_{ki} \, x_{kit} + \sum_{i=1}^{I} u_{it} \quad (t=1, \ldots, T) \tag{3}$$

This expression is in general not identical to the macro-relation obtained by interpreting (1) in terms of macro-variables:

$$y_t = \alpha + \sum_{k=1}^{K} \beta_k \, x_{kt} + u_t \quad (t=1, \ldots, T) \tag{4}$$

in which α and β_k are the macro-parameters and u_t disturbance terms with zero means.

In this case of a simple linear model a necessary and sufficient condition for complete correspondence between (3) and (4) is (assuming that no restrictions are imposed on the distribution of the predetermined variables):

$$\beta_{ki} = \beta_{kj} \quad \forall i,j \tag{5}$$

When condition (5) holds, we speak of perfect aggregation. When condition (5) does not hold, while the macro-relation (4) is still used (because of the absence of data at the individual level) one commits a specification error which may give rise to misleading results.

Condition (5) can be relaxed when certain restrictions are imposed on the distribution of the predetermined variables. These restrictions can be formulated by means of so-called auxiliary equations. More details can be found in Theil (1954).

3. THE RELEVANCE OF PERFECT AGGREGATION

In the former section we concluded that only in rather exceptional cases an analysis at a high level of (spatial) aggregation is in agreement with behavioural relationships specified at the micro-level. Therefore, one might expect a strong orientation in research towards analysis at low levels of spatial aggregation. The reality is different, however: it appears that aggregate data are frequently used and often with reasonable results.

Possible explanations of this phenomenon are:

1. Macro-data are usually more reliable than data on more disaggregated levels (measurement errors may level each other out). Also the length of the time-series available for the former are often longer than for the micro-data.
2. The coefficients of the independent variables do not differ significantly at the disaggregate level. Then the condition of perfect aggregation is approximately satisfied.
3. The point of departure in the theory of perfect aggregation is the micro-relation (1). When this equation would have been improperly specified, the notion of perfect aggregation loses its value, however. For example, Kelejian (1980) argues that in the case of micro-relations which are non-linear, the notion of perfect aggregation loses its relevance. More importantly, Grunfeld and Griliches (1960) and

Green (1977) point to the fact that in equation (1) only variables of one particular level play a role. Thus, it assumes that behaviour at the disaggregate level is not influenced by variables at the aggregate level. Grunfeld and Griliches present results on forecasts of investments according to which the macro-approach (4) performs better than the disaggregate approach (3). Hence, they conclude that aggregation is not necessarily bad.

The last mentioned explanation, which is the most fundamental one, is closely related to some results obtained in the field of multiregional modelling. It immediately ties in with arguments concerning 'top-down' versus 'bottom-up' approaches in regional modelling. Therefore, we will now pay attention to the meaning of these approaches.

The following classification of multiregional models is fairly common (see among others Courbis, 1982, and Nijkamp and Rietveld, 1982).

1. Top-down models
2. Bottom-up models
3. Interregional models
4. Regional-national models

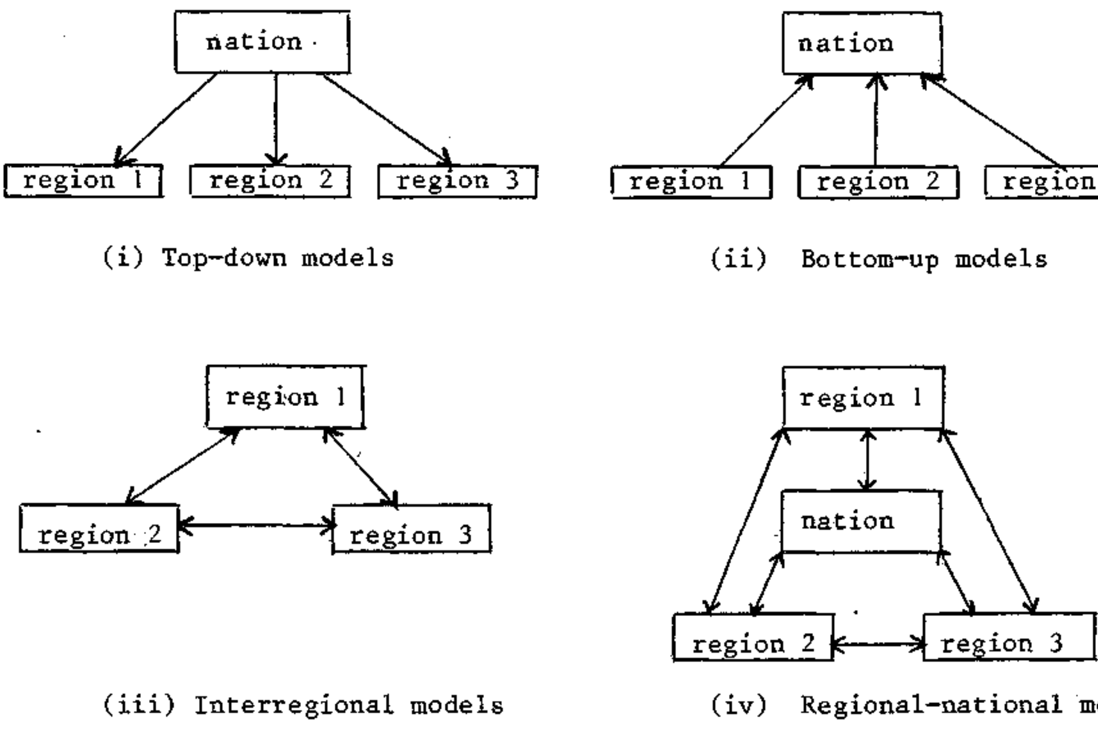


Figure 1. Structures of multiregional models.

(1) Top-down models

In the case of top-down models, the values of the national variables are assumed to be known. The values of the corresponding regional variables can be deduced from the known national variables by means of a disaggregation method. An advantage of this way of model building is that a multiregional model with a top-down structure can be linked immediately with already existing national economic models. One of the disadvantages of top-down models is the absence of feedback effects of regional development on national variables.

(2) Bottom-up models

In these models, first the variables on the regional level are determined: the corresponding national variables follow by aggregation. A disadvantage of bottom-up models is that they are often based on low-quality data. The availability and reliability of regional data leaves much to be desired compared with national data. An advantage of bottom-up models is that they can be used for investigating the possible conflict between national growth and interregional (in-)equality.

(3) Interregional models

In these models, the relationships between regions receive particular attention (usually by means of input-output models). These models are especially useful when the understanding of interregional relations is the main interest and when no restrictions on these relations are imposed because of national conditions.

(4) National-regional models

In this case, the top-down approach and the bottom-up approach are combined. The relations between various regions can also be considered. The advantage of this union is that on the one side there exists a feedback from the national to the regional level and on the other side, from the regional variables to the national variables. The values of the national and regional variables are determined simultaneously. A disadvantage of this type of model is the complexity of the model structure.

Comparing this typology with the approach to aggregation problems presented in section 2, we note that in section 2 essentially a bottom-up approach is employed. Interregional linkages and linkages from the national level to the regional level are ruled out by the specification of (1). Both types of linkages are important in spatial analysis, however. A top-down approach is appropriate when it concerns variables related to institutions (markets, governments, firms) which operate at the national level. A bottom-up approach is only appropriate for variables which refer to institutions operating at low levels of spatial aggregation.

4. A NUMERICAL EXAMPLE : THE REGIONAL DISTRIBUTION OF DISABILITY BENEFITS

In this section we will give an empirical example of the effects of spatial aggregation. We will examine how changes in spatial scale affect the determinants of the number of persons receiving disability benefits in the Netherlands.

In this analysis we will not deal with alternative compositions of regions of the same scale (see section 2). Only the scale of regions will be varied. Concerning the various types of modelling approaches mentioned in section 3, we will only deal with bottom-up approaches. Extensions in other directions will be published in a forthcoming paper.

We have selected a topic for which a bottom-up approach seems to be reasonable. It is closely related to the labour market, which typically operates at the regional level. We will focus on the number of people receiving disability benefits after having withdrawn from the labour market. The data used can be found in GMD 1977-1981 and CBS 1977-1981a and b. Substantial regional differences do exist in the share of people receiving disability awards. It is interesting to explore the extent to which these differences can be explained by economic indicators such as the unemployment rate. If such an explanation can be carried out, an implication would be that hidden unemployment exists among the recipients of disability benefits. This would be an important implication, since in the Netherlands (until recently) the number of unemployed has been much smaller than the number of recipients of disability benefits. For example, in 1978, the number of unemployed was approximately 200,000, while the number of disability benefits recipients amounted to approximately 550,000 (see also Van de Bosch and Petersen, 1982).

For the explanation of the number of disability benefit recipients (DB), the following stock-flow approach will be used. The number of recipients at the end of year t (DB_t) is by definition equal to the same number one year before (DB_{t-1}) plus the inflow during the year (DBI_t) minus the outflow (DBO_t):

$$DB_t = DB_{t-1} + DBI_t - DBO_t \tag{6}$$

The number of people leaving the labour market and entering the stock of disability benefit recipients (DBI_t) depends on the volume of employment at the end of the preceding period (E_{t-1}). We assume that yearly a certain proportion of the employed persons starts receiving benefits while this proportion depends on the unemployment rate u_t:

$$DBI_t = (\beta_2 + \beta_3 \left(\frac{u_{t-1}}{1-u_{t-1}} \right)) E_{t-1} \tag{7}$$

so that

$$DBI_t = \beta_2 E_{t-1} + \beta_3 U_{t-1} \tag{8}$$

where U_{t-1} denotes the number of unemployed at the end of year t-1.

Concerning the number of people leaving the stock of disability benefit recipients, we simply assume that yearly a constant fraction leaves this stock:

$$DBO_t = \gamma DB_{t-1} \tag{9}$$

Thus, after substituting (8) and (9) into (6) one arrives at:

$$DB_t = \beta_1 DB_{t-1} + \beta_2 E_{t-1} + \beta_3 U_{t-1} + \varepsilon_t \tag{10}$$

where $\beta_1 = 1 - \gamma$ and ε_t is a disturbance term (normally distributed with zero mean) which represents neglected variables). We hypothesize that all parameters are positive, and that β_1 is smaller than 1.

Estimations of (10) have been carried out at three spatial levels: nation, 12 provinces and 40 counties. See Appendix 1 for a map. For the dependent variable regionalized information is only available for the years 1977-1981. Thus, the three parameters in (10) can only be estimated on the basis of four observations, giving rise to only one degree of freedom. It is not surprising therefore that most estimation results give rise to very high outcomes for the coefficients of determination R^2 and that statistically significant results are rare. The estimation results at the national level have been represented in Table 1. The parameters have been estimated by means of linear regression.

$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	R^2
.577	.063	.049	.992
(.213)	(.025)	(.074)	

Table 1. Estimation results of (10) at the national level (standard errors between parentheses).

We note that the parameters have the right sign but that they are not significant at the 5% level. Especially the outcome for β_3 is remarkably low: it suggests that the level of unemployment has only a marginal effect on the number of disability benefit recipients.

Estimation results at the provincial and county level are presented in Appendix 2. Clearly, the values found for the parameters differ considerably, so that the condition of perfect aggregation is not satisfied.

We will now examine to what extent a disaggregate approach leads to projections of the dependent variable which is better than the macro-approach. Thus we compute:

$$\hat{DB}_t = \hat{\beta}_1 \, DB_{t-1} + \hat{\beta}_2 \, E_{t-1} + \hat{\beta}_3 \, U_{t-1} \tag{11}$$

for the four time periods and the various spatial levels. The outcomes at the national, provincial and county levels will be denoted by $\hat{DB}_t^n$, $\hat{DB}_t^{pi}$ and $\hat{DB}_t^{cj}$ respectively. Subsequently, we compare $\hat{DB}_t^n$, $\sum_{i=1}^{12} \hat{DB}_t^{pi}$ and $\sum_{j=1}^{40} \hat{DB}_t^{cj}$ with the observed outcomes DB_t^n to see whether a disaggregate approach yields better results. The results are summarized in Table 2.

	1978	1979	1980	1981
DB_t^n	557493	586327	618201	635904
$DB_t^n - \hat{DB}_t^n$	22	-2861	4375	-1588
$DB_t^n - \sum_{i=1}^{12} \hat{DB}_t^{pi}$	31	-2777	4105	-1393
$DB_t^n - \sum_{j=1}^{40} \hat{DB}_t^{cj}$	77	-2618	3787	-1261

Table 2. Comparison of national and regional projections.

The table shows that in three out of four cases the calculations based on the data on a county level give the most satisfactory results, followed by the results on the provincial level and finally the results on the national level. So by using disaggregated data the national level of the variable concerned can be projected better in three out of four cases.

When we look at the size of the differences between projections and observations, we find that the mean absolute error at the national level is 2211 . For the provincial and county level we find the following values: 2075 and 1937 .
This means that the mean absolute error can be reduced with approximately 6,5% by disaggregating towards the provincial level and another 6,5% when the county level is considered.

A similar comparison between aggregate and disaggregate results can be carried out at the provincial level, since all counties are part of only one province (see Appendix 3).
When we leave aside the two provinces which coincide with only one county, we observe that in 27 out of 40 cases (10 provinces and 4 time

periods), the calculations at the county level give the best results, while in 13 cases an aggregate approach gives better results.

5. CONCLUDING REMARKS

In this contribution some reflections concerning the explanation of aggregate variables by means of disaggregated data are given.
The choice of the appropriate aggregation level depends on the problem under consideration.
In section 3 we have given some reasons why a disaggregate approach does not necessarily produce better results than an aggregate one. Especially when the phenomenon concerned has a top-down structure, an aggregate approach may be preferable. For the numerical application we have selected a case in section 4, for wich a bottom-up structure is most plausible. Indeed, we find that a disaggregate approach leads to somewhat better results. Concerning the estimation results, it should be remarked that the small number of observations did not allow the testing of alternative model specifications. For example, we did not deal with interregional linkages (see White and Hewings, 1982).

We conclude that in spatial studies the information obtained for analytical, planning and forecasting issues may be highly sensitive for shifts in the level of spatial aggregation. In order to test this sensitivity, simulation experiments such as carried out in this paper are inevitable.

REFERENCES

Akdeniz, F. and G. Milliken, 1975, The Relationship between Macro and Micro Parameters, International Economic Review, no. 16, pp. 511-515.

Alker, H.R., Jr., 1969, A Typology of Ecological Fallacies, in M. Dogan and S. Rokkan (eds.), Quantitative Ecological Analysis in the Social Sciences, The M.I.T. Press, Cambridge, Massachusetts, and London, England, pp. 69-86.

Baumann, J., M.M. Fischer and U. Schmidt, 1983, Multiregional Labor Supply Models for Austria: The Effects of Different Regionalizations in Multiregional Labor Market Modeling, Papers of the Regional Science Association.

Bosch, F.A.J. van den and C. Peteren, 1982, An Explanation of the Growth of Social Security Disability Transfers, paper presented at the symposium 'Some Economic Aspects of Disability', October 1981, Erasmus University Rotterdam.

CBS, 1977-1981a, Sociale maandstatistieken, Centraal Bureau voor de Statistiek, Voorburg.

CBS, 1977-1981b, Statistiek werkzame personen, Centraal Bureau voor de Statistiek, Voorburg.

Courbis, R. 1982, Multiregional Modeling: a General Appraisal, in M. Albegov, Å.E. Andersson and F. Snickars (eds.), Regional Development Modeling: Theory and Practice, Norht-Holland Publ. Co., Amsterdam, New York, Oxford, pp. 65-84.

Daal, J. van, 1980, On Aggregation of Economic Relationships, Erasmus University Rotterdam, revised version (mimeographed).

Fischer, M.M., 1982, Eine Methodologie der Regionaltexonomie: Probleme und Verfahren der Klassifikation und Regionalisierung in der Geographie und Regionalforschung, Bremer Beiträge zur Geographie und Raumplaning, Schwerpunkt Geographie, Universität Bremen, Bremen.

GMD, 1977-1981, Regionale gegevens AAW- en/of WAO-uitkeringsgenietenden, Gemeenschappelijke Medische Dienst, Amsterdam.

Green, H.A.J., 1977, Aggregation Problems of Macroeconomics, in G.C. Harcourt (ed.), The Microeconomic Foundations of Macroeconomics, MacMillan, London, pp. 179-194.

Grunfeld, Y. and Z. Griliches, 1960, Is Aggregation Necessarily Bad?, Review of Economics and Statistics, Vol. 42, no. 1, pp. 1-13.

Harvey, D., 1969, Explanation in Geography, Edward Arnold, London.

Hordijk, L., 1979, Problems in the Estimation of Econometric Relations in Space, Papers of the Regional Science Association, Vol. 42, pp. 99-118.

Kelejian, H.H., 1980, Aggregation and Disaggregation of Nonlinear Equations, in J. Kmenta and J.B. Ramsey (eds.), Evaluation of Econometric Models, Academic Press, New York, London, pp. 135-152.

Lohmoeller, J.B., J.W. Falter, A. Lind and J. de Rijke, 1983, Unemployment and the Rise of National Socialism: Contradicting Results from Different Regional Aggregations, paper presented at the Workshop on Analysis of Qualitative Spatial Data, March 28 - april 1 1983, Free University Amsterdam.

Nijkamp, P. and P. Rietveld, 1982, Structure Analysis of Spatial Systems, in B. Issaev, P. Nijkamp, P. Rietveld and F. Snickars (eds.), Multiregional Economic Modeling: Practice and Prospect, Norht-Holland Publ. Co., Amsterdam, New York, Oxford, pp. 35-48.

Openshaw, S., and P.J. Taylor, 1981, The Modifiable Areal Unit Problem, in R.J. Bennett and N. Wrigley (eds.), Quantitative Geography: A British View, Routledge and Kegan Paul, London, pp. 60-69.

Orcutt, G., H. Watts and J. Edwards, 1968, Data Aggregation and Information Loss, American Economic Review, Vol. 58, pp. 773-787.

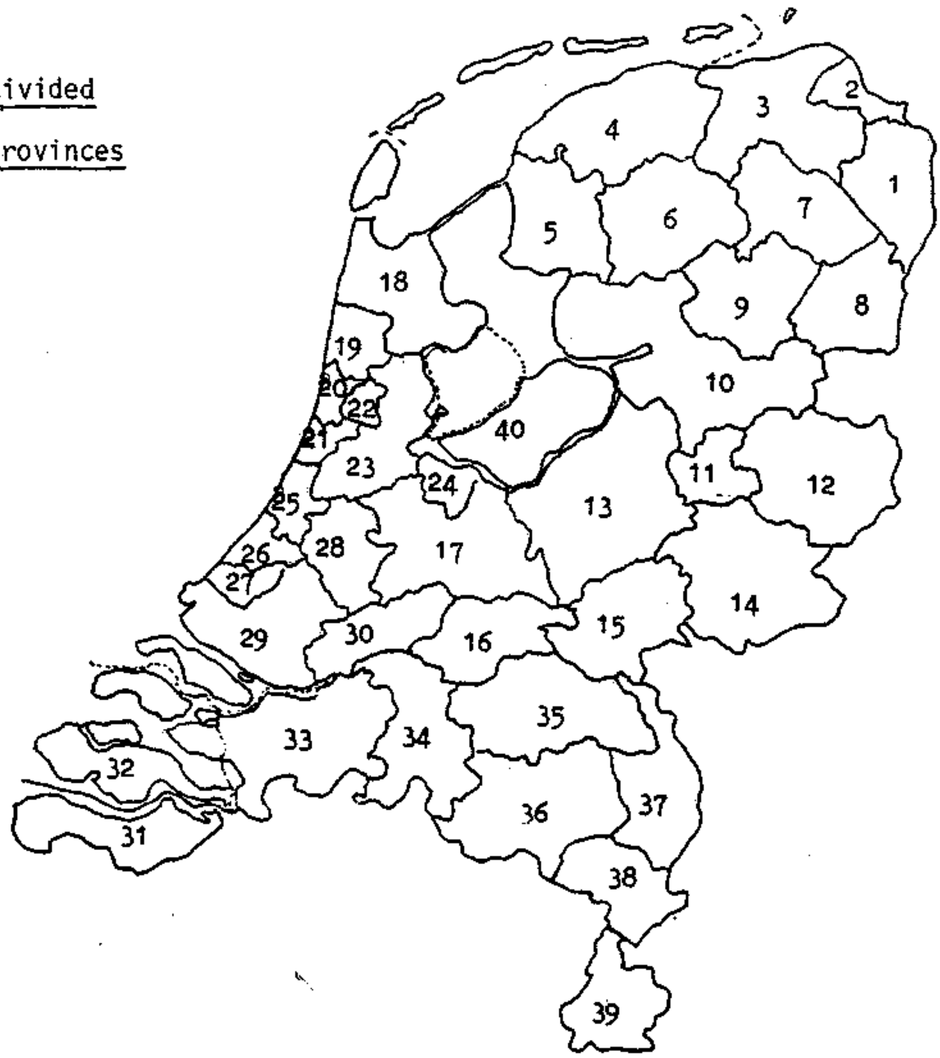
Theil, H., 1954, Linear Aggregation of Economic Relations, North-Holland Publ. Co., Amsterdam, New York, Oxford.

White, E.N. and G.J.D. Hewings , 1982, Space-Time Employment Modeling: Some Results Using Seemingly Unrelated Regression Estimates, Journal of Regional Science, vol. 22, no. 3, pp. 283-302.

APPENDIX I

The Netherlands divided

in counties and provinces



county	province	
01 Oost-Groningen 02 Delfzijl e.o. 03 Overig Groningen	Groningen	(1)
04 Noord-Friesland 05 Zuidwest-Friesland 06 Zuidoost-Friesland	Friesland	(2)
07 Noord-Drente 08 Zuidoost-Drente 09 Zuidwest-Drente	Drente	(3)
10 Noord-Overijssel 11 Zuidwest-Overijssel 12 Twente	Overijssel	(4)
13 Veluwe 14 Achterhoek 15 Arnhem/Nijmegen 16 Zuidwest-Gelderland	Gelderland	(5)
17 Utrecht	Utrecht	(6)
18 Kop van Noord-Holland 19 Alkmaar e.o. 20 IJmond 21 Agglomeratie Haarlem 22 Zaanstreek 23 Groot-Amsterdam 24 Gooi en Vechtstreek	Noord-Holland	(7)
25 Agglomeratie Leiden 26 Agglomeratie 's-Gravenhage 27 Delft en Westland 28 Oostelijk Zuid-Holland 29 Groot-Rijnmond 30 Zuidoost Zuid-Holland	Zuid-Holland	(8)
31 Zeeuws-Vlaanderen 32 Overig Zeeland	Zeeland	(9)
33 West Noord-Brabant 34 Midden Noord-Brabant 35 Noordoost Noord-Brabant 36 Zuidoost Noord-Brabant	Noord-Brabant	(10)
37 Noord-Limburg 38 Midden-Limburg 39 Zuid-Limburg	Limburg	(11)
40 Zuidelijke IJsselmeer-polders	Zuidelijke IJsselmeer polders	(12)

APPENDIX 2 The coefficients of equation (10) and R^2 for three aggregation levels (standard errors between parentheses)

1. National level

	β_1		β_2		β_3		R^2
Neth.	.577	(.213)	.063	(.025)	.049	(.074)	.992

2. Provincial level

Prov.	β_1		β_2		β_3		R^2
1	.106	(.144)	.146	(.021)	.187	(.046)	.997
2	-.168	(.410)	.172	(.056)	.203	(.092)	.984
3	.325	(.114)	.127	(.018)	.150	(.040)	.998
4	-2.697	(1.806)	.522	(.251)	.997	(.503)	.982
5	.379	(.200)	.096	(.026)	.080	(.060)	.991
6	.505	(.236)	.069	(.026)	.032	(.110)	.971
7	.614	(.141)	.057	(.016)	.045	(.071)	.991
8	.424	(.231)	.060	(.020)	.062	(.076)	.981
9	.787	(.482)	.030	(.040)	.030	(.249)	.971
10	.622	(.144)	.057	(.016)	.040	(.040)	.998
11	1.040	(.274)	.017	(.047)	-.139	(.121)	.993
12	.455	(.166)	.069	(.012)	.258	(.092)	.998

3. County level

county	β_1		β_2		β_3		R^2
1	.389	(.111)	.146	(.023)	.236	(.061)	.994
2	.365	(.135)	.105	(.019)	.165	(.053)	.999
3	-.364	(.071)	.186	(.009)	.194	(.015)	1.000
4	.190	(.509)	.111	(.063)	.124	(.107)	.981
5	.178	(.237)	.124	(.031)	.206	(.084)	.976
6	-.337	(.804)	.234	(.133)	.183	(.147)	.939
7	.133	(.110)	.172	(.020)	.139	(.026)	.999
8	.459	(.136)	.107	(.021)	.094	(.053)	.993
9	.172	(.568)	.138	(.084)	.273	(.239)	.945
10	-.338	(.240)	.176	(.030)	.192	(.054)	.987
11	-.010	(.577)	.145	(.078)	.103	(.095)	.974
12	-1.674	(.338)	.407	(.051)	.903	(.117)	.994
13	.116	(.213)	.132	(.029)	.148	(.056)	.996
14	.327	(.336)	.102	(.045)	.141	(.109)	.991
15	.433	(.188)	.088	(.023)	.050	(.060)	.981
16	.566	(.043)	.084	(.006)	.003	(.013)	1.000
17	.505	(.236)	.069	(.026)	.032	(.110)	.971
18	.473	(.000)	.083	(.000)	.123	(.000)	1.000
19	.604	(.047)	.063	(.006)	.084	(.016)	1.000
20	1.090	(.236)	-.001	(.022)	-.032	(.139)	.991
21	.443	(.153)	.104	(.023)	.038	(.118)	.951
22	.937	(.144)	.026	(.015)	-.115	(.104)	.991
23	.547	(.142)	.061	(.014)	.037	(.074)	.989
24	.819	(.001)	.032	(.000)	-.045	(.001)	1.000
25	.095	(.128)	.111	(.014)	.111	(.037)	.994
26	-.058	(.341)	.103	(.029)	.107	(.122)	.859
27	.610	(.182)	.038	(,013)	.023	(,103)	.974
28	.549	(.182)	.044	(.014)	.102	(.061)	.997
29	.698	(.175)	.034	(.015)	.010	(.050)	.991
30	.479	(.171)	.062	(.016)	.061	(.070)	.993
31	.968	(.409)	.020	(.032)	-.116	(.263)	.979
32	.683	(.515)	.036	(.045)	.091	(.213)	.970
33	.696	(.186)	.044	(.021)	.040	(.043)	.998
34	.490	(.085)	.075	(.009)	.062	(.031)	.998
35	.508	(.181)	.075	(.021)	.063	(.057)	.995
36	1.114	(.077)	-.006	(.009)	-.078	(.020)	1.000
37	2.025	(.023)	-.127	(.003)	-.490	(.007)	1.000
38	1.130	(2.695)	-.000	(.473)	-.168	(.503)	.932
39	.715	(.475)	.070	(.083)	.082	(.308)	.962
40	.445	(.166)	.069	(.012)	.258	(.092)	.998

APPENDIX 3 Comparison of projections at county and province level

		1978	1979	1980	1981
1.	DB_t^{p1}	26046	26909	28298	28853
	$DB_t^{p1} - \hat{DB}_t^{p1}$	-50	* -10	111	-54
	$DB_t^{p1} - \sum_{i=1}^{3} \hat{DB}_t^{ci}$	* -28	-15	* 73	* -31
2.	DB_t^{p2}	21140	21910	23024	23859
	$DB_t^{p2} - \hat{DB}_t^{p2}$	* 131	* 32	-244	94
	$DB_t^{p2} - \sum_{i=4}^{6} \hat{DB}_t^{ci}$	230	-65	*-235	* 92
3.	DB_t^{p3}	18514	19423	20669	21478
	$DB_t^{p3} - \hat{DB}_t^{p3}$	* -23	-41	* 98	* -37
	$DB_t^{p3} - \sum_{i=7}^{9} \hat{DB}_t^{ci}$	-39	* -30	111	-42
4.	DB_t^{p4}	41109	43875	46306	47496
	$DB_t^{p4} - \hat{DB}_t^{p4}$	-203	*-215	582	-183
	$DB_t^{p4} - \sum_{i=10}^{12} \hat{DB}_t^{ci}$	* -25	-221	* 358	*-119
5.	DB_t^{p5}	67503	70590	74018	75540
	$DB_t^{p5} - \hat{DB}_t^{p5}$	* -24	-291	476	*-169
	$DB_t^{p5} - \sum_{i=13}^{16} \hat{DB}_t^{ci}$	-26	*-272	* 465	-173
6.	DB_t^{p6}	35510	37405	39489	40631
	$DB_t^{p6} - \hat{DB}_t^{p6}$	120	-448	465	-134
	$DB_t^{p6} - \hat{DB}_t^{c17}$	120		465	-134

*) means: smallest difference between observed and computed value.

		1978	1979	1980	1981
7.	DB_t^{p7}	98685	104368	110329	113865
	$DB_t^{p7} - \hat{DB}_t^{p7}$	118	-663	789	-247
	$DB_t^{p7} - \sum_{i=18}^{24} \hat{DB}_t^{ci}$	* 102	* -580	* 673	* -200
8.	DB_t^{p8}	99073	102914	107488	109754
	$DB_t^{p8} - \hat{DB}_t^{p8}$	50	-650	888	-293
	$DB_t^{p8} - \sum_{i=25}^{30} \hat{DB}_t^{ci}$	* 47	* -589	* 809	* -271
9.	DB_t^{p9}	9701	10296	11178	11622
	$DB_t^{p9} - \hat{DB}_t^{p9}$	-38	-99	218	-84
	$DB_t^{p9} - \sum_{i=31}^{32} \hat{DB}_t^{ci}$	* -34	* -97	* 205	* -77
10.	DB_t^{p10}	76889	81683	86685	89776
	$DB_t^{p10} - \hat{DB}_t^{p10}$	* -82	-132	354	-149
	$DB_t^{p10} - \sum_{i=33}^{36} \hat{DB}_t^{ci}$	-99	* -99	* 321	* -133
11.	DB_t^{p11}	61779	65045	68415	70037
	$DB_t^{p11} - \hat{DB}_t^{p11}$	* 20	-285	* 412	* -151
	$DB_t^{p11} - \sum_{i=37}^{39} \hat{DB}_t^{ci}$	-182	* -226	589	-189
12.	DB_t^{p12}	1544	1909	2392	2993
	$DB_t^{p12} - \hat{DB}_t^{p12}$	12	25	-44	14
	$DB_t^{p12} - \hat{DB}_t^{c40}$	12	25	-44	14

*) means: smallest difference between observed and computed value.

Serie Research Memoranda (1)

1980-1	P. Nijkamp and H. Voogd	New Multicriteria Methods for Physical Planning by Means of Multidimensional Scaling Techniques
1980-2	H.P. Smit	Medium- and Long-Term Models for the ESCAP-Region - A Review of exisging models and a proposal for a new model system
1980-3	P.v. Dijck en H. Verbruggen	Productive Employment in Developing Countries' Exporting Industries
1980-4	P. Nijkamp and L. Hordijk	Integrated Approaches to Regional Development Models; A survey of some Western European Models
1980-5	P. Nijkamp	Soft Econometric Models; An Analysis of Regional Income Determinants
1980-6	P. Nijkamp and F.v. Dijk	Analysis of Conflicts in Dynamical Environmental Systems via Catastrophe Theory
1980-7	E. Vogelvang	A short term econometric Model for the Consumer demand of Roasted Coffee in The Netherlands
1980-8	N.van Hulst	De effectiviteit van de geleide loonpolitiek in Nederland
1980-9	P. Nijkamp	A survey of Dutch integrated Energy-Environmental-Economic Policy Models
1980-10	P. Nijkamp	Perspectives for Urban Analyses and Policies
1980-11	P. Nijkamp	New developments in Multidimensional Geographical data and Policy Analysis
1980-12	F.C. Palm, E. Vogelvang and D.A. Kodde	Efficient Estimation of the Geometric Distributed Lag Model; some Monte Carlo Results on Small Sample Properties

1981-1	E. Vogelvang	A quarterly econometric model for the Price Formation of Coffee on the World Market
1981-2	H.P. Smit	Demand and Supply of Natural Rubber, Part I
1981-3	R. Vos	The political Economy of the Republic of Korea; A proposal for a model framework of an open economy in the ESCAP-region, with emphasis on the Role of the State
1981-4	F.C. Palm	Structural Econometric Modeling and Time Series Analysis - Towards an Integrated Approach