The delimitation and coherence of functional and administrative regions

Cörvers, Frank

The Delimitation and Coherence of Functional and Administrative Regions

Journal:	*Regional Studies*
Manuscript ID:	CRES-2006-0019.R2
Manuscript Type:	Main Section
JEL codes:	J61 - Geographic Labor Mobility|Immigrant Workers < J6 - Mobility, Unemployment, and Vacancies < J - Labor and Demographic Economics, R12 - Size and Spatial Distributions of Regional Economic Activity < R1 - General Regional Economics < R - Urban, Rural, and Regional Economics, R23 - Regional Migration|Regional Labor Markets|Population < R2 - Household Analysis < R - Urban, Rural, and Regional Economics
Keywords:	regional disparities, Travel-To-Work-Areas (TTWA), commuting, functional regions

The Delimitation and Coherence of Functional and Administrative Regions*

FRANK CÖRVERS, Research Centre for Education and the Labour Market (ROA), Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands, email: f.coervers@roa.unimaas.nl

MAUD HENSEN, Hogeschool Zuyd, P.O. Box 550, 6400 AN Heerlen, The Netherlands, email: m.hensen@hszuyd.nl

DION BONGAERTS, Universiteit van Amsterdam, Roeterstraat 11, 1018 WB Amsterdam, The Netherlands, email: d.g.j.bongaerts@uva.nl

May 2007

Abstract

We test whether functional regions in the Netherlands show more labour market coherence between the municipalities included in them than the Dutch administrative regions. It turns out that regional disparities are not significantly smaller within functional than within administrative regions with respect to income level, housing prices, employment rate and unemployment rate. We argue that the numerous functional delimitations of the labour market that have been made for many countries in other studies are only useful for policy making if they clearly outperform the administrative delimitations with respect to some relevant indicators of labour market coherence or regional disparities.

Keywords: functional regions, commuting, Travel-To-Work-Areas (TTWA), regional disparities

JEL-codes: R23, J61

* This research has been granted by the Dutch Ministry of Economic Affairs. The authors would like to thank Jacco Hakfoort, Marcel van Wijk, Christoph Meng and three anonymous referees for their valuable comments on earlier versions of this paper.

1 Introduction

The definition of a regional labour market is very important in the light of regional labour market policies (Ball, 1980). In particular for research and policy-making purposes, the delimited areas should exhibit functional similarities. The economic diversity within an administratively defined region may be so large that comparison between regions is not justified. The decisions made concerning the planning, distribution and allocation of resources among the various regions derived, are not likely to be the most effective and meaningful relative to the decision that would be made if the underlying regional patterns were known (Amedo, 1968). For example, areas with high unemployment rates but administratively falling within regions with low average unemployment rates may receive no assistance from the national government or the European Union.

The dominant concept in defining functional regions is that of labour markets, as is illustrated by the substantial literature in this field by e.g. Andersen (2002), Baumann et al. (1996), Coombes et al. (1986), Casado-Díaz (2000), Eurostat (1992), Fox and Kumar (1965), Killian and Tolbert (1993), Newell and Papps (2002). For the delimitation of functional labour market regions commuting flows are used in most OECD countries (OECD, 2002). Commuting conditions like distance, closeness, commuting thresholds, travel times determine the magnitude of the commuting flows between areas. On the basis of commuting flows, a functional region can then be defined as a region in which a large proportion of the workers both live and work.

However, it may be difficult for local planning authorities to interpret statistical information as well as to set policy goals with regard to a different division of areas than the administrative division (see Coombes and Openshaw, 1982 and Green and Coombes, 1985). It may be even more difficult and costly to reorganise local government structure according to a particular functional division of regions. Therefore a functional regionalisation should have clear benefits over the administrative regionalisation to make it really valuable for policy makers. This may be the case if the labour market areas within the functional regions are significantly more coherent than within the administrative regions.

The issue of regionalizing countries into functional regions can be regarded as a Modifiable Areal Unit Problem (MAUP, see e.g. Openshaw, 1984 and Unwin, 1996), which consists of both a scale and an aggregation problem. The scale problem refers to the choice of the appropriate number of regions, whereas the aggregation problem refers to the choice of an appropriate regionalisation (Baumann et al., 1996). Labour market studies in which data on basic spatial units - in this paper municipalities - is aggregated by using administratively defined regions do generally not inform about the use of other delineations. Due to the potentially infinite options to aggregate the data, one should be aware of the spatial variation of the data, particularly when using more or less arbitrary and 'modifiable' boundaries between areas.

The aim of this paper is to examine whether the coherence of the functional labour market regions, which are carefully delimited without using arbitrary criteria, is larger than the coherence of the administrative defined labour market regions. The larger the coherence of the areas *within* the delimited regions, the larger the heterogeneity *between* the delimited regions is expected to be for particular measures related to the economy and the labour market. The comparison of different regionalisations using economic indicators has hardly been examined in the field of labour economics. We will attempt to make a contribution in this field by testing for the labour market coherence of different functional and administrative regionalisations for the case of the Netherlands.

Labour market coherence will be measured by four economic indicators that are commonly used in studies on regional disparities (see e.g. Chapter 2 in OECD, 2005). These indicators are: (i) income level; (ii) housing prices; (iii) employment rate, and (iv) unemployment rate. Moreover, the use of these indicators will be justified in a commuting model of the labour market. For these indicators we will test whether functionally defined regions show more coherence between the municipalities included in it than the administratively defined regions. According to the commuting model presented in this paper, low commuting flows between functional regions should go hand in hand with large interregional differences in wages, housing prices, employment and unemployment rates. The new approach in the paper is that the administrative and functional regions are linked to these measures of regional economic performance to get an idea of which of the regionalisations should be preferred.

The paper falls into two parts. Section 2 refers to the first part and discusses the relevant literature on delimitation studies. Furthermore, Section 2 explains the delimitation method used in this paper and shows the results of applying this method on the commuting data of the Netherlands at four different scale levels. Section 3 refers to the second part and starts with a commuting model of the labour market. The model underpins the use of the four economic indicators mentioned above for the empirical testing on labour market coherence. Next Section 3 discusses the results of this testing. In Section 4 we conclude.

2 Regionalization based on travel to work flows

2.1 Previous studies

The use of delimitations of functionally defined regions varies between countries (OECD, 2002). In Great Britain labour market areas have been defined to analyse labour market phenomena, calculate unemployment rates, identify assisted areas for industrial policies and reorganize local government. These labour market areas are known as Travel-To-Work-Areas

(TTWA, see Coombes et al., 1986, ONS and Coombes, 1998, and Coombes, 2005). They are the result of a delimitation procedure using the direct and indirect relationships between municipalities by analysing the behaviour of individual commuters. Also for a number of other European countries the same regionalisation algorithm has been undertaken (Eurostat, 1992). For some countries the results of more or less the same regionalisation algorithm have been published in separate studies. These countries include Denmark (Andersen, 2002), the region of Valencia in Spain (Casado-Díaz, 2000), Italy (Sforzi et al., 1997) and New Zealand (Newell and Papps, 2002).

The delimitation procedure of TTWAs was developed to generate the maximum possible number of areas with a self-containment level of at least 75% (see Smart, 1974) and a minimum size of the area of 3,500 resident workers. Within the area, at least 75% of the jobs should be fulfilled by the residents of that area (demand-side self-containment) and at least 75% of the residents should work in the area (supply-side self-containment). In addition, a 70% threshold was accepted if the size of the area exceeded 20,000 residents. The municipalities with the highest self-containment levels are selected as the starting point for the delimitation procedure. However, the determination of the threshold values determines to a great extent the number of local labour market areas defined.[1] Lower threshold values would yield more local labour market regions, as a result of which the usefulness of the delimitation for policy-making may be reduced. Other absolute threshold values to select employment centres are used, for example, by Giuliano and Small (1991), who defined contiguous employment areas in the Los Angeles region of the U.S. as areas with at least ten workers per acre or more than 10,000 workers. Moreover, Van der Laan and Schalke (2001) argued that the use of situation-dependent absolute figures is responsible for different classifications that depend on the country and the period of analysis. To avoid the problems related to the use of absolute figures when defining TTWAs, they used relative instead of absolute criteria to delimit local labour market areas in the Netherlands. Nevertheless, also the choice of these relative criteria seems to be rather arbitrary.

The more or less arbitrary nature of many delimitation procedures is illustrated by the fact that the regionalisation algorithm had to be adapted with respect to some thresholds to get 'satisfactory' results in the studies that used the algorithm by Coombes (see above). The 'fine tuning' for each country is considered to be necessary to deal with the wide variety of local labour market areas. However, such fine tuning can be rather arbitrary since "the TTWAs form only one of the innumerable possible different aggregations …… to achieve the goal of 75% self-containment."[2] Coombes and Openshaw (1982, p. 142) Therefore we do not agree with Coombes et al. (1986) that this flexible multi-stage aggregation approach should be preferred above the alternative approach applied by e.g. Brown and Holmes (1971) and Masser and Brown (1975). Although the alternative approach has been criticised for being too deterministic and solely based on numerical taxonomy principles and statistical objectives (Coombes et al., 1986), we prefer this approach since it does not require the modification of criteria in a rather arbitrary way.

The different methods in the alternative approach are reviewed and tested by e.g. Masser and Scheurwater (1980), Fisher (1980) and Baumann et al. (1996). From these reviews it turns out that there is no clear a priori advantage of one or the other method. We have chosen to delimit regions in the Netherlands by using the Markov analytic functional distance approach, which transforms the interaction matrix of commuting flows between municipalities into a mean first passage time (MFPT) matrix. This method is one of the most widely used regionalisation methods in the alternative approach, and is intuitively appealing because the cells of the MFPT matrix represent functional distances between municipalities.

In this paper we are mainly interested in whether our functional division of regions can outperform the administrative division with respect to the four economic indicators mentioned in the introduction. “The fundamental question of which regionalisation should be chosen should be decided on the basis of a set of relevant criteria, such as R^2-, *t*-values and a priori signs, etc.” (Bauman et al., 1996, p. 380) By using the economic indicators as the relevant criteria in our analysis we deal with the aggregation problem. We will cope with the scale problem by performing the analysis on the comparison between the functional and the administrative division at different scale levels.

2.2 Methodology used

Following Brown and Holmes (1971) and Baumann et al. (1996) we first transform the standard interaction matrix of commuting flows between municipalities into a mean first passage time matrix (MFPT matrix, see Appendix A), and then cluster municipalities that have more interaction with each other than with municipalities outside the cluster. This aggregation method aims to maximize within-region commuting flows by merging the two adjacent municipalities (or clusters) with the smallest distances expressed by the mean first passage time indicators, that is the greatest mutual interaction in commuting flows. Setting the maximum number of clusters preferred, this aggregation method leads to an optimal delimitation of functional regions.

Given the initial distance matrix the clustering procedure can be started. We follow the method proposed by Ward (1963), which has been used by e.g. Masser and Scheurwater (1980) and Baumann et al. (1996) as well. Two municipalities/clusters i and j are only clustered if they are adjacent. The procedure subsequently clusters the two municipalities with the smallest d (functional distance). The new distances from a to all other clusters are calculated by minimizing the variance within clusters. The new distance $d(a,r)$ between a new cluster a consisting of p and q to another region/cluster r is calculated according to the formula:

$$d(r,a)=d(r,p\cup q)=\frac{(n_p+n_r)\cdot d(p,r)+(n_q+n_r)\cdot d(q,r)-n_r\cdot d(p,q)}{n_p+n_q+n_r}$$

where n_x is the number of elements in cluster x.

2.3 Commuting data

For the delimitation analyses, we have used the travel-to-work[3] data (OVG, 'Onderzoek VerplaatsingsGedrag' for 2001, 1991 and 1992) from Statistics Netherlands, which observed the travelling behaviour of a sample from the Dutch population. This travelling behaviour can be classified according to the motivation of the mobility decision. Apart from the decision to travel to work, other motives to travel are also observed, such as shopping or sports. To delimit the Netherlands, only the home-to-work journey is used as a motive for the mobility decision.

The number of observations used in the delimitation analysis of 2001 was 39,280. Since the number of observations was substantially lower for 1991, we also used the OVG data for 1992 as if the data sets are from one year. In the remainder of the paper we will refer to '1991' when we use the data of 1991/92. The total number of observations for these two years was still only 5,875. Given the number of the 484 remaining municipalities, the regionalisation of 1991 should be regarded as less reliable than the regionalisation of 2001. Due to lack of data, the five islands in the north ('Waddeneilanden') could not be clustered in the 2001 and 1991 delimitation analyses.

The average travel distance the workers travelled to reach their work location, was about 16 kilometres in 2001. In 1991 workers travelled on average 13 kilometres. The commuting behaviour of workers has therefore changed over time. Hence the delimitation of regions may have changed over time.

2.4 Results of the delimitations with commuting flows

The method described above allows us to produce any number of functional regions. The number of functional regions to be generated in the delimitation procedure has been set equal to the number of administratively defined regions in the Netherlands in order to compare the coherence of the regions in the administrative and functional divisions in the next section. Below we will discuss four different administrative divisions of the Netherlands. For reasons of space and readability we only show the figures of the 4-region administrative and functional divisions of 2001. In Bongaerts et al. (2004) the figures for 1991 and for the 12-, 24- and 40-divisions of administrative and functional regions can be found.

Eurostat uses the Nomenclature of Territorial Units for Statistics (NUTS) to divide countries into regions. According to NUTS1 the Netherlands is divided into 4 country parts: north, south, east and west. Figure 1 shows this administrative delineation of the Netherlands. Since the Netherlands is divided into 4 administrative regions in NUTS1, we generated 4 functional regions in accordance with the delimitation procedure from the preceding section.

[Insert Figure 1 about here]

Figure 2 presents the derived delineation for 4 functional regions in 2001. The functional division into 4 regions of the Netherlands is evidently different from the administrative 4-region division. It appears that the Utrecht region is a separate regional labour market according to the delineation into functional regions. Furthermore, the Zeeuws-Vlaanderen region, consisting of 3 municipalities, can also be seen as a separate – more homogeneous – region. This can be easily explained by the absence of a bridge or a tunnel across the Westerschelde estuary to connect Zeeuws-Vlaanderen and Zuid-Beveland.[4]

Moreover, the functional delineation based on commuting flows suggests that the northern part of the Netherlands interacts more with the western and the middle part of the Netherlands than is suggested by the administrative division. The same is true for the southern part. Apparently there is more north-south than east-west distinction between regions. This can be partly explained by the river Rhine flowing from east to west into the North Sea. Probably related to the course of the Rhine, the border between the functional regions Utrecht and South is almost the same as in the administrative 4-division or the administrative 12-division of provinces (see below). However, in the west the border between the functional regions North and South follows the administrative border between the provinces North Holland and South Holland instead of the Rhine.

The functional division into 4 regions also illustrates the strong polycentricity of the Randstad Holland region (see e.g. Musterd and Van Zelm, 2001), which consists of the cities of Amsterdam, Utrecht, the Hague and Rotterdam. In the functional division the latter two cities are part of the south and the cities of Amsterdam and Utrecht are situated in different functional regions, whereas in the administrative division the Randstad cities belong to one region (i.e. west, see Figure 1). A recent report by the Netherlands Institute for Spatial Research (Ritsema van Eck et al., 2006) concludes that the Randstad cannot be regarded as a single cohesive whole, although this holds to a lesser extent for commuting patterns than for business relationships and shopping expeditions between the urban regions.

[Insert Figure 2 about here]

The regional division of the Netherlands at the NUTS2 level refers to the 12 provinces of the Netherlands, which fall within the boundaries of the NUTS1 regions. These provinces represent the administrative layer in between the national government and the local municipalities. A large share of the regional budgets for policy planning is distributed over the provinces. The division of the Netherlands into 12 functional regions has been compared to the Dutch division in 12 provinces.

Other administrative delimitations that are compared to the functional delimitations include the RBA division of 28 regions and the COROP division of 40 regions (NUTS3). The RBA division refers to a delimitation of labour market areas, formerly used by the national employment agency. The COROP regions were delimited according to the nodal division principle, which means that every region contains a central municipality. Although the COROP regions can be considered more or less as functional regions, an additional requirement for this delimitation was that the COROP regions were situated within the boundaries of the provinces. Both the RBA and the COROP divisions have been widely used in structural analyses of labour markets, for analyzing territorial disparities, but also by specific administrative bodies to plan their policies.

In general, the delimitations using 1991 commuting flows lead to more small regions than in 2001. An explanation for this could be the shorter travel distances. In 1991 workers may have been less able – due to a less favourable infrastructure or their not having a car – or less willing to commute to reach their work location than in 2001. If commuting distances further increase during the next decade in the Netherlands, then there will hardly be left small regions that represent more or less closed labour markets.

3 Testing for the coherence of regions

3.1 A commuting model of the labour market

In this subsection we will analyse how commuting reduces wage inequality between regions.[5] It is hypothesized that large commuting flows between regions will reduce interregional wage differences, and interregional differences in housing prices, employment and unemployment rates. Therefore regions can be aggregated to one coherent region if the commuting flows between them are large.

Our starting point is a classical supply-demand representation of the labour market with α and β representing the exogenously given parameters. In this model the labour demand function can be characterized by:

$$L_{D,t}^{R} = \beta_{1}^{R} - \beta_{2}^{R} * W_{t}^{R} \tag{1}$$

where W_t^R is the real average wage earnings per worker and R denotes regions A, B. Furthermore, t refers to the period before and after commuting is allowed, where in period 1 (t = 1) no commuting is allowed and in period 2 (t = 2) it is. The labour supply function can be characterized by:

$$L_{S,t}^{R} = \alpha_{1}^{R} + \alpha_{2}^{R} * W_{t}^{R} \tag{2}$$

When regional labour supply and demand in period 1 are equal, the labour market in region R is in equilibrium. That is, $L_{D,1}^{R} = L_{S,1}^{R}$. Setting (1) equal to (2) yields the equilibrium wage level for both regions (A and B) in period 1:

$$W_{1}^{R*} = \frac{\alpha_{1}^{R} - \beta_{1}^{R}}{-\beta_{2}^{R} - \alpha_{2}^{R}} \tag{3}$$

where $\alpha_2^R > 0$ and $\beta_2^R > 0$ are the wage elasticities of supply and demand, respectively. It follows that for W_1^{R*} to be positive, $\alpha_1^R < \beta_1^R$.

In period 2 commuting is allowed. Assume that in period 1 $W_1^{A*} > W_1^{B*}$. This will stimulate workers to commute from B to A. However, workers usually have to make costs to travel or to cross natural or psychological barriers. These costs are related to the so-called commuting conditions mentioned in Section 1. Suppose that individuals are confronted with fixed costs F if they commute from region B to region A. The equilibrium wage level (if commuting between A and B takes place), becomes:

$$W_2^{A*} = W_2^{B*} + F \tag{4}$$

where $F \geq 0$ and $W_2^{A*} < W_1^{A*}$ and $W_2^{B*} > W_1^{B*}$. The higher wage level in region A in period 1 leads to a commuting flow from region B to A, which in turn decreases the wage level in region A and increases the wage level in B. It can be proved that there will be no commuting if the fixed costs F are too large, i.e. $W_1^{A*} \leq W_1^{B*} + F$. Thus, commuting between regions A and B starts only if $W_1^{A*} - W_1^{B*} > F$ and stops in period 2 if equation (4) is fulfilled.

The commuting flow L_C^A from region B to region A is equal to the difference between labour demand and supply in region A *(or B)*. In period 2, labour demand is larger than labour supply, due to the decrease in the wage level in region A, and vice versa for region B. For region A this is illustrated by equation (5).

$$\begin{aligned} L_C^A &= (L_{D,2}^A - L_{D,1}^A) - (L_{S,2}^A - L_{S,1}^A) \\ &= -(\beta_2^A + \alpha_2^A)(W_2^{A*} - W_1^{A*}) \\ &= -\gamma^A (W_2^{A*} - W_1^{A*}) > 0 \end{aligned} \tag{5}$$

L_C^A is larger than 0 since $\gamma^A > 0$ and $W_2^{A*} < W_1^{A*}$. The parameters γ^A and γ^B can be interpreted as the sensitivity of commuting flows to wage adjustments within a region. A large γ refers to high wage elasticities of supply and demand. By definition it holds that:

$$L_C^A + L_C^B = 0 \quad (6)$$

Therefore:

$$-\gamma^A (W_2^{A*} - W_1^{A*}) - \gamma^B (W_2^{B*} - W_1^{B*}) = 0 \quad (7)$$

$$\frac{W_1^{A*} - W_2^{A*}}{W_2^{B*} - W_1^{B*}} = \frac{\gamma^B}{\gamma^A}$$

Equation (7) implies that regions with relatively low wage elasticities are confronted with relatively large changes in the regional equilibrium wage due to commuting. The equilibrium wage levels in period 2 for region *A* and *B* are identical if there are no fixed costs. To commute from region *B* to region *A*, workers have to incorporate the fixed costs *F*. Therefore, the difference between the equilibrium wage levels in period 2 consists of the fixed costs *F* (see equation (4)). Combining (7) with (4) results in:

$$W_2^{A*} = \frac{\gamma^B}{\gamma^B + \gamma^A}(W_1^{B*} + F) + \frac{\gamma^A}{\gamma^B + \gamma^A} W_1^{A*} \quad (8)$$

Equation (8) shows that the new equilibrium wage level in region *A* is the weighted average of the old equilibrium wage levels in regions *A* and *B* corrected for fixed costs. The region with the largest wage elasticities has the largest weight. From equation (5) it follows that:

$$W_2^{A*} = \frac{-L_C^A}{\gamma^A} + W_1^{A*} \quad (9)$$

In the next subsections we want to test the relationship between the wage variance (in period 2, i.e. based on the observed wages) and the commuting flows. This relationship is predicted by the next equation, which can be derived by combining equations (6) and (9) for regions *A* and *B*:

$$W_2^{A*} - W_2^{B*} = \qquad (10)$$

$$\left(\frac{-L_C^A}{\gamma^A} + W_1^{A*}\right) - \left(\frac{-L_C^B}{\gamma^B} + W_1^{B*}\right) =$$

$$-\frac{\gamma^B + \gamma^A}{\gamma^B \gamma^A}(L_C^A) + (W_1^{A*} - W_1^{B*})$$

From equation (10) it can be readily understood that the interregional wage difference in period 2, which is equal to the fixed costs of commuting as follows from equation (4), is dependent on the wage elasticities and the interregional wage difference in period 1, which are both predetermined. It follows that for given wage elasticities and interregional wage differences when regions are closed (in period 1), the magnitude of the observed commuting flows is negatively related to the observed interregional wage differences (in period 2, i.e. when regions are open). The larger the commuting flows between regions, the lower the interregional wage differences. Both commuting flows and interregional wage differences reflect the commuting conditions mentioned before.

Since higher wages will raise housing prices and will pull more individuals to the labour market, higher costs of commuting may also be reflected in larger interregional differences in housing prices and labour participation (i.e. employment rates). Finally, since job searchers face relatively high costs of commuting between municipalities of different regions, low commuting flows between regions may be related to large interregional differences in unemployment rates.

The four economic indicators that follow from the above analysis are widely used indicators when analysing regional disparities. This is not only evident from a study by the OECD (2005) on the persistence of regional disparities in OECD countries, but also from a number of recent publications for the Netherlands. These studies include Atzema en Van Dijk (2005) on unemployment rates, Vermeulen (2005) on regional employment and unemployment rates, and Vermeulen and Van Ommeren (2006) on unemployment, housing prices and wages. Moreover, for many other small countries contributions on regional disparities with respect to these indicators can be found in Felsenstein and Portnov (2005). Finally, Hazans (2004) has found empirical evidence in line with the above commuting model. He shows that commuting significantly reduces wage disparities between areas in Estonia, Latvia and Lithuania, and increases the employment possibilities of the residents in these countries.

3.2 Descriptive statistics of economic indicators for labour market coherence

Below we will describe the mean and the standard deviation of the indicators following from the commuting model. This will be done for the 4 administrative and functional regions distinguished in the previous subsection. The calculation of the mean and the standard deviation per administrative or functional region is based on the indicators of the municipalities for which data is available. The first indicator concerns the average wage level per worker. This indicator can be measured by for example gross earnings paid by the employer. This data is, however, not available at the municipality level for the Netherlands. Therefore we have used the net personal income of workers, which incorporates the gross wages earned as well as income taxes, tax allowances and fiscal deductions. The income data has been drawn from the same survey as the travel-to-work data of Subsection 2.3 (OVG of Statistics Netherlands).

Moreover, data on housing prices was drawn from Statistics Netherlands that processed the data collected by the tax authorities. The housing prices are based on the tax declarations by home owners in the Netherlands. Finally, the employment and unemployment rates are based on the Labour Force Survey of Statistics Netherlands. However, data on the employment and unemployment rate at municipality level was only available for municipalities with more than 10,000 inhabitants. For these two indicators, data for 300 municipalities was used in the analysis. For housing prices, employment and unemployment rates no data was available for 1991. Table 1 gives an overview of the average values and standard deviations with regard to the four economic indicators for the municipalities in the 4 administrative regions of the Netherlands. The North region traditionally has the lowest labour participation, as is indicated by the employment and unemployment rate. In the West region the income level, the housing prices and the labour participation rates are the highest.

[Insert Table 1 about here]

The same overview is presented in Table 2 for the 4 functionally delimited regions. In the functional division the regions of Utrecht and Zeeuws Vlaanderen have the most extreme values for the four indicators. In Utrecht the income level, housing prices and labour participation are the highest, in Zeeuws Vlaanderen the lowest (except for the income level).

[Insert Table 2 about here]

3.3 Specification of the test

To test for the coherence of the regions, we carried out regression analyses using standard. Ordinary Least Squares (OLS) regression. The average values of the economic indicators of the municipalities in the functional or administrative regions have been regressed on the dummy variables of the regions to account for the differences in average income levels, housing prices, employment and unemployment rates between regions. The question in this context is whether there are any significant differences between the different clusters of municipalities with regard to these economic indicators. The following equation has been estimated to reveal the average income differences between functional or administrative regions:

$$Income_m = \beta_0 + \beta_{1,..k} * delimitation$$

where m stands for the municipalities, k is the number of regions minus 1, β_0 represents the average income level of the reference region, and $\beta_{1,..k}$ represent the differences between the average income level of the other regions and the reference region. The regressions are repeated for all couples of regions (i.e. taking different reference regions) of the same functional or administrative division. The number of couples in the 4, 12, 28 and 40 divisions is 6, 66, 378 and 780, respectively (i.e. $(k+1)*k/2$). Similar regression equations are estimated for housing prices, employment and unemployment rates. Significant differences in economic indicators between regions indicate that the delimitation is based on coherent regions. Finally we have counted the number of significant differences between the average levels of the economic indicators of regions.

3.4 Results

The higher the number of significant differences between the regions in the estimated equations, the lower the interaction of workers between these regions, and the higher the coherence of the municipalities within the regions. Tables 3 and 4 show the mean differences of the four economic indicators of all possible combinations of the four administrative and functional regions, respectively. Remarkably, the differences in income levels are not significant between the administrative and functional regions. On the other hand, housing prices are significantly different for almost all of the 6 pairs of regions. The employment and unemployment rates are significantly different for about half of the 6 pairs.

[Insert Table 3 about here]

The positive and negative signs of the differences across the four economic indicators are generally in accordance with the predictions that follow from the commuting model presented before. In most cases a region with a lower average income level than another region, also has a lower average housing price, a lower average employment rate and a higher average unemployment rate relative to the other region. The North region, for example, has a lower income level than the West region - although not significantly so -, a significantly lower housing price and employment rate, and a significantly higher unemployment rate than the West region.

[Insert Table 4 about here]

Tables 5 and 6 below give a complete overview of the percentages of significant differences for the four economic indicators with respect to the 4, 12, 28 and 40 functional and administrative delimitations in 2001 and 1991, respectively. For the income level in the 12, 28 and 40 division of regions, the functional delimitation performs slightly better than the administrative delimitation. It appears that, in terms of average income level, the functionally defined regions have slightly more coherence than the administrative regions of the

Netherlands. The functional division of 12 regions has the best score, since 27% out of the 66 pairs of regions have a significantly different income level. For the other three economic indicators, the performance is generally much higher for both the administrative and the functional delimitation. However, for these economic indicators the functional delimitation is not better than the administrative delimitation.

[Insert Table 5 about here]

With respect to the four economic indicators it follows that the number of regions that are significantly different from each other is more or less equal for the administrative and the functional 2001 delimitations. Only in the case of the 28 division of regions the performance is slightly better for the functional delimitation. The differences in performance between the administrative and functional divisions of 4, 12, 28 and 40 regions in 2001 are, however, small.

Table 6 shows the percentages of significant differences based on the 1991 delimitation of the Netherlands. For the interregional income differences in both 1991 and 2001, the functional 1991 delimitation performs better than the administrative delimitation (except for the 28 division with the 1991 average income level). As in 2001, the performance of the income level as an economic indicator of interregional differences is low. For the three other economic indicators we again find relatively large percentages of significant differences between administrative and functional regions. The functional division performs slightly worse for these indicators with respect to the 12- and 24-division.

[Insert Table 6 about here]

It is generally true that the more differentiated the delimitation is, the worse the relative performance. This holds for both the administrative and the functional delimitations. Although the absolute number of significant differences usually increases when the delimitation is more differentiated, we conclude that there is no advantage to differentiate between relatively small regions.

4 Conclusion

In many delimitation studies the procedures and algorithms are adapted with respect to some thresholds to get 'satisfactory' results. This 'fine tuning' is considered to be necessary but can also be regarded as rather arbitrary. The method applied in this paper avoids the use of a set of more or less arbitrary criteria by transforming the standard interaction matrix of commuting flows between municipalities into a mean first passage time (MFPT) matrix and applying the Ward clustering procedure. To compare the functionally defined regions with the administrative ones, the number of functional regions was tuned to the number of regions of the administrative delineation.

In this paper we argue that no matter what delimitation method is used, the resulting functional division of regions should be tested against the administrative division using a set of relevant criteria. Both from the commuting model presented in this paper and the reviewed empirical studies on regional disparities, it follows that we can distinguish four economic indicators of labour market coherence: (i) income level; (ii) housing prices; (iii) employment rate, and (iv) unemployment rate. For these indicators, we tested whether the municipalities within the functionally defined regions show more coherence (i.e. smaller disparities) than the municipalities within the administratively defined regions.

For both 1991and 2001 it appears that, in terms of income level, the functional regions have slightly more coherence than the administrative regions. The performance of income level as an economic indicator of differences between regions, however, was much worse than for the other economic indicators. A possible reason for the low percentage of significant differences in the income levels between regions, is that income is in fact an approximation for the wages of individual workers. For the other three economic indicators, the functional and the administrative regions showed, on average, the same coherence for both 1991 and 2001. It can be concluded that the administrative delimitation of the Netherlands performs, on average, equally well as the functional delimitation. The hypothesis that the municipalities within the administratively defined regions show less coherence than the municipalities within the functionally delimited regions, cannot be rejected.

Our results imply that it is important in delimitation studies to the test the functional divisions of regions against the administrative division with respect to a set of relevant indicators. If the functional divisions do not outperform the administrative division with respect to these indicators, there is not much to be gained in policy making by using a particular functional division. Other reasons for particular delimitations of regions, such as the existence of regional administrative and governmental bodies and the managerial control over regions, may then be more important. However, in the numerous delimitation studies for many different countries and regionalisation procedures there is hardly given any information on whether the functional divisions of regions can outperform the administrative division of regions. Therefore future research on this subject should include this information.

Finally, our results imply that it may be better for regional labour market policies not to use a highly differentiated division of regions for small countries like the Netherlands. In general, the regionalization of the Netherlands into four regions seems to be sufficient. On the one hand this conclusion is supported by the empirical study by Vermeulen (2005), who finds that the differences between Dutch regions in employment and unemployment rates are rather limited. On the other hand, Vermeulen (2005) finds larger disparities between regions for women, low age groups and lower educated. Moreover, according to Felsenstein and Portnov (2005a) there are no a priori arguments to expect that small countries will have less disparities between regions than larger countries. More research is required to find appropriate regionalisations of the labour market for both small and large countries.

References

AMEDO, D. (1968) An Optimization Approach to the Identification of a System of Planning Regions, *Regional Science Association* 23, 25-44.

ANDERSEN, A.K. (2002) Are Commuting Areas Relevant for the Delimitation of Administrative Regions in Denmark? *Regional Studies* 36, 833-44.

ATZEMA, O. and DIJK, J. VAN (2005) The Persistence of Regional Unemployment Disparities in The Netherlands, in D. FELSENSTEIN and B.A. PORTNOV (Eds) *Regional Disparities in Small Countries,* pp.147-67. Springer, Heidelberg.

BALL, R.M. (1980) The Use and Definition of Travel-To-Work Areas in Great Britain: Some Problems, *Regional Studies* 14, 125-39.

BAUMANN, J.H., FISCHER, M.M. and SCHUBERT, U. (1996) A Multiregional Labour Supply Model for Austria: The Effects of Different Regionalisations in Multiregional Labour Market Modelling, in FISCHER, M.M, BUTTON, K.J. and NIJKAMP, P. (Eds) *Regional Housing and Labour Markets*, pp. 366-96. Elgar Reference Collection, vol. 5, Cheltenham, U.K. and Lyme, N.H.

BONGAERTS, D., CÖRVERS, F. and HENSEN, M. (2004) *The Delimitation and Coherence of Functional and Administrative Regions*, Research Series 04O19, Ministry of Economic Affairs, The Hague.

BROWN, L.A. and HOLMES, J. (1971) The Delimitation of Functional regions, Nodal Regions, and Hierarchies by Functional Distance Approaches, *Journal of Regional Science* 11, 57-72.

CASADO-DÍAZ, J.M. (2000) Local Labour Market Areas in Spain: A Case Study, *Regional Studies* 34, 843-56.

COOMBES, M.G. and OPENSHAW, S. (1982) The Use and Definition of Travel-to-Work-Areas in Great Britain: Some Comments, *Regional Studies* 16, 141-49.

COOMBES, M.G., GREEN, A.E. and OPENSHAW, S. (1986) An Efficient Algorithm to Generate Official Statistical Reporting Areas: The Case of the 1984 Travel-To-Work-Areas Revision in Britain, *Journal of Operational Research Society* 37, 943-53.

COOMBES. M. (2005) Multiple Dimensions of Settlement Systems: Coping with Complexity in CHAMPION, T. and HUGO G. (Eds) *New Forms of Urbanization, Beyond the Urban-Rural Dichotomy*, International Union for the Scientific Study of Population, pp. 307-24. Ashgate, Hants/Burlington.

CORPELIJN, A.W.F. and HEERSCHOP, M.J. (2002) Inkomende en Uitgaande Pendel, 1998, *Sociaal-Economische Maandstatistiek*, no. 2002/12, pp. 53-57. Statistics Netherlands, Voorburg/ Heerlen.

EUROSTAT (1992) *Study on Employment Zones,* E/LOC/20, Luxembourg.

FELSENSTEIN, D. and PORTNOV, B.A. (2005, Eds) *Regional Disparities in Small Countries,* Springer, Heidelberg.

FELSENSTEIN, D. and PORTNOV, B.A. (2005a) Understanding Regional Inequalities in Small Countries, *Regional Studies* 39, 647-58.

FISCHER, M.M. (1980) Regional Taxonomy: A Comparison of Some Hierarchic and Non-hierarchic Strategies, *Regional Science and Urban Economics* 10, 503-37.

FOX, K.A. and KUMAR, T.K. (1965) The Functional Economic Area: Delineation and Implications for Economic Analysis and Policy, *Regional Science Association Papers* 15, pp. 57-85.

GIULLIANO, G. and SMALL, K.A. (1991) Subcenters in the Los Angeles Region, *Regional Science and Urban Economics* 21, 162-82.

GREEN, A.E. and COOMBES, M.G. (1985) Local Unemployment Rates: Statistical Sensitivities and Policy Implications, *Regional Studies* 19, 268-73.

HAZANS, M. (2004) Does Commuting Reduce Wage Disparities? *Growth and Change* 35, 360-90.

KILLIAN, M.S. and TOLBERT, C.M. (1993) Mapping Social and Economic Space: The Delineation of Local Labor Markets in the United States in SINGELMANN, J. and DESERAN, F.A. (Eds) *Inequalities in Labor Market Areas*, pp. 69-79. Boulder and Oxford Westview Press.

LAAN, L. VAN DER and SCHALKE, R. (2001) Reality versus Policy: The Delineation and Testing of Local Labour Market and Spatial Policy Areas, *European Planning Studies* 9, 201-21.

LEMAY, P. (1999) *The Statistical Analysis of Dynamics and Complexity in Psychology: A Configural Approach*, http://tecfa.unige.ch/~lemay/thesis/THX-Doctorat, University of Lausanne.

MASSER, I. and BROWN P.J.B. (1975) Hierarchical Aggregration Procedures for Interaction Data, *Environment and Planning A* 7, 509-23.

MASSER, I. and SCHEURWATER, J. (1980) Functional Regionalisation of Spatial Interaction Data: An Evaluation of Some Suggested Strategies, *Environment and Planning A* 12, 1357-82.

MUSTERD, S. and ZELM, I. VAN (2001) Polycentricity, Households and the Identity of Places, *Urban Studies* 38, 679-696.

NEWELL, J.O. and PAPPS, K.L. (2002) *Identifying Functional Labour Market Areas in New Zealand: A Reconnaissance Study using Travel-to-Work Data*, IZA DP No. 443, Institute of the Study of Labor, Bonn.

OECD (2002) *Redefining Territories, The Functional Regions*, Paris.

OECD (2005) *OECD Employment Outlook,* Paris.

ONS and M.G. COOMBES (1998) *1991-Based Travel-to-Work Areas,* Office for National Statistics, London.

OPENSHAW, S. (1984) *The Modifiable Areal Unit Problem*, Concepts and Techniques in Modern Geography, no. 38, Geo Books, Norwich.

RITSEMA VAN ECK, J., OORT, F. VAN, RASPE, O. DAALHUIZEN, F., BRUSSEL, J. VAN (2006) *Veel Steden maken nog geen Randstad (Many Towns do not a Randstad Make)*, Netherlands Institute for Spatial Research, NAi Uitgevers RPB, Rotterdam/Den Haag.

SFORZI, F., OPENSHAW S. and WYMER. C. (1997) Le Procedura di Identificazione dei Sistemi Locali del Lavoro in SFORZI, F. (Ed.), *I Sistemi Locali del Lavoro 1991,* pp. 235-42. ISTAT, Rome.

UNWIN, D.J. (1996) GIS, Spatial Analysis and Spatial Statistics, *Progress in Human Geography* 20, 540-51.

SMART, M.W. (1974) Labour market areas: uses and definitions, *Progress in Planning* 2, 238-353.

VERMEULEN, W. (2005) *Regional Disparities in Small Countries?* Paper presented at the ERSA conference 2005 in Amsterdam, Amsterdam/The Hague.

VERMEULEN, W. and OMMEREN, J. VAN (2006) *Compensating of Regional Unemployment in Housing Markets,* CPB Discussion Paper no. 57, The Hague.

WARD, J.H. (1963) Hierarchical Grouping to Optimise an Objective Function, *Journal of the American Statistical Association* 58, 236-44.

Appendix A: MFPT method

From Markov Chain to MFPT

To compute the MFPT matrix, we regarded our daily travel-to-work commuting flows of workers as a Markov chain. A Markov chain is a stochastic process that describes the transition from one state to another over time using probabilities. By using a Markov chain it is possible to re-enter each state at any point in time (i.e. $\sum_j P_{i,j} = 0$ for some i where $p_{i,j}$ is element (i, j) of matrix $lim_{t\to\infty}(P^t)$), and compute the average number of transitions needed to arrive from origin i in destination j for the first time. Note that because the probabilities to go from i to j are in general not equal to the probabilities to go from j to i by construction, theMFPT matrix is asymmetric. Since Masser and Scheurwater (1980) argued in favour of using a destination based instead of an origin-based probability matrix when handling travel-to-work flows, we use the destination-based approach throughout this paper.

Computation of the MFPT

For the Markov chain with a single-period transition matrix P, the j-period transition matrix is defined by P^t. If we let this process run for an indefinite time span, we end up in an equilibrium state. The proportions of time spent in each state are then $lim_{t\to\infty}P^t = A$. A is called the limit matrix. Having these two matrices, we can compute the so-called fundamental matrix Z of the process. Z can be computed by $Z = (I - (P - A))^{-1}$, where I is the identity matrix. From the limit matrix, matrix D is defined by $1/a_i$ on its diagonal and zeros for all other elements. The MFPT matrix can then be computed by $M = (I - Z + EZ_{diag})D$ where E is a matrix containing ones everywhere and Z_{diag} is the matrix containing the diagonal elements of Z and zeros for all other elements (see also Lemay, 1999).

Properties and interpretation of the MFPT matrix

First of all, the diagonal elements of the MFPT matrix are very small; this indicates that there are many travel-to-work flows within a region, something quite intuitive. Furthermore, all other values in the columns are relatively close to the column average, that is to say, they are of the same order. These column averages are indicators for how much attraction a region has to work in. The lower the column average, the more attractive the region is.

From MFPT to distances

The asymmetry that we observed for the MFPT matrix is particularly inconvenient to cluster regions, as clustering procedures often implicitly assume symmetric distances. Another problem arising from the MFPT is that the order of the column averages differs considerably among columns, which may result in the clustering of all larger regions together, even though the distance in kilometres between these regions is very large. In fact, we want to cluster the regions in such a way that the variation within clusters is minimal. Therefore we need appropriate measures of variation. The problem of differences in the order of column averages is solved by taking the z-values, which are defined by $z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}$. Note that although the diagonal values of this z-matrix can be computed, they make no sense and should be equal to zero or even nonexistent. From these z-values we obtain a measure for how close regions i and j are to each other. This is done with a so-called squared distance matrix. For each column k, we compute the difference between z_{ik} and z_{jk} (where $i,j \neq k$) and square it. This is the marginal contribution from k to the squared distance. In formula this can be written as:

$$d_{i,j} = \sqrt{\left(\sum_{k:i,j \neq k} (z_{ik} - z_{jk})^2 \right)}$$

By construction, d_{ij} is equal to d_{ji}, so this transformation also handles all other problems.

Table 1

Overview of statistics of the four administrative regions of the Netherlands, 2001

Administrative regions	Income level €	Std. Dev.	Housing prices €	Std. Dev.
North	17,688	3,154	47,774	9,992
East	17,362	3,617	66,402	10,432
West	18,247	4,624	71,312	20,969
South	17,747	4,138	70,785	10,909
Total	17,874	4,143	67,051	17,611

Administrative regions	Employment rate %	Std. Dev.	Unemployment rate %	Std. Dev.
North	61.20	3.54	5.01	2.16
East	64.49	3.87	3.31	1.33
West	66.19	4.29	3.07	1.37
South	64.16	3.52	3.17	1.44
Total	64.67	4.22	3.39	1.62

Source: Statistics Netherlands

Table 2

Overview of statistics of the four functional regions of the Netherlands, 2001

Functional regions	Income level €	Std. Dev.	Housing prices €	Std. Dev.
Zeeuws Vlaanderen	17,893	2,909	38,721	3,780
South	17,911	4,166	69,043	13,645
Utrecht	17,974	4,170	84,961	22,598
North	17,775	4,107	60,984	16,899
Total	17,874	4,143	67,051	17,611

Functional regions	Employment rate (%)	Std. Dev.	Unemployment rate (%)	Std. Dev.
Zeeuws Vlaanderen	62.53	4.20	4.90	2.00
South	65.07	4.13	3.20	1.41
Utrecht	66.77	4.17	2.74	1.58
North	63.74	4.18	3.76	1.75
Total	64.67	4.22	3.39	1.62

Source: Statistics Netherlands

Table 3

The mean differences of the four economic indicators for the four administrative regions (NUTS1) of the Netherlands, 2001

		Income level		Housing price	
		Mean dif.	Sig.	Mean dif.	Sig.
(I) Region	(J) Region	(I-J)	p-value	(I-J)	p-value
1) North	2) East	325	0.622	-18,625*	0.000
	3) West	-560	0.343	-23,535*	0.000
	4) South	-60	0.926	-23,011*	0.000
2) East	3) West	-884	0.079	-4,909*	0.011
	4) South	-385	0.497	-4,385*	0.044
3) West	4) South	499	0.303	524	0.778
Number of sig. dif.	Out of 6	0		5	

		Employment rate		Unemployment rate	
		Mean dif.	Sig.	Mean dif.	Sig.
(I) Region	(J) Region	(I-J)	p-value	(I-J)	p-value
1) North	2) East	-3.292*	0.000	1.698*	0.000
	3) West	-4.996*	0.000	1.939*	0.000
	4) South	-2.967*	0.000	1.835*	0.000
2) East	3) West	-1.704*	0.003	0.241	0.270
	4) South	0.325	0.613	0.136	0.578
3) West	4) South	2.029*	0.000	-0.105	0.218
Number of sig. dif.	Out of 6	5		3	

* = Significantly different at the 5%-level

Table 4

The mean differences of the four economic indicators for the four functional regions of the Netherlands, 2001

(I) Region	(J) Region	Income level Mean dif. (I-J)	Income level Sig. p-value	Housing price Mean dif. (I-J)	Housing price Sig. p-value
1) Zeeuws Vlaanderen	2) South	-17	0.992	-30,319*	0.000
	3) Utrecht	-82	0.964	-46,237*	0.000
	4) North	117	0.946	-22,262*	0.001
2) South	3) Utrecht	-65	0.924	-15,918*	0.000
	4) North	134	0.741	8,058*	0.000
3) Utrecht	4) North	199	0.773	23,976*	0.000
Number of sig.dif.	Out of 6	0		6	

(I) Region	(J) Region	Employment rate Mean dif. (I-J)	Employment rate Sig. p-value	Unemployment rate Mean dif. (I-J)	Unemployment rate Sig. p-value
1) Zeeuws Vlaanderen	2) South	-2.532	0.297	1.697	0.066
	3) Utrecht	-4.235	0.094	2.164*	0.025
	4) North	-1.203	0.620	1.143	0.216
2) South	3) Utrecht	-1.702*	0.047	0.467	0.151
	4) North	1.329*	0.008	-0.555*	0.004
3) Utrecht	4) North	3.031*	0.001	-1.021*	0.002
Number of sig. dif.	Out of 6	3		3	

* = Significantly different at the 5%-level

Table 5

Percentages of significant differences (at the 5%-level) between the means of the economic indicators, delimitations of 2001

Indicator	4 region division		12 region division	
	Administrative	Functional	Administrative	Functional
	%	%	%	%
Income level (2001)	0	0	0	27
Housing price (2001)	83	100	74	67
Employment rate (2001)	83	50	56	33
Unemployment rate (2001)	50	50	50	36
Total	54	50	45	41

Indicator	28 region division		40 region division	
	Administrative	Functional	Administrative	Functional
	%	%	%	%
Income level (2001)	3	5	2	8
Housing price (2001)	58	62	55	49
Employment rate (2001)	38	36	24	20
Unemployment rate (2001)	28	40	29	24
Total	32	36	28	25

Table 6

Percentages of significant differences (at the 5%-level) between the means of the economic indicators, delimitations of 1991

Indicator	4 region division		12 region division	
	Administrative	Functional	Administrative	Functional
	%	%	%	%
Income level (1991)	0	0	0	8
Income level (2001)	0	17	0	18
Housing price (2001)	83	83	74	51
Employment rate (2001)	83	67	56	27
Unemployment rate (2001)	50	50	50	36
Total	43	54	36	28

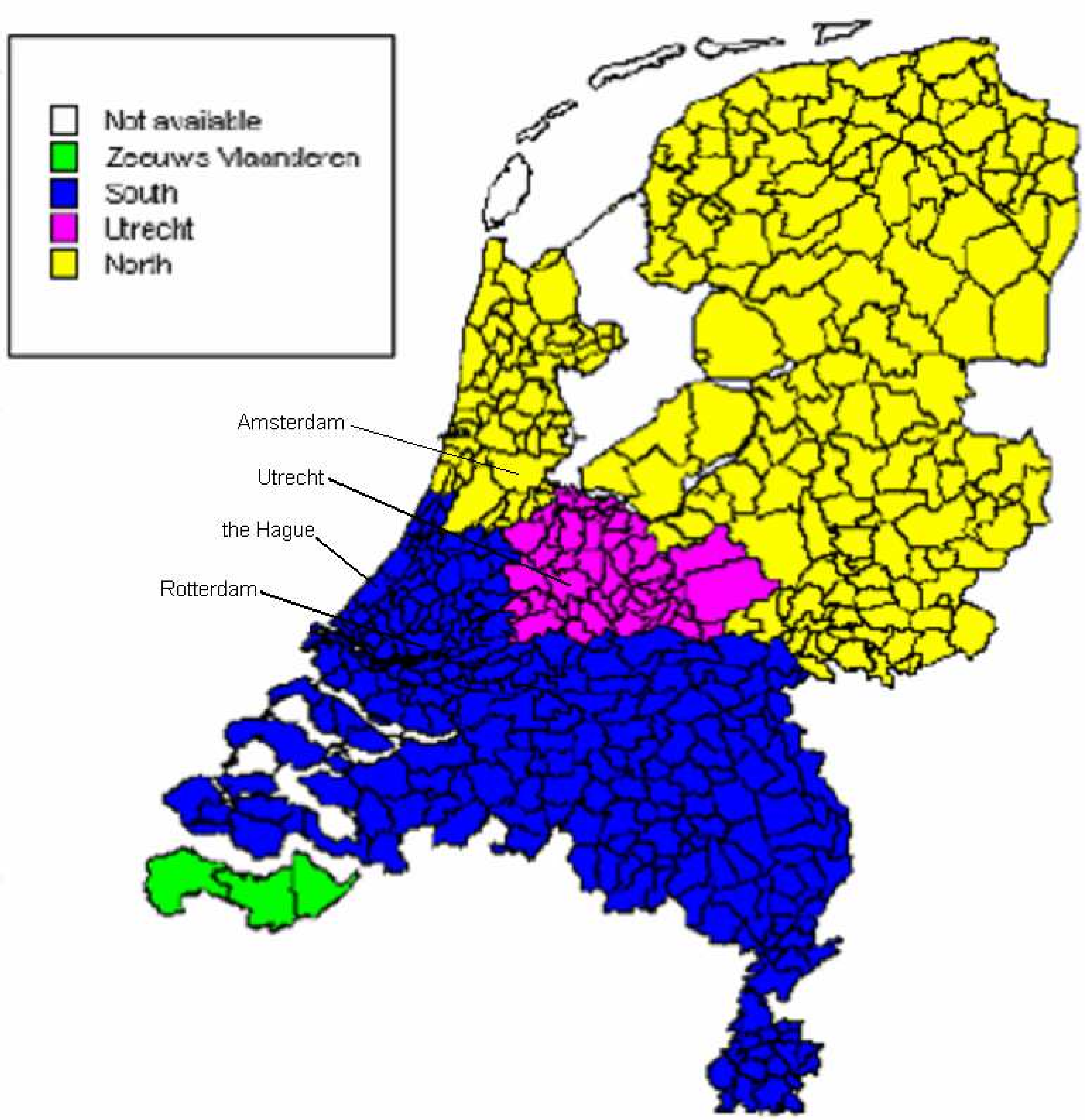
Indicator	28 region division		40 region division	
	Administrative	Functional	Administrative	Functional
	%	%	%	%
Income level (1991)	6	3	2	10
Income level (2001)	3	8	2	6
Housing price (2001)	58	50	55	56
Employment rate (2001)	38	28	24	29
Unemployment rate (2001)	28	21	29	29
Total	27	22	22	26

Figure 1

The four administrative regions (NUTS1) of the Netherlands

Figure 2

The four functional regions of the Netherlands, 2001



[1]. This problem refers to the 'scale problem' mentioned in the introduction.

[2]. This problem refers to the 'aggregation problem' mentioned in the introduction.

[3]. See Corpelijn and Heerschop (2002) for more details on the commuting flows in the Netherlands.

[4]. In 1991 and 2001 two car ferries were running across the Westerschelde estuary. In 2003 the Westerschelde tunnel was put into use and the car ferry services were stopped.

[5]. In this paper we abstract from other factors that may determine wage inequality between regions, like regional differences in the educational structure of the population and the economic structure of regions with respect to sectors of industry and occupations. Also the analysis of the impact of national collective agreements on reducing wage inequality is beyond the scope of the paper.