Paper presented at the ECOMOD conference
Istanbul July 3 - 5 2003

A regional labour market model for the Netherlands

by: Eugene Verkade and Wouter Vermeulen *
CPB, Netherlands Bureau for Economic Policy Analysis
PO BOX 80510
2508 GM The Hague
The Netherlands
Email : e.m.verkade@cpb.nl , w.vermeulen@cpb.nl

The CPB is building a regional labour market model for the Netherlands. This model will be used to construct long term scenarios and to analyse labour market developments with a regional component. The focus is on the interaction between regional population and employment, against the background of given national developments. The regional unit is the Eurostats NUTS 3 level.

Regional employment follows a shift share approach. The share component depends on the sectoral composition (18 sectors) of the regional economy. The shift component depends on demographic and geographic variables (locations factors). Population is divided into gender and 7 age classes. Commuting, migration, labour participation rates and unemployment form important mechanisms to equilibrate discrepancies between regional population and employment.

*With special thanks to Carel Eijgenraam for useful comments and Jelte Haagsma for data processing

1 Introduction

Policy analyses with a spatial component have attracted renewed interest in the Netherlands. For example, considerable attention is paid in recent years to the consequences of a number of proposed large infrastructural projects, like building high speed railway connections within the Netherlands or connecting Dutch railways to the European system. Other issues are related to the strong urbanisation in the west of the Netherlands or the regional issues in the Northern provinces. On a lower spatial scale there are the issues on larger cities. All these studies have raised the demand for up to date regional economic models.

With intermediate intervals of 5-7 years the CPB, Netherlands Bureau for economic Policy Analysis, produces long term studies for the Dutch economy. Since the beginning of the nineties the focus has been shifted from forecasting to scenario building. According to CPB, long term studies should do justice to the fundamental uncertainty about the future by developing several alternatives. A small number of scenarios, each one internally consistent but quite different from the others, should cover an useful and reliable margin of uncertainty for policy makers. In the latest long term scenarios, published in 1997, renewed attention was paid to regional and spatial developments, but their quantitative elaboration on these topics was rather rudimental. Quality improvements can certainly be made by use of economic model. A first clear and straightforward improvement can be obtained through a disciplinary definition framework of an economic model. Further improvements can be made by incorporating economic behaviour.

CPB used to have a regional model some 20 years ago (Suyker(1981),Kwaak (1985)). After regional analysis became less popular for some time, the model has been shifted away to outside CPB. Since then economic theory about regional modelling has progressed, new techniques for spatial econometrics became available, while computer facilities improved.

2 Purpose and position of the model

There exists a variety of regional models to analyse these issues. Notably, an applied spatial general equilibrium model for the Netherlands is under construction (Oosterhaven et al. (2001)), in which the Netherlands is divided in 40 regions. There also exists some models with a much higher detailed regional scale to advise provinces and municipalities. At the CPB a new Spatial and Regional Economics Unit has been created. From its start this unit has expressed the ambition to build a new econometric regional labour market model for the Netherlands, replacing an existing simple shift-share model. The new model should generate time paths for regional developments for the medium and longer term, while at the same time tracking fully with national developments as generated elsewhere in the CPB. The new CPB regional model intends to incorporate some of the latest economic insights about model specification. Current computers hardly put any restriction anymore on complexity or number of equations. Advanced econometric techniques allow for complex estimation procedures and the use of spatial interaction models, which can be production constraint, attraction constraint o doubly constraint. However, to keep the model manageable and the regional forecast plausible the number of regions should remain limited.

By its medium and long term scope, the regional aggregation level and its dynamic structural specification, this model would fill a gap in the set of existing instruments for regional analysis. The new model will be used for simultaneous regional forecasts of population, labour force, employment, unemployment, commuting and migration. These results will be input for (internal and external) estimates for demand for land use (housing and working), demand for mobility and regional welfare. Besides long term ‘forecasting' the model should be suitable for analyses with a regional component.

Top down approach

The national developments are considered to be exogenous. National developments for the short term, the medium term or the long term, are constructed elsewhere within the CPB. Adding up the regional variables to the national level the model should reproduce these national forecasts. The model will follow a top down distribution.

Sector breakdown

The model will make use of sector information. The CPB Athena model generates sector forecasts for 18 branches of industries[1]. The model will take this sector information into account. Regional sector structure is often expected to be an important determinant of regional employment growth. We will see that in the Netherlands during the last decades this is only the case in a limited way. That does not mean we are going to leave out all sector information. The sector composition of regional employment remains important for the regional claims on land and therefore we want to keep it inside the model.

Regional break up

The regional unit follows the Eurostat NUTS-3 classification. The corresponding Dutch classification is the COROP region. It is the lowest regional level, for which Netherlands Statistics constructs data, which are consistent with the National Accounts. The Netherlands is divided in 40 NUTS-3 areas with an average population of 400 thousand people. A Dutch

[1] The 18 branches of industry in the Athena model are agriculture, 4 manufacturing industries (food processing industry, chemical industry, metal industry, other manufacturing) , mining, oil refineries, construction, 9 service industries (trade, transport, telecommunications, banking & insurance, exploitation of real estate, commercial services, household services, health & care services) and the government sector

COROP area is not an administrative unity in the Netherlands, like provinces and municipalities. Its area is built up of municipalities, which are located within the same province. Its boundaries have been chosen on the basis of coherency between living and working population. These municipalities are added together according to a hierarchic clustering method. In the joint area the percentage of the residents, who also work there and the percentage of the workers, who also live there, should both exceed a certain threshold percentage. On average these percentages are around 70%. Automatically that implies that on average around 30% of the population commutes between the chosen regions

Personal characteristics

Population is classified by personal characteristics. These characteristics are relevant for decisions on labour supply, commuting and migration. The most relevant characteristics are gender and age. To keep the model manageable we distinguish seven age groups. The working population between 15 and 65 is divided in 5 age groups of 10 years each. Other important characteristics with respect to the labour market are education level and household composition, but we take these only indirectly into account. The main reason is a lack of reliable data on the regional level. Besides, expanding the model ambitiously in too many directions can make it fastly too complex to manage. For the same reason, we neglect ethnicity for the time being, although the population increase in the long term, like elsewhere in Western Europe, is expected to come mainly from people, who either themselves or their parents are born elsewhere in the world.

3 Bird's eye view of the model

The central variables are the regional employment and regional population. The regional developments of employment and population will influence each other. Regions with faster growing population may expect to attract and generate more jobs. Regions with a fast growing employment may be an attractive destination for migration. Employment and population will be determined simultaneously. There are a number of interconnections between these two variables, but their strength may gradually change over time.

The model will generate time paths. It is not likely that in each region employment and population will grow in line automatically. In the short term additional commuting between nearby regions can bridge part of the gap. Commuting allows some diversion between employment and population. But distance causes limitations on commuting, although gradually falling transportation costs, both in money and time, bring regions 'closer' together.

Another mechanism to absorb shocks in regional employment growth without changes in population size is via the regional labour participation rate. In that case the regional population size remains unaffected, but more people enter the labour market and accept a job, increasing the regional working population size. Participation rates have increased sharply during the last decade in the Netherlands, especially among women in middle and higher age groups. Regional participation has increased as well during this period, but their levels still vary among regions, with some regions approaching ceiling levels for certain age groups. The population size in the Netherlands will hardly grow anymore in the near future and may even fall in the long term. However, overall national participation rates are expected to increase further. Given the fact that certain regions have nearly reached ceiling level for certain age groups this means a deviation in regional growth potentials. A good description of the regional participation rate becomes important

As commuting distances increase, migration becomes more and more an option. Part of the commuters may move to the work region after some time. This so called labour induced migration implies that people will follow jobs. In the Netherlands labour induced migration is mostly related to persons with a job in the destination region. There maybe also some opposite movement. As income increases workers may move to a more attractive region to live in and become commuters. Commuting and migration will be partly related then. In the model a direct relation is not incorporated so far.

Jobs will follow people as well. A faster growing regional population will in itself generate faster employment growth to match the average level of facilities (shops, medical services etc). A faster growing population can also be an attractive factor for new business locations in other sectors as agglomeration forces may appear.

In the long run equilibrium tendencies may appear. Regional employment does not have to match regional population everywhere in the same way, as commuting flows and differences in participation rates may occur. Regional differences in participation rate between age groups may gradually phase out (section 4.2). Overall differences in participation rate will mainly be attributed to regional population composition. Commuting flows will keep on increasing over time (section 4.3).

A more useful indicator for regional disparities seems to be the regional unemployment levels, scaled by the regional labour force. These regional unemployment rates will operate as a local ‘tension' variable to bring living and working closer together. People in regions with a relatively high unemployment rate will have to look further away to find a job compared to regions with a low unemployment rate. Differences in regional unemployment rate will influence commuting flows (section 4.3). It will also have an impact on (labour induced) migration (section 4.4).

Figure 3.1 Overview regional model

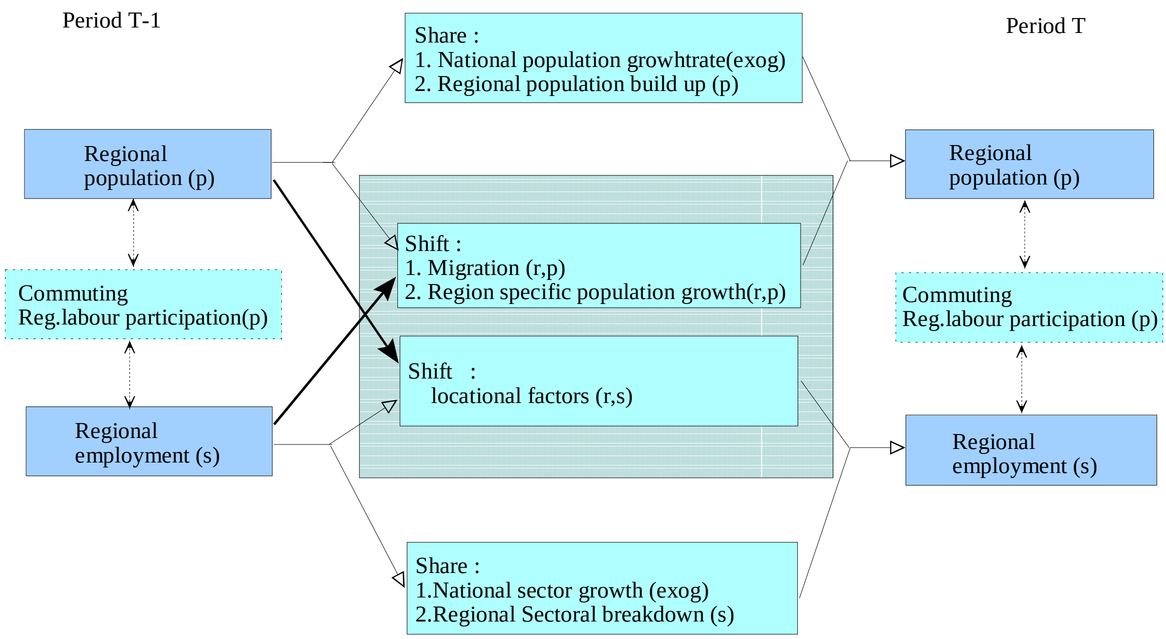


Figure 3.1 gives an overview of the model. The figure indicates how the core variables population and employment elapse over time. The left hand side stands for period (T-1) and the right hand side for period T. Both variables are determined in an equation, which is formed by a sum of a shift and a share component.

The share component describes the impact of national developments. In case of employment the share component includes the national growth effect and the regional sectoral composition effect. The idea is that regions with a high share of fast growing sectors on a national level will grow faster. This would automatically be true, if growth per sector is homogenous for all regions. In case of population the share component confronts the national ageing developments with the regional population build-up. Regional build–up can differ for example because of the presence of education facilities. Because national developments are given, the share parts are exogenous in the model

The shift component describes the purely regional circumstances. In case of employment favourable location or agglomeration effects can make employment grow faster in certain regions. In the case of population attractive living conditions can make regional population grow faster through (net) immigration. The shift components enclose the interactions between population and employment. Therefore the shift parts form the core of the model and encompass all the simultaneities.

4 Model description

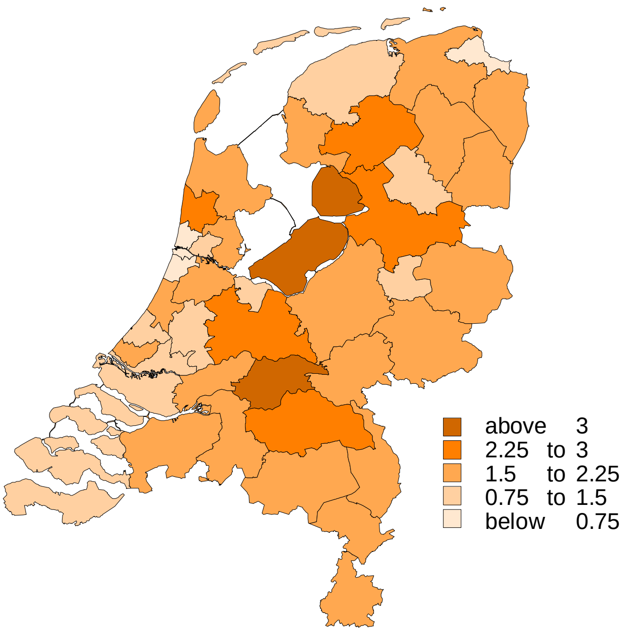
4.1 Regional employment

Regional employment in the model has been determined through a shift share equation. The shift component measures the impact of the national level sectoral developments. The shift component takes the locational advantages into account. Differences between regional employment growth rates have been considerable in the recent past. Differences in regional sectoral decomposition do not explain this observation. This puts extra weight on a good perception of the shift component. In the current model version we use regional population and historic trends in the sectoral shifts. In future model versions we will take other locational factors into account.

4.1.1 Some stylised facts on regional employment in the Netherlands

Netherlands Statistics provides rather detailed sectoral information on regional employment. The raw data distinguishes 37 branches of industry (CBS 2001a). Data are available for the period 1987-2000. The data refer to the wage earners only. The self employed, which accounts for around 10% of the labour volume on a national level, are not included. The sectoral regional

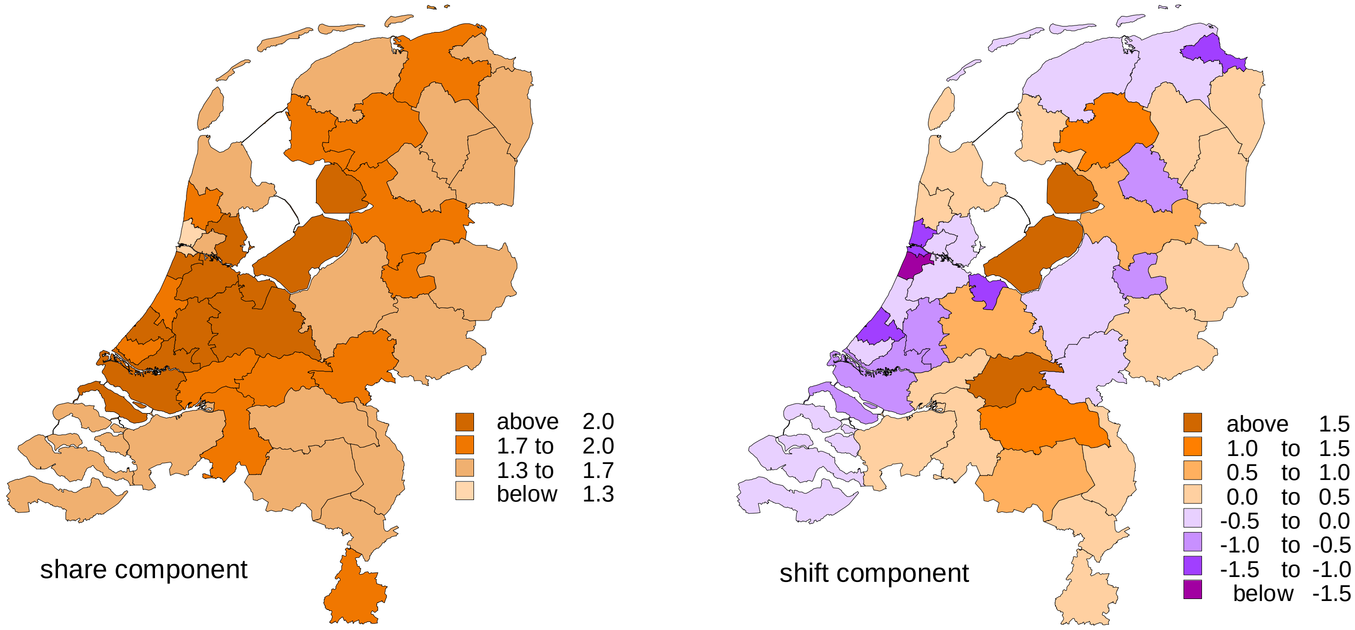
Figure 4.1 Average annual regional employment growth wage earners (1987-2000)



data form a valuable historic source. The sectoral breakdown is in particular essential for the estimation of regional claims on land from employment, produced elsewhere within the CPB. In our further analysis and in the model we reduce the number of industries to 18 in line with the CPB Athena model (see section 2).

There have been considerable differences in employment growth between regions. Figure 4.1 shows the average yearly growth rate of employment of wage earners over the 1987-2000 period. On average, employment in the Netherlands grew with slightly under 2% annually during this period. Employment growth was the highest (above average) in the centre regions of the country and some adjoining regions.

Figure 4.2 Share and shift component of the regional employment growth of wage earners (1987-2000)



In figure 4.2 the average regional growth rate is split in a share (left) and a shift (right) component. The share component is calculated assuming homogenous sectoral growth over the country. In other words, the regional sectoral employment growth follows the national sectoral growth fully. During the 1987-2000 period employment in the agricultural and manufacturing sectors shrank while employment in a number of service sectors especially in the later years boomed. The homogeneity assumptions generate only some differentiation between regions. Sectoral decomposition was rather favourable in the densely populated West including the centre province of Utrecht, where nearly half the Dutch population lives.

The shift component is calculated as the residual. Does the sectoral decomposition contribute in explaining part of the regional growth differences? It looks hardly the case. Growth in the West lagged well behind the national average, despite a favourable sectoral decomposition. In most parts of the eastern and the southern regions employment growth performed rather well, notwithstanding an unfavourable sectoral breakdown. However, the performance of the centre provinces of Utrecht and Flevoland suggest that they have the best locational advantages during this period.

It is known from literature that the shift share outcomes are sensitive to the detail in sectoral classification. Our 18-sector decomposition puts extra weight on a perception of the shift component. What are the factors behind these shifts and how stable are they over time? In the current model version we will only use a rather simple description of the shift component. Further research will certainly be needed. In particular it is worthwhile to look further into the direction of the causality between regional employment and regional population , for example like in Carlino & Mills (1987).

4.1.2 The regional employment equations

In the regional employment equations the region is indicated with a subscript. The subscript NL stands for the national total of the Netherlands. Employment consist of wage earners and self-employed. On a national level nearly 90% of the labour is related to wage earners.

Regional employment in labour years

The regional labour data from Netherlands Statistics (CBS (2001)) are expressed in labour years of wage earners only[2]. We indicate employment in labour years with an E with a score on top. When we later convert the labour years into numbers of persons to make regional labour market accounts, we skip the score above the E. On a regional level, we have:

$$\overline{E}_r = \overline{EW_r} + \overline{ES_r} \quad (1)$$

where : $\bar{E}_r$ = total employment in labour years in region r

$\bar{E}W_r$ = employment in labour years of wage earners in region r

$\bar{E}S_r$ = employment in labour years of self-employed in region r

Regional employment is broken down into 18 sectors. The sectors are indicated by a superscript.

$$\overline{E}_r = \sum_s \overline{E}_r^s \quad (2)$$

where : $\bar{E}_r^s$ = employment in region r in sector s

The regional scale determines the nature of the relevant regional variables. The NUTS-3 classification in the Netherlands implies a mixture, varying from relatively thinly populated

[2] Regional data on sectoral employment of self employed in labour years has been estimated by CPB

areas in the north to densely populated areas around the major cities in the west. In smaller regions it becomes more difficult to recognize national sectoral developments. We then have to rely more on local circumstances.

During the last decades the employment growth in the manufacturing sector in the relatively densely populated West of the country (the area around the 4 major Dutch cities Amsterdam, Rotterdam , The Hague and Utrecht) lagged well behind in favour of the adjoining regions in the East en South. Lack of space and (corresponding) higher land prices will play a role. One may expect that these developments will continue, may be gradually at a lower rate. Employment growth in a number of service industries, in particular trade and medical services, may also be influenced by the growth in the population size in the concerning and adjoining regions. Finally, a spreader type of variable should guarantee that regions sum up to the (exogenous) national total.

The employment equations are expressed in annual growth rates. A growth rate is indicated by an ° above the name. So far we get :

$$\overset{\circ}{\overline{EW}}{}_r^s = \overset{\circ}{\overline{EW}}{}_{NL}^s + f(\overset{\circ}{POP}, (\overset{\circ}{\overline{EW}}{}_r^s - \overset{\circ}{\overline{EW}}{}_{NL}^s)_{t-1}, other) \tag{3}$$

In later model versions we may introduce other spatial variables, which are related to the location of the region. One can think of the locations near economic centres of production or densely populated areas or locations near the Dutch main ports of the Rotterdam harbour and the Schiphol Airport. Cluster effects can play a role as well: sectors which are well represented in a region can give a reason for faster growth.

Self employed are mainly concentrated in a few industries, notably agriculture, trade and other services. In all industries the share of the self employed is below 15%, with the exception of the agricultural sector. We assume that the growth rate of the self employed follow those of wage earners. Hence :

$$\overset{\circ}{\overline{ES}}{}_r^s = \overset{\circ}{\overline{EW}}{}_r^s \tag{4}$$

Regional employment in persons

Employment in persons[3] is between 10 to 20% higher than in labour years because of the large appearance of part time jobs in the Netherlands, in particular in some parts of the service sector. We know for each sector the conversion factor only on a national level. In general, this conversion factor is higher for the self employed, which includes for example participating

[3] Persons with a job of at least 12 hours a week

family members. We assume the conversion factors are more or less homogenous over the regions. There is no indication that there is much regional variation. We then get:

$$cw_r^s = cw_{NL}^s \tag{5}$$

$$cs_r^s = cs_{NL}^s \tag{6}$$

$$E_r = \sum_s (cw_r^s \,.\, \overline{EW}_r^s + cs_r^s \,.\, \overline{ES}_r^s) \tag{7}$$

where: cw_r^s = conversion factor wage earners volume to persons, region r and sector s

cs_r^s = conversion factor self employed volume to persons, region r and sector s

E_r = employment in number of persons in region r

4.2 Regional population and labour supply

Regional population is classified in 14 categories, according to gender and age. Regional population developments per category follow a national and a regional component. The national aging developments are exogenously translated to the regions. We use information from an external model, which describes the regional population in 1 year cohorts and uses region specific birth and death rates. Migration is described in a section 4.4.

Population numbers are translated into labour force through so called labour participation rates, On a national level these participation rates differ considerably among population categories. Regional participation rates follow national trends as long as they have not reached ceiling levels.

4.2.1 Some stylised facts on regional population and labour supply

National data as well as regional data on population are available for a long historical time period. We take into account 14 age categories, according to gender and age. The working population between 15 and 64 is split into 5 age cohorts of 10 years. These cohorts are in general sufficiently homogenous, with respect to supply to the labour market.

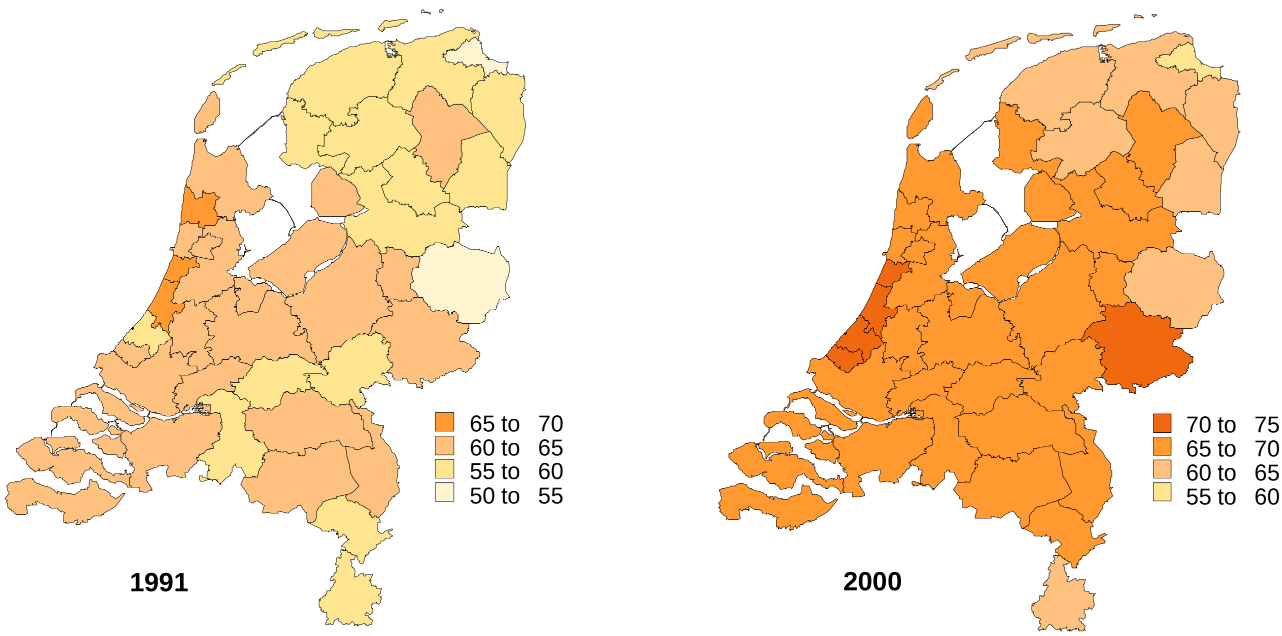
Table 4.1 Population developments, index with 1960 = 100

	North	East	West	South	Total
1960	100	100	100	100	100
1970	111	118	111	118	114
1980	123	134	116	132	123
1990	126	147	122	139	130
2000	131	160	129	148	139

Table 4.1 gives an overview of the population growth in four parts of the Netherlands[4]. In 1960 the Dutch population accounted for 11,4 million people, of which around 50% lived in the West, 20% in the East and in the South and the remaining 10% in the North. Population growth was below average in the relatively thinly populated North and densely populated West and well above average in the East and the South. Although there has been some small variation in regional aging processes, the main differences have to be caused by net migration.

The regional labour force is defined as the sum of the employed and the unemployed[5], both expressed in number of persons. The employed regional labour force is calculated as the regional employment corrected for net commuting. The relation between regional population and labour supply is expressed in the gross participation rate, which is defined as the labour force divided by the population level. Figure 4.3 shows that the overall participation rate (all age categories) varies among NUTS-3 regions, but the variation has become smaller over time.

Figure 4.3 Overview of the overall participation rates in NUTS 3 areas for 1991 and 2000



The demographic composition of the population (gender and age) may explain part of the differences in overall participation rates. Other important features of the regional labour force are schooling levels and household composition, in particular with respect of willingness to commute. These features will also be taken into account.

[4] For a description of these 4 regions, see appendix

[5] In the Netherlands a persons is counted as unemployed, if he has no job, is looking for one and immediately available for at least 12 hours a week.

4.2.2 The regional population equations

The increase in the regional population is assumed to consist of three components. The first one is the natural growth, which encompasses births, aging and deaths. The second one is net domestic migration and the third one net foreign migration.

National demographic projections are exogenous. That means that on a national level both the aging process as well as the total level in- and outflow of migrants to abroad is given for the 14 categories. The regional aging parameters may for a larger part follow the national ones, but there is some space left for regional variation. Fertility rates may differ between urban and rural areas and between autochthonous and allochthonous habitants. A proper year on year accounting of the composition of the population actually requires a description with 1 year cohorts, but this would increase the size of the model enormously. Therefore, we have chosen a second best solution. All regional aging components are kept exogenous in the model, but their values are obtained from interactions with a detailed regional population model. Regional population is calculated by:

$$\Delta POP_r(g,a) = NG_r^{ex}(g,a) + \left(\sum_s M_{s,r}(g,a) - \sum_s M_{r,s}(g,a)\right) + \left(EI_r^{ex}(g,a) - EO_r^{ex}(g,a)\right) \quad (8)$$

where: $POP_r(g,a)$ = population in region r, by gender and age
$NG_r(g,a)$= natural growth in region r, by gender and age
$M_{r,s}(g,a)$ = domestic migration , from region r to region s by gender and age
$EI_r(g,a)$ = migration from abroad to region r by gender and age
$EO_r(l,s)$ = migration to abroad from region r by gender and age

Migration, both domestic and foreign is described in section 4.4.

4.2.3 The regional labour force equations

A number of personal characteristics are relevant for the supply decision on the labour market. The most important ones are gender and age. Labour participation rates among women are generally much lower than those of men, in particular in the higher age cohorts. But the differences become smaller and have nearly been vanished in the lower age cohorts. Other characteristics are education level and ethnicity. The regional labour force is derived form the regional population, using age and gender specific labour participation rates. These rates also differ among regions. The national levels are given. Regional circumstances like unemployment, fertility rates, wage level, part timers may play a role (Van der Knaap e.a. (1995)). So far we have not taken these into account. Instead we follow a simple calculation rule. Regional participation rates follow the direction of the national one. There maybe ceiling levels, which can not be exceeded. A spreader should guarantee consistency. Because employment and unemployment size on a national level are exogenous, so is the national labour force implicitly. Therefore, participation rates are increased in age classes and in regions, where they have not reached a ceiling yet, until the provoked labour supply matches the required level.

$$L_r(g,a) \quad = lp_r(g,a).POP_r(g,a) \tag{9}$$

$$\Delta lp_r(g,a) = \Delta lp_{NL}(g,a) + \alpha_r(g,a).(\Delta u_{NL} - \Delta u_r) + \beta_r(g,a).(lp_r(g,a) - lp_r^{\max}(g,a)) + spreader \tag{10}$$

$$lp_r(g,a)_t = \min(lp_r(g,a)_{t-1} + \Delta lp_r(g,a)\ ,\ lp_r^{\max}(g,a)) \tag{11}$$

$$spreader = spreader + (\sum_r U_r - U_{NL}^{ex}) / L_{NL} \tag{12}$$

where $L_r(g,a)$ = regional labour force, by gender and age

$lp_r(g,a)$ = regional labour participation rate, by gender and age

$lp_r^{max}(g,a)$= ceiling level for the regional participation rate, by gender and age

u_r = regional unemployment rate

For the time being we have calibrated α on .2 and β on .3 for all regions. These values seem to be rather plausible for the Netherlands.

4.3 Commuting and regional unemployment

In the model we generate a full commuting matrix. The commuting flow equations are based on estimation results of a spatial interaction models (Fotheringham & O'Kelly, 1989). The model assumes an ongoing decreasing influence of the distance in kilometres over time (distance decay), reflecting further improvements in infrastructure, which has caused distances in travelling time to become smaller. Demographics will play an additional role (Rouwendal & Rietveld, 1994). However, estimation results have shown only a demonstrable influence of the schooling level.

4.3.1 Some stylised facts on Dutch commuting behaviour

Data on commuting come from the labour force survey (EBB) of Netherlands Statistics. Each year, this institute questions about one percent of Dutch households. The interviewees indicate personal characteristics (notably, gender, age, ethnicity, level of education, etcetera) and their employment status. Also, the NUTS3 region where they live and the region where they work are known. The period from 1992 to 2001 is covered. The sample is blown up to a full commuting matrix, which matches national data. However, when more personal characteristics are added cells can become blank for lack of records or for privacy reasons of the interviewees.

Commuting flows and distance

The relation between the size of commuting flows and distance (between regions) is a key issue in the model. We use a distance matrix that has been constructed under the authority of AVV Transport Research Centre. This institute has a large dataset of traffic flows in 1995. A subset has been taken that consists of all work trips by car. The average distance of these trips from one region to another is considered the distance between these regions. This method also yields a measure for average distance within (or for size of) a region.

The far majority of he intraregional commuters commute between 25 and 75 km. Above this distance the number of commuters falls drastically.

The interregional commuting surge

Between 1992 until 2001 employment and labour force have increased by 19% and 16% respectively, but the number of (interregional) commuters (in our definition) has increased by 58%. In fact the sharp surge in commuting started already in the mid 80s (Ekamper and van Wissen (2000)). The increase in commuting does not imply a similar increase in average distance from home to work. Much has to do with the typical size of a region. A gradually increase in home to work distance may initially be absorbed within the region. At a certain moment it will have to cross region borders. It is clear that the distance deterrence in commuting behaviour has decreased over the past decade (Vermeulen (2003a)).

To some extent, the rise in average commuting distance should be attributed to the business cycle. The second half of the nineties is characterised by a considerable upswing and it is a common observation that the number of job changes is higher under favourable economic developments. Rouwendal and Rietveld (1994) introduce job search as a theoretical background for the analysis of commuting behaviour, they find that people who change job are likely to commute over a longer distance after the change occurred.

Commuting flows and personal characteristics

Table 4.2 gives an overview of the average home to work distance (all workers) and the percentage of the working labour force that does not work in their region of residence (commuters) to these personal characteristics. On average, women commute over shorter distances than men. The pattern over age groups reveals that distances for the age group 15 - 24 are relatively short. A more pronounced relation is between commuting and education level. The well-educated are more specialised and can be expected to have more specific job preferences, for which they have to search in a larger area. Another explanation could be that the well-educated have relatively high incomes and consequently make a different trade-off between residential quality and commuting cost. Some distribution effects may be present. For example, the average level of education has risen over the past decades, so younger people are generally higher educated, which somewhat obscures the above relations.

Table 4.2 Average distance from home to work and percentage of commuters to characteristics (year 1998)

Personal characteristic	share (in %)	average distance	percentage of commuters across NUTS-3 borders
Total	100	29.0	19
- male	61	30.0	21
- female	39	27.3	16
- age group 15-24	12	27.8	14
- age group 25-34	31	30.4	23
- age group 35-44	28	28.9	20
- age group 45-54	23	28.3	17
- age group 55-64	7	27.6	15
- lower education	28	26.7	12
- middle education	44	28.4	18
- higher education	28	32.1	28

Commuting flows and households composition

Rouwendal and Rietveld (1994) have studied commuting distances of households in the Housing Demand Survey (WBO). Clearly, one-person households have shorter commuting distance than heads of household of larger households. An explanation put forward is that these households can more easily adapt their housing situation to their work location. However, the authors do not find much evidence in favour of the hypothesis that households with both the head and the partner employed have larger commuting distances. In our model, we have disregarded household characteristics.

4.3.2 The commuting model

Assume that the number of people C_{rs} who live in region r and work in region s depends on the size of the labour force L_r in r and the size of employment E_s in s. The commuting flow will decrease with the distance d_{rs} between the two regions. A simple form is then:

$$C_{rs} = L_r^{\alpha} \, . \, E_s^{\beta} \, . \, d_{rs}^{\gamma} \qquad (13)$$

where : C_{rs} = Number of commuters between r and s

E_s = Employment in s (destination region)

L_r = Labourforce in r (origin region)

d_{rs} = distance between r and s

Living and working region can coincide (r=s).We have then 40x40=1600 equations. One may expect commuting flows that originate in a region to be roughly proportional to the size of its labour force. Therefore, the coefficient α is expected to be close to unity. The same reasoning holds for the coefficient β for local employment. The distance deterrence coefficient γ should be negative. We have chosen for a power function of distance decay. An often seen alternative is exponential distance decay, but the power function is recommended for longer distances (Fotheringham and O'Kelly, 1989).

An attraction constrained version

The sum of all incoming flows (including from the region itself) equals the total employment in the region. Since regional employment is determined elsewhere in model (section 4.1), its level is exogenous in the commuting block. Technically this leads to a restriction for each of the 40 regions. Equation (13) under these restrictions form an attraction constrained spatial interaction model. Equation (14) shows the attraction constrained version of model (13). It is easily verified that flows to region s sum up to regional employment in s.

$$C_{rs} = E_s \cdot \frac{L_r^{\alpha} \, . \, d_{rs}^{\gamma}}{\sum_r L_r^{\alpha} \, . \, d_{rs}^{\gamma}} \qquad (14)$$

Taking into account personal characteristics and a time trend

The commuting matrix has been estimated on total flows, without making a distinction in the personal characteristics or household composition of the commuters. The main reason is the lack of sufficient detailed estimation data (section 4.3.1), in particular when applying more personal characteristics at the same time. However we have seen in table 4.2 that these personal characteristics do matter, although the size of their impact varies. The strongest impact comes from education level, while there may be some interdependency with the others. We therefore choose to take only education level into account. We make the distance decay parameter γ dependent of the share of the higher educated in the regional labour force. We also add a time trend in the distance decay. We then get:

$$\gamma(t, c_r(t)) = \gamma_0 + g. \, t + \gamma_1 \, c_r(t) \qquad (15)$$

where: γ = distance decay parameter in region r

c_r = share higher educated in the labour force of region r

We found an average value of the distance decay parameter of around -3.5 in 1998. Estimation of equation (15) indicates a trend parameter of 0,026 and an impact of the education level (γ_1) on distance decay of 0,033.

Incorporating the impact of regional unemployment rates

W, we have brought in regional unemployment rates to serve as a tension variable to bring population and employment closer together. The idea is that people in regions with a relatively high unemployment rate will have to look further away to find a job compared to regions with a low unemployment rate. More in general, differences in regional unemployment rate will influence additional commuting flows.

In a dynamic specification of a linearised form of (15) we introduce the regional unemployment rate U_i in deviation of the national rate U_{NL} as an error correction mechanism. We then get :

$$\Delta \log(C_{rs}(t)) = (1-\beta)\cdot(\alpha \cdot \Delta \log L_r + \Delta\gamma \cdot \log(d_{rs}(t))) + \beta \cdot (U_r - U_{NL})(t-1) + \varepsilon \qquad (16)$$

Estimation of the dynamic system (16) for the 1992-2001 period suggests a strong impact of the unemployment rate. The parameter β is estimated to be around 0.8. At first glance this value seems to be rather large. However, although unemployment rates may vary over the country as a whole, the differences between nearby region are much smaller in general. The impact α of the labour force in the origin region is close to 1.0.

4.3.3 Regional unemployment

The regional unemployment level is by definition the difference between the regional labour force and the regional employment, corrected for net commuting. All variables are expressed in number of persons. Because we do not distinguish personal characteristics in employment we can not do it for unemployment as well.

$$U_r = L_r - E_r + \sum_s C_{s,r} - \sum_s C_{r,s} \qquad (17)$$

where: U_r = regional unemployment

The regional unemployment rate is defined as the unemployment level as a share of the labour force.

$$u_r = \frac{U_r}{L_r} \qquad (18)$$

4.4 Migration

The modelling framework for (domestic) migration is a generation - distribution approach. For each age group, we formulate a generation model for outgoing migration based on population characteristics. Unemployment rate, average education level and share of 1- person households in total households function as push factors.

In a production constrained spatial interaction model, region-specific pull factors determine the destination region of these migrants. Explanatory variables cover housing, labour market and study motives. Housing market related variables dominate short distance moves. Our estimation results show that labour market variables indeed play a significant role in describing long distance moves of young adults.

4.4.1 Some stylised facts on migration

Data on domestic and foreign migration are based on population statistics by municipalities as published by Netherlands Statistics. We take all residential moves, which cross municipality borders into account[6]. Data on these moves are recorded over a long period. People may have many motives to move. When they move to get closer to their new job, we speak of labour migration. Neighbourhood adjustment and housing quality are motives for residential migration. We assume that residential migrants do not change of job.

Compared to commuting domestic migration numbers between NUTS-3 regions have been rather stable during the last decades (Ekamper & van Wissen, 2000). Historically domestic migration was for a larger part labour induced. However, the impact of working seems to become smaller, as we have also seen from the increase in commuting. An attractive area of residence seems to become a more significant motive (housing migration). Nowadays in the Netherlands, about 10% of all moves is labour induced and residential migration constitutes about 25%, but the former share increases with migration distance (Ekamper and Van Wissen, 2000). Some young people have to move to get nearby their school or university, this is called study migration. Obviously, there can be many other personal reasons (like a marriage or the desire to live closer to ones relatives) for a migration. This is especially the case for the age group of above 65.

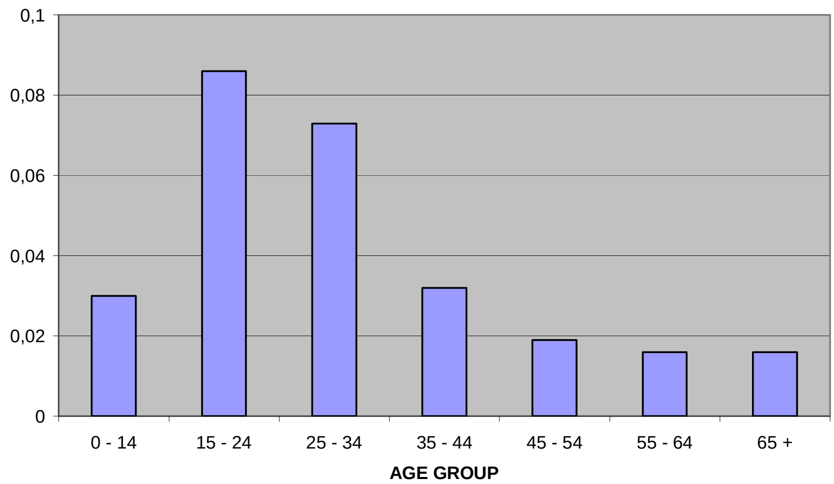
The life cycle and personal characteristics

The probability that someone decides to migrate can depend on many personal characteristics, of which age is likely to be the most important. Figure 4.4 presents migration propensity (number of migrations divided by population size) over seven age groups in the Netherlands in the year 2000. The life course approach (for example Plane and Rogerson, 1994) state that until

[6] Technically, this means that in our case, if both municipality of origin and destination lie in the same NUTS-3 region, migration can take place within the region.

the age of 18, most people live with their parents. In the years following, young persons move either for an education (study migration) or when they become employed after finishing their education (working migration) or just to get a place for themselves. When they get children they may look for a larger house and a quieter place to live. When they finally stop working, they face another natural moment to move (retirement migration).

Figure 4.4 Migration propensity in The Netherlands in the year 2000 by age group



A related observation is that the longer people stay at the same place the less they are inclined to move. This implies that the probability to migrate decreases with age. Empirically Eichperger and Gordijn also found that this probability in the Netherlands jumps up at the age of 18 and then decreases almost in a monotonic way. Unemployment, level of education and income are other personal characteristics that seem to influence migration. It is generally believed that the higher educated are less tied to a region and search jobs at the national rather than at the local level. Pecuniary migration costs are a smaller barrier to people with a higher income

Finally it is important to note that although we discuss personal characteristics here, migration is a household decision. A substantial share of migrations consists of actually tied moves. One-person households are generally the most mobile, with migration propensity decreasing with household size. Two earner households choose a destination such that both partners are at acceptable distance from their work. Obviously, a household that owns their dwelling is less likely to move than a household that rents its house. Van der Vlist et al. (2002) research the effect of other dwelling characteristics on household mobility. They also find that the impact of household characteristics on mobility differ in urbanised areas from their effects in rural areas. Every year a considerable number of people move into a new house. A lot of these moves are over a short distance only. For migration in the model, we only take into account the moves which cross a municipality border. We distinguish domestic migration and net foreign migration.

4.4.2 The generation - distribution approach

In a generation-distribution approach of migration, the number of people that decide to move (the generation) and the distribution of these migrants over regions of destination are treated separately. Generation is triggered by so called push factors, attraction to the destination region by so called pull factors.

The generation model

Equation (19) formulates a simple generation model based on regional age group and gender specific migration propensities. Let M_r be the total of out migration flows that originate in region r. The probability that someone of gender g and age group a decides to move out of region r is denoted $p_{r,g,a}$. POPr (g,a) refers to the specific regional population size

$$MO_r = \sum_{g,a} p_{r,g,a} \, . \, POP_r \, (g,a) \qquad (19)$$

where : MO_r = outgoing migration from region r

$p_{r,g,a}$ = probability to move

POP_r = population size in region r

The generation is triggered by a number of push factors. Beside a constant, these are: overall regional unemployment rate age in deviation of the national one, share of higher educated in the labour force and share of one person household in the total number of households. Equation (20) is estimated for all 7 age categories. Gender within each age category differs only in constant term.

$$p_{r,g,a} = \alpha_{g,a} + \beta_a \cdot (u_r - u_{NL}) + \gamma_a \cdot (c_r - c_{NL}) + \delta_a \cdot (H1_r - H1_{NL}) \qquad (20)$$

with : u_r = unemployment rate, region r

c_r = share of higher educated, region r

$H1_r$ = share of one person households in total number of households, region r

Table 4.3	The generation model for outgoing migration				
Age category	male	female	unemployment rate	% higher educated	% 1-p households
	(α_m)	(α_f)	(β)	(γ)	(δ)
0-14	.031	.030	.237	.045	-
15-24	.072	.100	.519	.016	.021
25-34	.076	.070	.173	.069	.110
35-44	.036	.028	.133	.048	.039
45-54	.020	.018	.070	.015	.031
55-64	.017	.015	.051	.010	.029
65+	.015	.016	.040	.007	.043

In table 4.3 we present the estimated results. The values of the constant term α show a decreasing influence of age. Unemployment rates have the largest influence on young people. Education level end share of 1 person households have the biggest impacts on the group between 25 and 34. The coefficients of the youngest age group resemble those of their parents in the 35-44 age group.

The distribution model

The outgoing migration flows are dispersed over all regions in the Netherlands, including the originating one[7]. Comparable to gravitation models in physics population size of the destination regions plays a (proportional) role in the attraction, as does the (squared) distance.

Migration will take place between all region pairs, although some flows may be rather small or even zero in certain years. Like in the commuting block we describe all the 40x40 flows. A migration distribution equation is estimated for each age group. From the generation model the size of the outflows is fixed for each region. Again we have 40 restrictions, this time on the production (of migrants) side. We then use a production constraint spatial interaction model. Comparable to (14) this can be formulated as :

$$M_{r,s} = MO_r \cdot \frac{POP_r^{\alpha} \, . W_{r,s}^{\beta} \, . \, d_{rs}^{\gamma}}{\sum_s POP_r^{\alpha} \cdot W_{r,s}^{\beta} \, . \, d_{rs}^{\gamma}} \qquad (21)$$

where: $M_{r,s}$ = domestic migration flow from region r to region s
POP_r = regional population in r
$W_{r,s}$ = attraction factor from region r to region s
$d_{,r,s}$ = distance between region r and region s

[7] see footnote 6 on page 19

Short distance and long distance pull factors

We make an explicit distinction between short distance and long distance motives. The idea is that up to a certain distance migration will be mainly driven by residential reasons. Above such distance it will be particularly labour market driven. After people have found a new job too far away to commute they will move to their new working region. We have set the separation distance on 75 km. This level is chosen somewhat arbitrary and particularly inspired by the fact, that above 75 km commuting numbers decrease drastically (section 4.3)

Fotheringham (1989) has put forward the concept of a hierarchical destination choice. People look first to a cluster of regions and decide then to settle in one of them. Population in the neighbouring regions is then a relevant factor. The chance to settle in a certain region becomes smaller if other densely populated regions are nearby as alternative destinations. We add a competition variable (COM) to the model.

Furthermore, we have found acceleration in the housing stock HS (exogenous in the model) and population density to be short distance pull factors. Faster growth attracts more immigrants. A low population density is a proxy for a more comfortable and maybe greener living area. Relevant long term pull factors are now job growth (E) and job opportunity (JO) in the destination region and its surroundings. The job opportunity is defined as the number of jobs per head of the labour force in a region as well as its surroundings weighted by distance.

By taking first differences of the logarithms we have rewritten equation (21) in growth rates. A spreader guarantees consistency.

$$\begin{aligned}\Delta \log M_{r,s} &= \alpha . \Delta \log POP_s + \beta . \Delta \log COM_s + \gamma . \Delta^2 \log HS_s^{ex} \\ &+ \delta . \Delta \log \left(\frac{POP}{km^2}\right)_s + \varepsilon . \Delta \log JO_s + \zeta . \Delta \log E_s\end{aligned} \tag{22}$$

$$COM_s = \sum_{r \neq s} \frac{POP_r}{d_{rs}^{\gamma}} \tag{23}$$

$$JO_s = \sum_r \frac{E_r}{d_{rs}^{\gamma}} \Big/ \sum_r \frac{L_r}{d_{rs}^{\gamma}} \tag{24}$$

where : COM_s = competing destination in destination region s
HS_s = housing stock acceleration in destination region s
JO_s = job competition growth in destination region s
E_s = employment growth in destination region s
` L_r = labour force in region r

Table 4.4 describes the estimation results for the seven age groups (Vermeulen 2003b). The coefficients for population are close to unity. The impact of lower population density becomes more popular for the older age groups. Job opportunity and job growth are important for the younger age groups

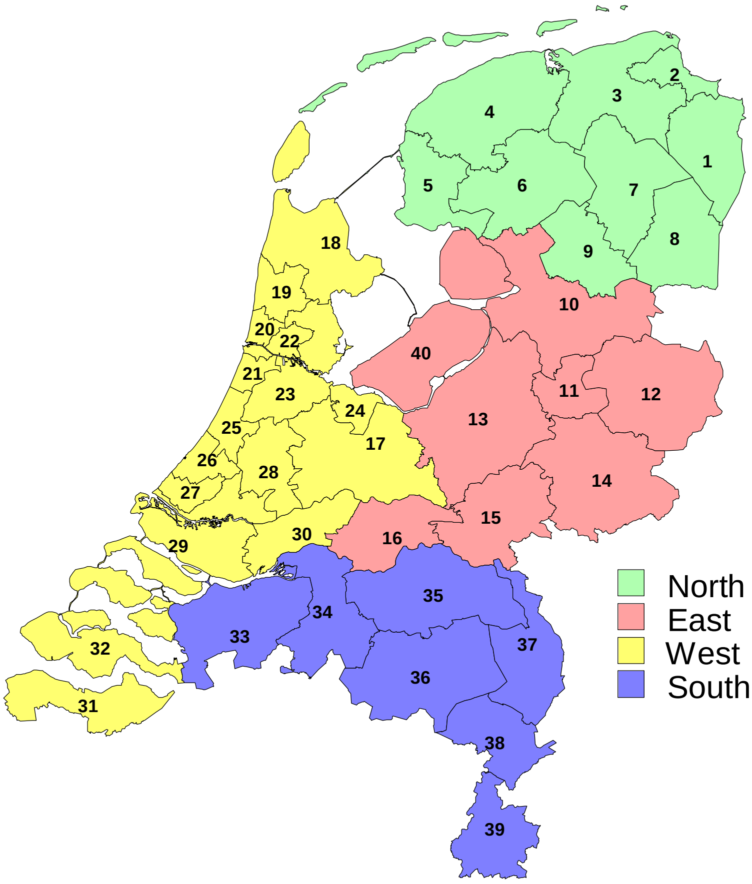
Table 4.4 Coefficients distribution model migration

Age category	population	competing population	acceleration housing stock	population density	job opportunity	job growth
	(α)	(β)	(γ)	(δ)	(ε)	(ζ)
0-14	.86	-.61	18.49	-.37	-	0.90
15-24	.93	-.83	17.30	-.09	1.20	3.67
25-34	.91	-.41	15.26	-.27	1.70	5.36
35-44	.86	-.50	13.39	-.37	0,73	3.33
45-54	.82	-.53	19.34	-.36	0.65	.72
55-64	.77	-.45	23.06	-.48	0.20	.63
65+	.90	-.08	3.97	-.68	0.10	-.14

4.4.3 Foreign migration

The size of the (net) foreign migration is exogenous in the model. Foreign migration is expected to play an important role in long term scenarios for the Netherlands. Their inflow is expected to contribute considerably to the otherwise possibly shrinking population. Migrants may come for all parts of the world. Their education level may differ. It is difficult to foresee on which region they will settle. Their size and breakdown in population categories and their distribution over the NUTS-3 regions is based on external information.

Appendix : Overview of the NUTS 3 regions in the Netherlands



NUTS-3 region numbers :

	North		East		West		South
1	East Groningen	10	North Overijssel	17	Utrecht	33	West North-Brabant
2	Delfzijl & surroundings	11	Southwest Overijssel	18	North of North Holland	34	Middle North-Brabant
3	Other Groningen	12	Twente	19	Alkmaar &surroundings	35	Northeast North-Brabant
4	North Friesland	13	Veluwe	20	IJmond	36	Southeast North-Brabant
5	Southwest Friesland	14	Achterhoek	21	Haarlem agglomeration	37	North Limburg
6	Southeast Friesland	15	Arnhem/Nijmegen	22	Zaanstreek	38	Middle Limburg
7	North Drenthe	16	Southwest Gelderland	23	Greater Amsterdam	39	South Limburg
8	Southeast Drenthe	40	Flevoland	24	Gooi & Vechtstreek		
9	Southwest Drenthe			25	Leiden & Bollenstreek		
				26	The Hague agglom.		
				27	Delft & Westland		
				28	East South-Holland		
				29	Greater Rijnmond		
				30	Southeast S.-Holland		
				31	Zeeuws Vlaanderen		
				32	Other Zeeland		

References

Broersma, L. and J. Van Dijk (2002), Regional labour market dynamics in the Netherlands, Papers in Regional Science, number 81, 343 - 364

Carlino, G.A. and Mills,E.S. The determinants of country growth, Journal of Regional Science (1987) Vol.27, No 1 p39-54

CBS(2001), Regionaal economische jaarcijfers 1998

Dijk, J. van, H. Folmer, H.W. Herzog Jr. and A.M. Schlottmann (1988), Efficiency of job matching mechanisms: a cross-sectional comparison, Papers of the Regional Science Association 64: 79 – 94

Dijk J. van, H. Folmer, H.W. Herzog Jr. and A.M. Schlottmann (1989), Labour market institutions and the efficiency of interregional migration: a cross-sectional comparison, in: Migrations and labour market adjustment, edited by Van Dijk, J., H. Folmer, H.W. Herzog Jr. and A.M. Schlottmann, Kluwer Academic Publishers

Eichperger L.H.& Gordijn H.(1994),A regional demographic model for the Netherlands,RPD

Ekamper P. en Van Wissen L.(2000),Regionale arbeidsmarkten, migratie en woon-werkverkeer, Nederlands Interdisciplinair Demografisch Instituut

Fotheringham A.S. (1993), Migration and spatial structure: the development of the competing destinations model, Migration models, macro and micro approaches, edited by J. Stillwell and P. Congdon, Bellhaven Press, 1991, 57 - 72

Fotheringham A.S. & O'Kelly M.E. (1989), Spatial Interaction Models: formulations and Applications, Kluwer academic publishers

Knaap, G.A. van der, F.G. van Oort and H. Scholten(1995), Een multiregional arbeidsaanbod voor Nederland (EGI onderzoekspublicaties)

Kwaak A.(1985) Regambev, Een model voor de provinciale ontwikkeling van arbeidsmarkt in Nederland, CPB Onderzoeksmemorandum no 8

Lowry, I.S. (1966), Migration and metropolitan growth: two analytical models, San Fransisco, Chandler

Ommeren, J.N. van (2000), Commuting and relocation of residences. A search perspective, Ashgate

Ommeren, J.N. van, W. Vermeulen and J. De Vries (2003), Commuting, spatial interaction models and panel data: how constant is the distance-decay parameter over time?, mimeo

Oosterhaven J., Knaap T. , Ruijgrok C.J., Tavasszy L.A. (2001), On the development of RAEM, the Dutch spatial equilibrium model and its first application, 41th European regional Science Association Conference, Zagreb.

Plane, D.A. & P.A. Rogerson (1994), The Geographical Analysis of Population: With Applications to Planning and Business, John Wiley & Sons, New York

Suyker, W.B.C. (1981), An analysis of interprovincial migration in The Netherlands, De Economist 129, Nnumber 3, 394 - 411

Verkade, E.M. (2002), A new regional labour market model, CPB Report 2002-4

Vermeulen,W. (2003a), A model for Dutch commuting, CPB Report 2003-1

Vermeulen, W. (2003b), International migration in the Netherlands: an aggregate analysis, Internal CPB paper

Vlist, A.J. van der (2001), Residential mobility and commuting, Tinbergen Institute

Vlist A.J. van der, C. Gorter, P. Nijkamp and P. Rietveld (2002), Residential mobility and local housing market differences, Environment and Planning A, number 34, 1147 - 1164