Papers in Evolutionary Economic Geography

05.02

Variety and regional economic growth in the Netherlands

Koen Frenken, Frank G. van Oort, Thijs Verburg & Ron A. Boschma

Utrecht University
Urban & Regional research centre Utrecht

Variety and regional economic growth in the Netherlands

Final report to the Ministry of Economic Affairs

Date: 24 December 2004

Koen Frenken[+,*]
Frank G. van Oort[*,o]
Thijs Verburg*
Ron A. Boschma*

[+] Corresponding author (k.frenken@geog.uu.nl)

* Section of Economic Geography, Urban and Regional research centre Utrecht (URU)
Faculty of Geosciences, Utrecht University, http://econ.geog.uu.nl

[o] The Netherlands Institute for Spatial Research (RPB), The Hague, http://www.rpb.nl

Acknowledgements

This research has been funded by the Ministry of Economic Affairs in The Netherlands as part of the research programme on 'Space and Productivity'. We thank Herm van der Beek, Henri de Groot and Jacco Hakfoort for helpful comments on earlier versions. We also thank Bart Los for the data on technological similarity between sectors. The usual caveats apply.

Faculteit Geowetenschappen

Universiteit Utrecht

Variety and regional economic growth in the Netherlands

Koen Frenken, Frank G. van Oort, Thijs Verburg and Ron A. Boschma

Summary

In economic theory, one can distinguish between variety as a source of regional knowledge spillovers, called Jacobs externalities, and variety as a portfolio protecting a region from external shocks. We argue that Jacobs externalities are best measured by related variety (within sectors), while the portfolio argument is better captured by unrelated variety (between sectors). We introduce a methodology based on entropy measures to compute related variety and unrelated variety. Using data at the COROP level for the period 1996-2002, we find that Jacobs externalities enhance employment growth, while unrelated variety dampens unemployment growth. Productivity growth, by contrast, can be explained by traditional determinants including investments and R&D expenditures. Implications for regional policy in The Netherlands follow.

Table of contents

1. Introduction

The relationship between variety[1] and economic development has been a neglected research area in economics. For long, economic theory has been focused on explaining economic growth by a combination of growth in inputs and efficiency improvements (Solow 1957). The underlying qualitative nature of economic development, for example, in terms of the variety of sectors or the variety of technologies, has been addressed only rarely.

One can distinguish between three types of relationships between variety and economic development. The first approach centres on variety, spillovers and growth, which has become a central theme in what is called new growth theory. It has been argued that, apart from spillovers occurring between firms within a sector, spillovers also occur between sectors. Following this argument, the present variety in an economy can be an additional *source* of economic growth (Jacobs 1969; Glaeser *et al.* 1992; Van Oort 2004). This means that not only the stock of inputs affects growth, but also the precise composition in a qualitative sense. Only some technologies and sectors are complementary in that their joint presence within an economy causes additional growth. And, since spillovers are geographically bounded, differences in regional growth should be related to qualitative differences in an economy's composition at the regional level. A region specialising in a particular composition of complementary sectors will experience higher growth rates than a region specialising in sectors that do not complement each other.

A second way to relate variety to regional economic development, and more specifically, to unemployment, is to view variety as a portfolio strategy to protect a region from external shocks in demand (Attaran 1985; Haug 2004). In this context, one also speaks of regional diversification analogous to corporate diversification as a risk spreading strategy. A high sector variety of a regional economy implies that a negative shock in demand for any of these sectors will have only mild negative effects on growth and employment. By contrast, a region specialising in one sector, or a group of sectors with correlated demand, runs to risk of serious slowdown in growth and high rates of unemployment as a result of a demand shock.

Finally, a third type of relationship between variety and economic development is more of an evolutionary nature (Pasinetti 1981; Saviotti 1996). An economy that does not *increase* the variety of sectors over time, will suffer from structural unemployment, and will ultimately stagnate. In this view, the development of new sectors in an economy is required to absorb labour that has become redundant in pre-existing sectors. This labour has become redundant due to a combination of productivity increases and demand saturation in pre-existing sectors, characterising the product lifecycle dynamics in each sector. These processes underlying long-term growth also have geographical implications, as new sectors typically emerge in urban areas while the older sectors are more dominant in rural areas. This means that labour becomes redundant primarily in rural areas, while new employment is primarily created in urban areas. This imbalance is counteracted by labour migration from rural to urban areas and by firm migration in the opposite direction.

The three approaches are not mutually exclusive, but can be considered to be complementary. The different views relate variety either to the economy's supply-side, or demand-side, or to both. The spillover approach argues that variety can be an additional *supply* source of economic growth, while the portfolio approach views variety as a buffer from external shocks in sectoral *demand*. The evolutionary approach combines supply-side and demand-side arguments arguing that the combination between productivity increases and demand saturation in pre-existing sectors necessitates the creation of new sectors leading to an ever-increasing variety in the economy. The approaches also differ in how they treat time. The first two approaches are static (or short-term) approaches in the sense that they assume variety in a region at time t to be related to regional

[1] We prefer to use the term variety rather than diversity, as diversity is closely linked to biology, while variety is the common term in economics. By variety we mean sector variety, unless stated otherwise.

growth or unemployment at time t (or $t+1$). The evolutionary approach takes a long-term perspective on regional development in arguing that the increase in variety is a necessary condition for long-term growth and full employment. The effects of variety may therefore become apparent only after considerable time lags.

Another issue, which is closely related but analytically distinct from the issue of variety and regional economic growth, is the relationship between variety and urbanisation. There is a wide agreement that variety is positively related to the degree of urbanisation, the reason being that a variety of products and sectors can only be sustained with sufficient local demand, both for intermediate inputs and final products. With urbanisation being positively related to variety, and variety being positively related to economic growth, urbanisation will generally have a positive impact on economic growth. However, it is important to distinguish, both theoretically and empirically, between urbanisation as a source of economic growth and variety *per se* as a source of economic growth (that is, when controlling for urbanisation).

Below, we present a survey of relevant literature on variety and regional economic development. We will discuss the main theories, measurement issues, and the outcomes of empirical research. Approaching the question of regional economic development from the concept of variety, we will not provide a comprehensive review of regional growth theory. Rather, we will zoom in on those theories that have something to say about the role of variety in economic growth. Following the three approaches distinguished above, we will discuss, respectively, theories of spillovers including the new growth theory and the economics of agglomeration (*section 2*), portfolio theory and regional diversification (*section 3*), and evolutionary growth theory and urban life cycle theory (*section 4*). We then discuss a number of recurrent measurement problems concerning variety (*section 5*), and relate those to the main outcomes of previous empirical research (*section* 6). We then turn to our empirical analysis using data on regional employment growth, productivity growth and unemployment growth for The Netherlands (*section 7*). Concluding remarks and policy reflections (*section 8*) finish up this report.

2. The economics of agglomeration

2.1 New growth theory

The important role of spillovers in explaining economic growth is central to a family of economic models headed under the label of new growth theory. The 'new' is to be understood with reference to the standard growth model developed by Solow (1957) and others. In this older view, the economy is represented by a production function transforming inputs into outputs. The growth of an economy can then be related to a combination of (i) the growth of inputs such as capital, labour and land which lead to a growth in output, and (ii) technological change that increases the efficiency with which inputs are transformed into outputs. Technological change is treated as being dependent solely on time. The rate or nature of technological change underlying the process of economic growth is not addressed in these models.

During the late eighties, it was acknowledged that the Solow-type of growth accounting lacks a theory of innovation (Romer 1986, 1990; Lucas 1988). The determinants of technological change became subject of further theorising, which led to the advent of new growth models. These models include education, research & development (R&D) or learning-by-doing as additional inputs affecting growth. Knowledge is regarded as an endogenous core input for economic growth.

Since, debates in economic growth theory have shifted from material to immaterial inputs, and, in particular, to the positive externalities arising from knowledge spillovers (Jaffe 1986; Griliches 1992). And, as spillovers imply the possibility of under-investment in knowledge (as firms recognise the danger of knowledge leaking towards competitors), government policy increasingly focused on stimulating on knowledge and innovation (*e.g.*, by providing subsidies for R&D or incentives to diffuse knowledge). Related to this, recent policy discourse centres on potential spillovers between public research institutions and the business sector, both in Europe and in the United States.

2.2 Agglomeration economies

A research area that is related to the new growth theory, though it is only partially overlapping, concerns the economics of agglomeration. In its crudest form, this field of research aims to explain the changing spatial distribution of economic activity over time. The core idea underlying the economics of agglomeration holds that clustering of economic activity occurs because firms experience some form of benefit from locating near one another. One of the reasons for firms to cluster is the existence of positive spillovers between firms that are located nearby. Evidence has been accumulating that suggests that knowledge spillovers are indeed geographically localised at a regional level (Jaffe 1989; Jaffe *et al*. 1993; Audretsch and Feldman 1996a). This means that firms profit most from spillovers if they are located near to other firms or public research institutions producing knowledge.

Before debating the possible causes of spillovers, it is useful to give a broader definition of agglomeration economies, and to distinguish between different types of agglomeration economies. A broad definition of *agglomeration economies* is that it concerns those external economies from which a firm can benefit by being located at the same place as one or more other firms. Many standard textbooks on economic geography and urban economics have incorporated the distinction between localisation and urbanisation economies. Localisation economies differ from urbanisation economies in that localisation economies are associated with benefits for firms that arise when locating near to other firms in the same industry, while urbanisation economies are associated with benefits for firms that arise when locating near to firms irrespective of their activity. Urbanisation economies are, therefore, also viewed as being a function of the population density in general, hence the term urbanisation. The distinction between localisation and urbanisation economies links to a threefold classification, which goes back to Hoover (1948) and Isard (1956), in which the sources of agglomeration advantages are grouped together as:

(1) *Internal increasing returns to scale*. These may occur to a single firm due to production cost efficiencies realised by serving large markets. There is nothing inherently spatial in this concept other than that the existence of a single large firm in space implies a large local concentration of factor employment;

(2) External economies available to all local firms within the same sector: *localisation economies*;

(3a) External economies available to all local firms irrespective of sector and arising from urban size and density: *urbanisation economies*;

(3b) External economies available to all local firms stemming from a variety of sectors: *Jacobs externalities.*

Localisation economies usually take the form of what are called Marshallian (technical) externalities whereby the productivity of labour in a given sector in a given city is assumed to increase with total employment in that sector. Marshallian externalities arise from three sources: labour market pooling, creation of specialised suppliers, and the emergence of technological knowledge spillovers (Henderson 2003; Feser 2002). The strength of local externalities is thus assumed to vary, so that these are stronger in some sectors and weaker in others. Localisation economies apply when the industry to which the firm belongs (rather than the firm itself) is large. Under further assumptions on crowding (congestion costs that increase with population triggers dispersion), perfect product and labour mobility within and between locations and the influence of large agents, the urban system is composed of (fully) specialised cities, provided that the initial number of cities is large enough (Henderson 1974; Richardson 1973).

Urbanisation economies reflect external economies passed to enterprises as a result of savings from the large-scale operation of the agglomeration or city as a whole and independent from industry structure (3a). Relatively more populous localities are also more likely to house universities, industry research laboratories, trade associations and other knowledge generating organisations. It is the dense presence of these organisations (not solely economic in character, but also social, political and cultural) that supports the production and absorption of know-how, stimulating innovative behaviour, and contributes to differential rates of interregional growth (Harrison *et al.* 1997).

The diverse industry mix in an urbanised locality also improves the opportunities to interact, copy, modify and recombine ideas, practices and technologies across industries (3b). Important innovations stem from the recombination of knowledge present in different industries. Geographical proximity between firms in different industries renders such recombination more likely to occur, in particular, if firms also operate under similar institutional conditions. The functional specialisation of firms in heterogeneous industries in close proximity of each other is supposed to generate spatial interdependencies and generates benefits (and costs such as congestion) for everyone in that specific location (Quigly 1998). Thus, variety in itself may be an extra source of knowledge spillovers and innovation. As this was first suggested by Jacobs (1969), this type of agglomeration economies is often referred to as Jacobs externalities.

2.3 Variety, economies of scale, and division-of-labour

To disentangle the different sources of agglomeration economies analytically, it is useful to go into the relationship between variety and economies of scale more deeply. Traditionally, variety in the economics literature has been referred to as the number of variants within a specific *product group*. The debate in this area addressed the relationship between product differentiation, market structure, economies of scale and consumer welfare within a neoclassical framework of complete information and equilibrium analysis.[2]

Two traditions of thought, going back to Chamberlin (1933) and to Hotelling (1929), have given rise to this literature. Chamberlin stressed differentiation and limited substitutability of products within an industry/product group, and the effects it would have on competition, by giving each firm a degree of monopoly (called monopolistic competition). Hotelling examined product differentiation from a strategic perspective, now better known as game theory, addressing the problem of where different sellers of a given commodity would locate in a one-dimensional space (*e.g.* a street or a product characteristic). In this case the growing dispersion of sellers would imply a greater product differentiation or variety. Interestingly, the two core models on product variety reach different conclusions. From a welfare perspective, Chamberlin's model of monopolistic competition would lead to an excessive variety, each firm monopolising a different niche, while according to Hotelling's model, too little differentiation would take place as each firm 'moves to the middle'. Though outcomes of equilibrium models on product variety may sometimes be sensible to the specific assumptions made, for example, regarding the market

[2] Some of this equilibrium-based economic literature recently introduces geographical space in its analytical framework (e.g., Murata 2003).

structure,[3] a number of conclusions seem quite robust. In particular, product variety in equilibrium will be greater (Saviotti 1996, pp. 100-101):

1. the smaller the economies of scale;
2. the lower the substitutability of goods within the group and with outside goods;
3. the width of the market (the degree of dispersion of preferences);
4. the depth of the market (density of consumer purchasing power for each variant).

The inverse relationship between variety and scale economies is also underlying the idea of agglomeration economies stemming from regional specialisation, though the relationship here is subtler. As internal economies at the firm level stem from specialisation and scale, external economies (*i.e.* externalities) stem from a region specialising in one sector thus allowing for a greater extent of division-of-labour *among firms*. The finer degree of specialisation among firms corresponds to the two classical Marshallian sources of agglomeration: specialised suppliers and specialised labour. In both cases, the benefits from clustering stem from a greater division-of-labour (among suppliers and among labour) such that inputs are more efficiently transformed into outputs. Note that in the hypothetical absence of transport costs, such benefits fade away as specialised suppliers and specialised labour would then be available for all regions.

2.4 New economic geography

When we introduce the distinction between urban and rural areas, one can understand that because urban regions harbour wider and deeper markets, as defined above, they also allow for a greater variety of goods being produced. As long as transport costs are important, one expects a greater variety of products to be produced in the vicinity of wider and deeper markets. Thus, one can expect a greater variety of goods to be produced in urban regions compared to rural areas.

Using a neo-Chamberlin set-up, product variety in cities is also underlying the 'new economic geography' models, pioneered by Krugman (1991) and elaborated by, among others, Brakman *et al.* (2001) and Fujita and Thisse (2002). In these models, consumers are willing to accept higher cost-of-living in cities in order to be close to a large variety of goods as the presence of variety increases the chances of preferences being met more accurately (called the *love-of-variety effect*). At the same time, it pays for firms to locate near consumers as to minimise transport costs to final markets. Automatically, this also implies that firms optimise internal economies of scale, as they produce on one location only. In this way, a typical equilibrium is characterised by complete clustering of all firms and consumers in one city (the core) with only sectors using immobile inputs, such as agriculture and tourism, being located outside the agglomeration (the periphery).

2.5 Spillovers

From the previous discussion, we understand that variety is typically greater in urban areas than rural areas. The presence of wider and deeper markets, as well as the opportunities for internal economies of scale, explains that firms and consumers prefer locating in one central city leaving the hinterland for agriculture and tourism. Following the reasoning of the new economic geography, variety is solely an *outcome* of the decisions of economic agents to cluster, and is not regarded as a source of economic growth. More specifically, the new economic geography models do not take into account geographically localised spillovers (or, non-pecuniary externalities in general), which would account as an additional, or even alternative explanation for urbanisation to occur. An important reason for firms to cluster is the existence of positive

[3] Several neo-Chamberlin (Dixit and Stiglitz 1977) and neo-Hotelling (Eaton and Lipsey 1975; Lancaster 1979) models arrive at different results using different assumptions.

spillovers between firms (and research institutions) that are located nearby. Evidence has been accumulating that suggests that knowledge spillovers are indeed geographically localised at a regional level (Jaffe 1989; Jaffe *et al*. 1993; Audretsch and Feldman 1996a). This means that firms profit most from spillovers if they are located near the actor from which the knowledge originates.

If one accepts, following the new growth theory, that spillovers are an important source of urban and regional economic growth, an important empirical question holds whether these spillovers occur primarily when a region is specialised in few sectors (localisation economies), or diversified into a large variety of sectors (Jacobs externalities), or whether it is primarily related to city size and density as such (urbanisation economies). In principle, all three types of agglomeration economies can occur as a result of spillovers, as a firm can learn from firms in the same industry (localisation economies), from firms in other industries (Jacobs externalities), or from a concentration of actors other than firms, including consumers, universities, and governments[4] (urbanisation economies). Focusing on the question whether regional growth benefits most from localisation economies or Jacobs externalities, the issue at hand is one of composition. As the amount of spillovers differs, both within each sector, and between each pair of sectors, the question is which precise composition of sectors in a regional economy creates most spillovers. Ideally, a regional economy is specialised sectors that are 'related' in that R&D investment in one sector spills over to other sectors.

The distinction between the different sources of spillovers bears important implications on theorising, because different types of spillovers are expected to lead to qualitatively different types of benefits. Localisation economies are expected to spur incremental innovation and process innovation, as the knowledge that spills over originates from similar firms producing similar products. The impact of localisation economies is thus expected to filter down primarily in productivity increases. By contrast, Jacobs externalities are expected to facilitate more radical innovation and product innovation as knowledge and technologies from different sectors are recombined leading to complete new products or technologies (compare Schumpeter's concept of 'Neue Kombinationen'). And, since radical innovations and product innovation lead to the creation of new markets and employment, rather than productivity increases, their impact may be very different from the incremental and process innovations caused by localisation economies. These qualitative differences in the types of innovation are also taken up by evolutionary growth theory (*section 5*).

Given that different types of spillover effects have potentially different effects on innovation and growth, one should be careful in selecting the output variables in an empirical research design. When analysing the impact of agglomeration economies on productivity growth, one can expect localisation economies to be important, while Jacobs externalities are expected to be important to explain differences in employment growth. Thus, both localisation economies and Jacobs externalities are all expected to contribute to regional economic development, but in different ways. This leads us to formulate the following hypotheses:

Hypothesis 1: Jacobs externalities are positively related to employment growth

Hypothesis 2: Localisation economies are positively related to productivity growth

[4] The importance for firms to interact with actors other than firms to innovate successfully underlies the concept of (regional) innovation systems (Freeman 1987; Braczyk *et al*. 1998; Boschma *et al.* 2002).

3. Portfolio theory

3.1 Corporate diversification

As outlined in the introductory chapter, a second theory relating variety to economic growth concerns portfolio theory. The concept of portfolio stems from business economics. It is usually applied to the valuation of a collection of assets, or to the impact of product diversification on corporate profitability and growth. Whatever the context of application, the concept of portfolio amounts to saying that variety reduces risk. Placing bets on more than one horse reduces the risk of high losses (although it also reduces the probability of high profits).

The extent to which a portfolio reduces risk is dependent upon the correlation between economic outcomes associated with each of the elements within a portfolio. For example, a firm that diversifies its sales into twenty different products with correlated demand (say, twenty different holiday destinations in Greece) will not substantially reduce the risk of going bankrupt, as a sudden fall in demand will hit all twenty products. By contrast, a firm that diversifies into only ten different products with uncorrelated demand will be more effective in reducing risk, as a fall in demand in one product is most likely to be compensated by a rise in demand for another product.

Though diversifying into products with uncorrelated demand is preferable as a risk reducing strategy, economies of scope[5] will generally be lower in a portfolio with uncorrelated demand compared to economies of scope in a portfolio with correlated demand. Thus, diversification into related products is often more rewarding for firms as a firm's core competencies can be better exploited. This hypothesis is in accordance to the resource-based and evolutionary theory of the firm that explain corporate growth as a process of diversification in related industries thus exploiting economies of scope (Montgomery 1994; Teece *et al.* 1997). Ideally, a firm diversifies into products that are related (to exploit economies of scope), while uncorrelated, or negatively correlated, in terms of demand.

3.2 Regional diversification

The sectoral composition of a regional economy can be approached in a way analogous to corporate diversification in product portfolios. Regional variety can be considered as a portfolio strategy to protect regional income from sudden sector-specific shocks in demand (also called asymmetric shocks that hit only one or few sectors, such as oil price shocks, a trade war, a radical innovation). This will especially protect labour markets, and thus prevent sticky unemployment to occur. Even if inter-regional labour mobility is high preventing unemployment to occur, asymmetric shocks reduce economic growth as agglomeration economies and the tax base deteriorate (Krugman 1993). Following this reasoning, industrial variety at the regional level would reduce regional unemployment and would promote regional economic growth, while specialisation would increase the risk of unemployment and a growth slowdown.

As for firms, a central question is whether related or unrelated diversification is most rewarding for stability and growth (Baldwin and Brown 2004). One can expect that related industries more often (though, again, not as a rule) have correlated demand shocks. Therefore, spreading risk over unrelated sectors is to be preferred from the viewpoint of a portfolio strategy. However, one should take into account the possible benefits from related diversification as well. Analogous to economies of scope at the firm level, one expects positive externalities within the region to occur primarily among related sectors, and only to a limited extent among unrelated sectors. In terms of

---

[5] Economies of scope arise when the joint production of multiple products by one firm is cheaper than the production of products by different firms, due to the reuse or better use of inputs (such as knowledge or machinery).

agglomeration theory, Jacobs externalities are expected to be higher in regions with a related variety of sectors compared to unrelated variety of sectors, because knowledge spills primarily between firms that use similar technologies and knowledge.

The effects of related and unrelated sector variety, therefore, are expected to differ. Unrelated variety protects a region best against external asymmetric shocks in demand. By contrast, related variety in a sector is expected to be beneficial for urbanisation economies and knowledge spillovers, thus enhancing growth and employment. Combining the portfolio theory with the economics of agglomeration discussed in the previous chapter, we can expect unrelated variety to protect a region from rising unemployment, while related variety, as an indicator of Jacobs externalities, enhances employment growth as already stated in hypothesis 2. This leads us to the following additional hypothesis:

Hypothesis 3: Unrelated variety is negatively related to regional unemployment growth

4. Evolutionary economics and product lifecycle theory

The first two approaches discussing the relation between variety and regional growth, are static approaches in the sense that they assume variety in a region at time *t* to be related to regional growth or (un)employment at time *t* (or *t+1*) As explained in the introductory chapter, evolutionary growth theory approaches the topic of variety from a long-term perspective in arguing that the increase in variety is a necessary condition for long-term growth and full employment. Put differently, where the approaches discussed above deal with effects of current variety on short-term growth, an evolutionary approach is concerned with long-term economic development, *i.e.* economic growth plus structural change.

4.1 Qualitative change

Evolutionary economics[6] can be understood as an economic theory that, first and foremost, is interested in out-of-equilibrium dynamics and structural change in the economy. We will not go into the history of evolutionary economics neither into its foundations. This has been done elsewhere (Nelson 1995a; Boschma and Lambooy 1999; Boschma *et al.* 2002). Here, we limit the discussion to evolutionary growth theory, which is a particular branch of evolutionary economics. The main characteristic of evolutionary growth theory, which also differentiates it from neoclassical or Keynesian growth theories, is the underlying theory of structural change in the sector composition of economies. Evolutionary growth theory is meso-founded, in particular, founded on product lifecycle theory. By contrast, neoclassical theory is micro-founded (firms and consumers), while the older Keynesian theories are macro-based. The latter approaches deal with growth without an explicit and comprehensive theory of structural change.

Another differentiating characteristic of evolutionary economics is its emphasis on economic growth as a process of qualitative change. Economic development, properly understood, refers to both the quantitative growth of an economy and its qualitative changes accompanying this process. In an abstract sense, qualitative change involves the introduction of entities previously not present in an economy, and the disappearance of entities previously present in an economy. Not withstanding the importance of institutions, which can either hamper or facilitate structural

[6] Not to be confused with evolutionary game theory.

change, the main source of qualitative change is undoubtedly innovation in processes, products and services, changing the variety of the composition of an economy (Saviotti 1996).

Approaching the question of economic development from the angle of structural change, uneven regional growth can then be understood as resulting from regional differences in sector composition, and the changes herein over time. Additionally, there will be regional-specific effects on growth (*i.e.* differences between regions when controlling for sector composition). In an evolutionary framework, such differences are often related to different institutional frameworks, which co-evolve with the structural composition of an economy (*e.g.*, university-industry relations, legal frameworks, the role of unions, etc.) (Nelson 1995b).

4.2 Pasinetti's growth theory

The work by Pasinetti (1981, 1993) has laid the foundation of an evolutionary theory of growth and development. Although the frame of reference is a closed economy of a country, its implications to regional economies are relatively straightforward. Pasinetti makes use of product lifecycle theory, which holds that industry evolution is characterised by product innovation in the first stage and process innovation in a second stage (Abernathy and Utterback 1978; Klepper 1997). Following this two-stage logic, Pasinetti explains growth as a combination of structural change caused by process innovation within existing sectors and product innovation leading to new sectors in the economy. The basic features of evolutionary growth theory are probably best summarised by two central arguments (Saviotti 1996):

1. The growth in variety is a necessary requirement for long-term economic development;
2. Variety growth leading to new sectors and productivity growth in pre-existing sectors, are complementary and not independent aspects of economic development.

Concerning the first argument, it can be noted that the growth of variety, as measured by the number of goods and services characterising the economy, is an empirical fact. In particular, during the last two centuries the number of goods and services has exploded with the number of product innovations greatly outweighing the number of products and services that have disappeared. However, evidence of growing variety does not imply in itself that growing variety is a necessary condition of economic growth. Concerning the second argument, we typically observe, at different moments in economic history, the emergence of complete new sectors alongside the 'maturation' of existing sectors characterised by productivity increase and demand saturation. Again, without further investigation, we cannot infer from these stylised facts that the two processes are complementary.

The theoretical support for the two arguments comes from Pasinetti's models (1981, 1993) of economic development and structural change.[7] Pasinetti's models contain a variable number of sectors, that is, they take into account the composition of the economic system and its relationship to economic development, a problem that is not dealt with by old nor new growth models.[8] The thesis developed in Pasinetti's models holds that an economy with a constant composition and constant variety, a constant productivity growth and a saturation of demand in particular goods/services, would not be stable. Such an economy would generate underutilisation of resources, including labour. It would become possible to produce the whole demanded output using smaller and smaller percentages of existing resources.

[7] See also, Saviotti (2001), Saviotti and Pyka (2004a,b), and Montobbio (2002).

[8] Notable exceptions are Romer's model (1990), which included a growth in the number of capital goods amongst the determinants of innovation, and Weitzman's model (1998), which included a production function for new knowledge generated by means of recombination of old knowledge.

While technological change, in the form of process innovation leading to productivity growth, would be partly responsible for this imbalance, it could also provide a form of compensation. Technological change in the form of product innovation creates new goods and services, and new sectors, which can re-employ the resources made redundant by the imbalance arising in the pre-existing sectors. Of course, no actual unemployment or re-employment needs to be involved if there is sufficient coordination between productivity growth in pre-existing sectors and the emergence of new sectors caused by innovations. Thus, on the one hand, growing variety by product innovation can help to overcome the bottlenecks created by the imbalance between productivity growth and demand saturation in pre-existing sectors. On the other hand, new goods and new services can only be created as the resources become available from productivity growth in pre-existing sectors. These resources are required to invest in R&D and training, which creates new goods and services.[9] This inter-sectoral re-allocation of resources, however, is not automatic, and is expected to occur in different ways and with different rates of success in different regions (and is very much dependent on the institutional framework).

The theory of Pasinetti distinguished between two types of technological change, process innovation leading to productivity increases in existing sectors and product innovation in new sectors, which can be associated with quantitative and qualitative growth, respectively. This distinction is in line with the more widely known product lifecycle theory that predicts product innovation to characterise the first stage of a sector's evolution and process innovation a sector's second stage of evolution (Vernon 1966; Abernathy and Utterback 1978). Here, we mean by sector an individual industry such as the automobile industry. This distinction should not be taken to mean that product innovation occurs exclusively at the time of birth of a new industry with process innovation taking over hereafter. Rather, product lifecycle theory assumes product innovation to peak before process innovation peaks. Evidently, Pasinetti's theory marks an important step forward in the integration of meso theory, traditionally allocated to the field of industrial organisation, and macro theory, traditionally dominated by the growth literature.

It is important to recognise that Pasinetti's theory, in first instance, is assumed to hold for the highest level of aggregations. Spatially, the theory should hold at least at the level of a set of integrated countries, or a single country that is not integrated with other countries. On the regional level, as defined as sub-national spatial units of analysis, the two hypotheses do not hold necessarily. However, the theory bears important implications for the role of variety at urban and regional levels, as discussed in the next section.

4.3 Urban and regional product lifecycle theory

To understand the implications of evolutionary growth theory and structural change for regional growth, the central question becomes how the product lifecycle takes place in geographical space. Related to this question, is the analysis of technology diffusion through space. Until the 1970s the economic literature paid most attention to diffusion of knowledge mostly in time, while few studies addressed the topic of diffusion in space. The little attention to the *spatial* diffusion of knowledge was mainly advocated by economic geographers and not by

---

[9] Note that the complementarity between variety growth and productivity growth in pre-existing sectors bears a considerable similarity to that between productivity growth in agriculture and investment in the new industries during the process of industrialisation (Kuznets, 1965). However, contrary to traditional views on economic development, evolutionary economics does not assume all countries to follow a deterministic order of sectors. By contrast, less developed countries can leapfrog developed countries by immediately specialising in sectors that have been created most recently in the world (for example, Germany leapfrogging the UK at the end of the 19th century by being successful in the then emerging chemical and machine industries). On the regional level it is even harder to maintain focussing on predetermined growth-stages over time and sectors, as recently argued by Parr (2001) who applied Rostow's well-known theory of growth-stages to regions.

(regional) economists (Van Oort 2004). Diffusion of innovation over space has long been supposed to occur according to predictable, well-known proximity patterns. This conceptualisation is worked out with several concepts of agglomeration theory, the best known being growth pole theory (Perroux 1950; Boudeville 1966). Its main assumption is that economic growth (in the form of innovations) is spread throughout a growth centre's hinterland to smaller nearby cities and localities. Innovations and knowledge are (once generated in a certain central location) expected to spread among regions from one locality to its neighbours.

An evolutionary perspective predicts that new growth poles emerge as a result of the emergence of new technologies. The locations of new growth poles are fundamentally uncertain, because the development of new sectors is not triggered by specific location factors. Rather, the development of new technologies requires generic inputs like general knowledge, infrastructure and demand; the specific inputs of new technologies are not present in any region, as they need to be developed more or less from scratch. In this context, one speaks of open *Windows of Locational Opportunity* in the early stage of a new sector (Storper and Walker 1989; Boschma and Lambooy 1999; Boschma and Frenken 2003). The generic inputs of knowledge being important in the early stage of a lifecycle, cities are expected to be the source of most product-lifecycles *casu quo* growth poles.

Yet, as a key sector driving a growth pole develops over time, its growth is likely to slowdown, as production tends to shift from high-wage locations to lower wage locations over the product lifecycle. The shift from production from high-wage to low-wage locations can be understood as a consequence of the shift from product to process innovation, which implies a shift from innovation and production by high-skilled labour to standardised mass-production by low-skilled labour (Markusen 1985). In the original formulation of the product lifecycle theory by Vernon (1966), these shifts were assumed to occur from high-wage countries, to medium-wage countries, and later to low-wage countries. This process involves both firms setting set up plants in foreign countries, and new firms set up in locations with lower wages outcompeting established firms in locations with higher wages.

This reasoning can be translated into a geographical framework by assuming that new lifecycles start in urban environments and move to more rural environments over time. The knowledge of the urban labour force, capital services and product markets foster the function of nursery schools for starting firms (the incubation function of cities or agglomerations, Davelaar 1989). In accordance to the economics of agglomeration, evolutionary economists also stress the important role in variety to create new variety. In other words, Jacobs externalities are assumed to play an important role in urban areas to create new varieties and new sectors. In this context, the evolutionary notion of recombination, analogous to gene recombination in biology, of technologically related pieces of knowledge and technology is relevant (compare Schumpeter's concept of 'Neue Kombinationen').

When firms survive and become mature, they tend to standardise production and become more capital-intensive. The initial advantages of the urban agglomeration core now can become disadvantages: growth is difficult to be realised *in situ* and physical movement becomes apparent when limited accessibility and high wages become disadvantageous. Growing firms are expected to 'filter down' towards more peripheral locations and regions, where land, labour and transport costs are more economically justified[10]. This dynamic reasoning lies behind the notion of an 'urban product lifecycle' that new products are developed in large diverse metro areas with diversified skill base, but that matured firms eventually decentralise. Note that during migration processes of firms to more rural areas, innovation can still occur, a process that Davelaar (1989) called creative diffusion. Also note that migration does not necessarily occur solely from cities to rural areas, but also occurs hierarchically between cities (from larger to smaller cities).

[10] The term 'filtering down' in relation to spatial diffusion and occupational specialisation was first introduced by Thompson (1968).

Whether a city can sustain high growth in the long run depends on whether it can create new product lifecycles through product innovation. To some extent, this propensity to create new sectors within a city is endogenous, as current variety is expected to facilitate the creation of new varieties through Jacobs externalities (cf. Weitzman 1998). Note in this context that, in principle, new product life cycles can also be initiated in the pre-existing sectors opening new niches (say, moving from textile into fashion), or in complete new sectors. The first development contributes to related diversification, while the second process contributes to unrelated diversification. In either case, what is needed for sustained economic growth is a sufficient rate of product innovation creating employment, which can take place both in existing industries or in new emerging industries.

In evolutionary economics, the creation of new varieties leading to new sectors is thus considered to be a necessary condition for long-term economic growth of a country. Spatially, this 'function' is allocated to urban regions, while rural areas contribute primarily in terms of productivity growth in maturing sectors. The evolutionary model has important implications as a theory of regional convergence or divergence. The model predicts regional unemployment levels to diverge as urban areas are nodes of employment creation through product innovation, while rural areas are the source of unemployment through process innovation (except in the case that growth in export markets compensates, which in the longer run will not be the case as demand saturates). The regional divergence in unemployment can only be counteracted if net migration of workers takes place from rural to urban areas and/or net migration of firms takes place from urban to rural areas.

From our previous discussion, we can derive three hypotheses:

Hypothesis 4: Over the product lifecycle, ideal-typically, economic activity moves from urban regions to rural regions

Hypothesis 5: As employment growth is highest in the early stage of a lifecycle and productivity growth in the later stage, urban regions create most employment while rural areas create most productivity growth

Hypothesis 6: Regional divergence in unemployment can be counteracted by a net migration of (young) workers from rural to urban regions being trained sufficiently to contribute to product innovation and structural change

5. Measurement issues

Measuring diversification over sectors in regional economies is sensitive to the indicator applied. In our empirical analysis we apply an entropy measure. This measure provides one with a straightforward indicator of variety. As set out earlier by Theil (1972), Jacquemin and Berry (1979), Attaran (1985) and Frenken (2005), the main reason for using entropy is its decomposable nature. In the context of measuring regional variety to analyse the effects on growth, decomposition is informative as one expects entropy/variety at a high level of sector aggregation to have a portfolio effect on the regional economy protecting it from unemployment, while one expects entropy/variety at a low level of sector aggregation to generate knowledge spillovers and employment growth. Put differently, entropy at a high level of sector aggregation measures unrelated variety, while entropy at low level of sector aggregation measures related variety.

5.1 The entropy measure

The origin of the entropy concept goes back to Ludwig Boltzmann in 1877 and has been given a probabilistic interpretation in information theory by Claude Shannon (1948). In the 1960s, Henri Theil developed several applications of information theory in economics collected in *Economics and Information Theory* (1967) and *Statistical Decomposition Analysis* (1972).

The entropy formula expresses the expected information content or uncertainty of a probability distribution. Let E_i stand for an event and p_i for the probability of event E_i to occur. Let there be n events E_1, …, E_n with probabilities p_1,…, p_n adding up to 1. Since the occurrence of events with smaller probability yields more information (since these are least expected), a measure of information h should be a decreasing function of p_i. Shannon (1948) proposed a logarithmic function to express information $h(p_i)$:

$$h(p_i) = \log_2\left(\frac{1}{p_i}\right) \qquad (1)$$

which decreases from infinity to 0 for p_i ranging from 0 to 1. The function reflects the idea that the lower the probability of an event to occur, the higher the amount of information of a message stating that the event occurred. Information is here expressed in bits using 2 as a base of the logarithm, while others express information in 'nits' using the natural logarithm.

From the n number of information values $h(p_i)$, the expected information content of a probability distribution, called entropy, is derived by weighing the information values $h(p_i)$ by their respective probabilities:

$$H = \sum_{i=1}^{n} p_i \log_2\left(\frac{1}{p_i}\right) \qquad (2)$$

where H stands for entropy in bits.

It is customary to define (Theil 1972: 5):

$$p_i \log_2\left(\frac{1}{p_i}\right) = 0 \quad \text{if} \quad p_i = 0 \qquad (3)$$

which is in accordance to the limit value of the left-hand term for p_i approaching zero (Theil 1972: 5).

The entropy value H is non-negative. The minimum possible entropy value is zero corresponding to the case in which one event has unit probability (absence of any variety):

$$H_{\min} = 1 \cdot \log_2\left(\frac{1}{1}\right) = 0 \qquad (4)$$

When all states are equally probable $p_i = (1/n)$ the entropy value is maximum (all varieties are present with equal shares):

$$H_{\max} = \sum_{i=1}^{n} \frac{1}{n} \log_2(n) = n\frac{1}{n}\log_2(n) = \log_2(n) \qquad (5)$$

Proof of equation (5) is given by Theil (1972: 8-10). Maximum entropy thus increases with *n*, but decreasingly so. This property reflects the idea that each new variety contributes positively to the total variety in a system, yet this contribution is marginally decreasing with each additional variety added.

5.2 Indicators of related and unrelated diversification

An important measurement issue is how to distinguish between related and unrelated diversification. As explained above, the concept of related variety holds that some sectors are more related than other, and will generate relatively more inter-industry knowledge spillovers. To examine empirically the effect of related or unrelated diversification is not a trivial matter and sophisticated methodologies to measure inter-sectoral spillovers are scarce.

One methodology applied in the context of related and unrelated diversification, both at the firm level (Jacquemin and Berry 1979) and the regional level (Attaran 1985), concerns the entropy measure. The main advantage of the entropy measure, and the reason for its use in the context of diversification, is that entropy can be decomposed at each sectoral digit level (see the decomposition theorem in the appendix). This allows one to measure the effect of variety at different levels of aggregation. The decomposable nature of entropy implies that variety at several digit levels can enter a regression analysis without necessarily causing collinearity (Jacquemin and Berry 1979; Attaran 1985).

One expects, for theoretical reasons, that variety in a region at different digit levels of sectoral aggregation has different effects on economic performance variables of a region. For example, one expects variety at the highest level of aggregation (1-digit level), as an indicator of unrelated variety, to be negatively associated with unemployment as it dampens the effects of external demand shocks. And, one expects variety at the lowest possible level of aggregation (five-digit level), as an indicator of (strongly) related variety, to correlate positively with economic growth and employment growth.

A second methodology, applied to patent data rather than sectoral distributions, is Jaffe's (1986) methodology to use patent data to compare the patents of firms. His methodology assumes that the more similar the patents are of two firms, the more they mutually benefit from their R&D investments (possibly deflated by a distance-decay function to capture the beneficial effect of geographical proximity). Importantly, Jaffe's methodology applies to the firm level, and, consequently, requires high-quality data of firms and their patents. Also note that patents do not play an important role in all industries as many industries exist where most innovations are not patented (*e.g.*, services).

A more recent methodology, which we will use in our empirical study discussed below, is developed by Los (2000). His methodology aims to capture inter-industry spillovers (rather than inter-firm spillovers), by measuring the degree of technological similarity between two sectors. Similarity is measures by comparing the input mix of two sectors from input output tables. As input mixes reflect production technologies, a high similarity in input mixes of two sectors implies a small 'technological distance' between two sectors, and a high amount of spillovers. Conversely, two industries with very different input mixes are technologically distant, and, consequently, will hardly mutually benefit from spillovers. It must be noted, however, that this measure only deals with technological relatedness, as a proxy for inter-industry spillovers, and does not necessarily reflect how much sectors are related in terms of demand shocks (underlying portfolio theory). Thus, an additional indicator may be preferred to distinguish between technological relatedness, and, what one may call, demand relatedness. For example, one can use the correlation matrix of sector demand indicating to what extent the demand for products from two sectors correlate.

5.3 The example of regional variety in The Netherlands

To distinguish between the effects of unrelated and related regional diversification on regional growth, or, better stated, between degrees of relatedness in diversification, we concluded that one should use information on all levels of sector aggregation. In this context, the use of entropy is advocated because it can be decomposed in marginal entropy values at all levels of sectoral aggregation. An additional reason to rely on sector data at all digit levels, rather than solely on one level (as has been done in most empirical studies so far), holds that the levels of variety turn out to be very different at different levels of sector aggregation. Thus, outcomes of empirical tests may well be sensitive to the (otherwise arbitrarily) chosen level of sector aggregation. As Baldwin and Brown (2004, p. 538) suggest, the geographic unit of analysis when testing for variety and economic volatility relations should not oversize labour market regions, because it is on the level of such regions that product diversity matters and regional policy can be of influence. In our study we concentrate on the relation of variety with growth and unemployment on the level of labour market areas in the Netherlands, the so-called COROP-regions (n=40, see for a test on coherence of these regions Bongaerts *et al.* 2004). The maps of related and unrelated variety provided in figure 2 present two very different regional patterns for related variety (two-digit variety) and unrelated variety (marginal increase in entropy when moving from two- to five-digit). As it is clear from the maps, variety at high levels of aggregation shows little resemblance with variety at low levels, which strongly suggests that the choice of sector aggregation is not trivial. The absence of correlation between related and unrelated variety further supports this (see table 2).

6. A review of empirical research

6.1 Empirical research on agglomeration economies and economic growth

Research on agglomeration issues in relation to variety has grown rapidly in recent years (Glaeser *et al*. 1992; Feldman and Audretsch 1999; Duranto and Puga 2001; Acs 2002; Van Oort 2004; Van der Panne 2004). Agglomeration economies are at the explanatory core of uneven spatial development, in cities and regions. Research aims to explain differences in productivity, economic growth and innovation intensity in urban areas and metropolitan regions. As said, renewed attention is paid to economic externalities, in a pecuniary sense (as in the new economic geography) and in a non-pecuniary sense (knowledge spillovers). The paradox in urban economics in recent years, however, is that agglomeration economies (and diseconomies) are seen as the driving force behind explanations of geographical concentration of economic activity within cities, while they remain something of a black box.

An important part of the argument in Van Oort (2004) is indeed that little is understood on causal relations in economic agglomeration forces. The lack of consistent economic frameworks still causes considerable degrees of freedom in conceptualising within sub-disciplines like evolutionary, trade and growth economics, and prohibits an overall clear impact on urban economics. But the relevance and potential impact of the hypotheses stemming from these disciplines are becoming more and more accepted and further explored. In recent contributions (*e.g.*, Henderson 2003), innovation based spatial diffusion modeling and endogenous growth theory both emphasize the spatial proximity role of knowledge possessed by economic agents. These contributions aim at identifying knowledge spillovers between agents as crucial factors leading to external economies of scale in production. Empirical tests of this theory often have looked at cities, systems of cities and agglomerations of economic activities to identify settings in which agglomeration economies foster growth.

One of the most promising contributions in understanding the nature and content of agglomeration economies is provided in the theoretical and empirical framework developed by Glaeser *et al*. (1992). In their seminal contribution to the research field entitled 'growth in cities', an endogenous growth framework based on employment growth patterns in mixtures of industries in US cities is developed. Glaeser *et al*. (1992) interpreted Romer (1986) as predicting that knowledge spillovers will be most significant among firms in the same industry. The implied corollary is that industries that represent a high level of employment in a given city relative to that industry nationally will grow faster (Feser 2002). Missing in the contribution by Glaeser *et al.* (1992) is a proper treatment of space, distance and spatial dependence as surfaced from the regional science and geographical literature.

Since, empirical research on agglomeration economies has been rapidly expanding. Extensive reviews are available in Feser (2002), Parr (2002), Dissart (2003), Fingleton (2003) and Van Oort (2004). Here, we list the central conclusions:

- The role of sectoral diversity in relation to economic performance on the urban and regional level is not unambiguously clear. While most studies find some positive effects of variety on growth and innovativeness, other studies have found no such evidence. The ambiguity in results is probably due, at least in part, to problems of definition. Different definitions of variety, economic performance, spatial scale and spatial and sectoral linkages appear to be highly influential.

- The choice of the spatial unit of analysis is often not motivated theoretically (Feldman and Audretsch 1999; Acs 2002). In this context, Fingleton (2003) pleads to use the functional region as the spatial unit of analysis, because of the natural control for labour market induced heterogeneity. Moreover, knowledge spillovers are geographically bounded, primarily, in geographical labour market areas.

- The spatial conceptualisation of agglomeration externalities has become an important factor in regional growth theory. Proximity as well as regime-like spatial heterogeneous conceptualisations appears to be important (Van Oort 2004). Empirical analyses therefore have to incorporate the influence of proximate regions on 'own' growth more explicitly in order to capture spatial spillover effects. This aspect especially concerns the urban level, but the regional level as well (Fingleton 2003).

- Analysing spillover-effects within and between sectors gains from modelling economic transfer channels of technological change rather than solely depending on implicit interpretations of the importance of specialisation- and diversity patterns (Breschi and Lissoni 2001; Helsey and Strange 2002). To this end, one can make use of input-output data (*e.g.*, Los 2000) or patent citations (*e.g.*, Jaffe et al. 1993), or collaboration networks (*e.g.*, Breschi and Lissoni 2003).

- Other common problems concern the extremely broad conceptualizations of diversity-based urbanisation economies in terms of just population density and the belief that diversity is just the reciprocal of specialisation using one indicator for both aspects (Moomaw 1998; Rigby and Essletzbichler 2002; Henderson 2003).

- There is a lack of a common use of industry-specific scale (see for instance Fingleton 2003). A straightforward solution is to include multiple digit levels in regression models using the decomposable entropy measure.

6.2 Empirical research on portfolio theory

Empirical evidence on the role of diversity in portfolio theory as cushioning the adverse effects of economic cycles is as divided as the role of diversity in regional economic performance issues (Dissart 2003). The lack of consistency in research outcomes may well reflect the lack of standardised research methodologies rather than evidence that portfolio theory is falsified. Let us summarise the main results:

- Dissart (2003, p. 434), after reviewing studies on regional diversity, concludes that in general more diversity leads to more stability (also depending on how diversity and stability are conceptualised). But this meta-result is guised by the fact that most research uses aggregated sectoral employment data (Kort 1981).

- Dissart (2003) also concludes that most studies conclude that diversity is associated with less growth of unemployment.

- Results on income levels are mixed. There appears to be a need for testing the relation between economic diversity and income distribution on regional levels in order to better understand welfare effects of a diversified economy (who are socially winners and losers of diversification). Two different hypotheses on the diversity-income relation circulate (Izraeli and Murphy 2003; Attaran 1985). One hypothesis states that industrial diversity should trade off employment security for lower overall income. Another hypothesis states that diversity and income are positively correlated.

- Most portfolio measures lack a sound conceptual basis. In many cases, portfolios are compared to a national sectoral distribution. Why this national distribution is to be the benchmark is not always clear (Wasylenko and Erickson 1978). A diversified national economic growth policy may be of little use for regional economies (Bhattacharya 2003).

- Many diversity-measures are not comparable over different scales of industry-measurement. Entropy- and portfolio-variance analyses are. As explained, the entropy measure has the advantage of being decomposable in a straightforward way at any digit level. The portfolio variance measure has the advantage that besides the level of diversity, the level of economic activity in general (in a region) is integrated as a weighing mechanism (Brewer 1985).

- Kim (1987), and Baldwin and Brown (2004) argue that the spatial level of testing variety-volatility (unemployment) relations should be that of labour market regions, while Haug (2004) shows that the municipal level (in Germany) captures important sources of variation in this relationship. The spatial level of analysis is thus not unambiguously clear - municipal analysis seems to ignore spatial dependence with surrounding suburbs and regions - while analyses of labour market regions may ignore intraregional variation.

6.3 Empirical research on product lifecycles

Evolutionary growth theory is based on the assumption that sectors go through sector-specific lifecycles. Regional growth, then, is dependent on the precise industry-mix in a region. Regions where fast-growing sectors are dominant will grow faster than regions in which saturated sectors are dominant. Empirically, this industry-mix effect is typically accounted for using a shift-and-share analysis. Regional growth in the longer run, however, is not only dependent on the current industry mix, but also, and crucially, on the capacity of a region to creative new industries, or to find new niche markets in saturated industries. This thesis, in line with Pasinetti's (1981, 1993) thesis of increasing variety as a necessary condition for long-term growth, has not been put to the test rigorously hitherto. Although popular thinking would attribute more variety to modern economies compared to less developed economies, at least at high levels of spatial aggregation, systematic evidence is lacking. There is some evidence, as discussed by Laursen (2000), showing that export variety has grown over the past decades in OECD countries. Furthermore, Saviotti (1996) and Frenken *et al.* (1999) find evidence for increasing variety at the level of macro-technologies.

By contrast, research on urban product lifecycles is more developed, and actual as ever, according to recent elaborations on the theme. The spatial scale is often not so large as originally stated by Vernon (1966) or Hymer (1976), who argued that the filtering-down processes work at national or international scales. Recent research instead reveals that during the product lifecycle of a sector, firms moving out of urban areas often locate near the core location they leave, still in the influence sphere of the larger conurbation of a city. This is empirically confirmed by Carlton (1982), Chapman and Walker (1991) and Phelps *et al*. (2001). The latter authors named this the concept of 'borrowed size': while still gaining form agglomeration advantages, disadvantages like congestion and high land prices are avoided. More recently, this theoretical framework has been applied in agglomeration studies on innovation intensity (Henderson 1997; Brouwer *et al.* 1999), on employment growth (Henderson *et al.* 1995; Dumais *et al*. 2002), and on new firm formation (Van Oort and Atzema 2004).

Most of these studies use aggregate development patterns to show that the filtering-down processes of technology-driven diffusion stay close to home. The study by Bleichrodt *et al*. (1992) relatively accurately focuses on spatial detailed and longitudinal testing of filtering-down hypotheses in technology-driven diffusion in the Netherlands. From this study, the following spatial development trajectories can simultaneously be distinguished as important sources of spatial heterogeneity concerning economic growth: (1) national zoning; (2) urban-hierarchical effect (large cities are higher in the diffusion-rank than medium-sized and small cities); and (3) proximity-based neighbouring effects.

In combination with the introduction of more homogeneous sectors or industries, both Peneder (2001) and Knaap (2004) argue (both in vain of the new economic geography that age-determined lifecycle aspects of firms are important when interpreting filtering-down theories of regional development. Knaap (2004, chapter 3.3) introduces a core-periphery model on economic growth, in which the evolution or filtering-down of technological-based development is based on age-determined cohorts of firms. In line with Vernon's (1966) theory of regional specialisation, it is concluded that new firms usually come up in urban centres (cores) and that over time incumbent, mature technological firms move to peripheral regions.

Important aspects that came to the fore when summarising the literature on spatial lifecycles in relations to variety are:

- The aspect of variety in urban and regional lifecycle research is often only implicitly attached to the urban core region as a source of new varieties (Audretsch and Feldman 1996b).

- Explicit distinctions have to be made concerning sectors that react differently on spatial lifecycle hypotheses. Especially the distinction between services, manufacturing and distribution appears to be important. The aspect of within-sector or related sector growth potentials plays a role in this as well.

- Ideally, analyses incorporating time-lag estimations should be applied in empirical analysis. This in order to capture the age and cohort-effects attached to urban and regional lifecycle hypotheses. Our data in our forthcoming study, however, are not longitudinal enough to properly address this aspect.

7. Hypothesis testing for regional growth in The Netherlands

7.1 Hypotheses

We distinguished between three types of relationships between variety and economic development. The first approach centres on variety, spillovers and growth, and sees variety as a source of agglomeration economies stimulating growth. A second way to relate variety to regional economic development is to view variety as a portfolio strategy to protect a region from external shocks in demand (Attaran 1985). In this context, one can distinguish between related and unrelated diversification, where the former supports spillovers while the latter is the best buffer against external demand shocks. Finally, a third type of relationship between variety and economic development is based on evolutionary growth theory, and predicts that only economies with a sufficient increase in variety will have sustainable growth and full employment. In this view, the creation of new sectors is thus a *sine qua non* for economic growth. Related to this view is the product lifecycle theory of sectors, which predicts new sectors to emerge in cities and to migrate to rural areas only later.

From these theories, six hypotheses are derived:

H1: Jacobs externalities (related variety) are positively related to employment growth

H2: Localisation economies (Los-index or specialisations) are positively related to productivity growth

H3: Unrelated variety is negatively related to regional unemployment growth

H4: Over the product lifecycle, ideal-typically, economic activity moves from urban regions to rural regions

H5: As employment growth is highest in the early stage of a lifecycle and productivity growth in the later stage, urban regions create most employment while rural areas create most productivity growth

H6: Regional divergence in unemployment can be counteracted by a net migration of (young) workers from rural to urban regions being trained sufficiently to contribute to product innovation and structural change

Using the data available covering employment growth, productivity growth and unemployment growth in the period 1996-2002, we focus on hypothesis 1, 2 and 3 only. Hypothesis 4, 5 and 6 deal with long-run development and are not tested, because it would require a much longer time-series. However, we make use of evolutionary growth theory and urban lifecycle theory in interpreting some of the results that follow.

7.2 Data

In this section we give an overview of the indicators and data used in our empirical analyses. Maps at the COROP-level of all dependent variables (employment growth, productivity growth, unemployment growth and inactivity), as well as the four main independent variables (related variety, unrelated variety, LOS-index, population density) are presented in this section. Maps of all controlling variables can be found in Appendix II. Descriptive statistics including mean, minimum value, maximum value and standard deviation of all variables are given in table 1.

The choice of COROP as the spatial unit of analysis is motivated by the wish to deal with labour market regions, which are regarded as the most relevant unit of analysis in agglomeration research. In The Netherlands, the COROP level is commonly associated with spatial labour markets. A recent study on functional regions in The Netherlands by Bongaerts *et al.* (2004) confirmed that the functional coherence of the COROP classification is indeed statistically as not less coherent than the classification that can be obtained by empirical computation. We did not consider the use of data at an even lower level of spatial aggregation (municipalities), because the data required to compute sectoral variety were available to us at the COROP level.

7.2.1 Dependent variables

The four dependent variables are shown in figure 1 for all 40 COROP regions.

EMPLOYMENT GROWTH (1996-2002)

Computed as percentage growth over full-time employee equivalents (1996-2002) using data from Statistics Netherlands (CBS). These employment data are also used to construct the 'per fte' variables where employment levels are used as denominator (wage per fte, investment per fte, R&D per fte, see below). Note that the data include all economic activities except agriculture.

PRODUCTIVITY GROWTH (1996-2001)

Computed as percentage growth (1996-2001) and provided by the University of Groningen (Broersma and Oosterhaven 2004).

UNEMPLOYMENT GROWTH (1996-2002)

Computed as percentage growth (1996-2002) and computed from data from Statistics Netherlands (CBS).

INACTIVITY GROWTH (1996-2002)

Data stem from Statistics Netherlands similar to unemployment data and growth is computed as a percentage growth (1996-2002). Inactivity data include both unemployment numbers and the numbers of physically disabled workers (often seen as a hidden form of unemployment, see Broersma and Van Dijk 2002).

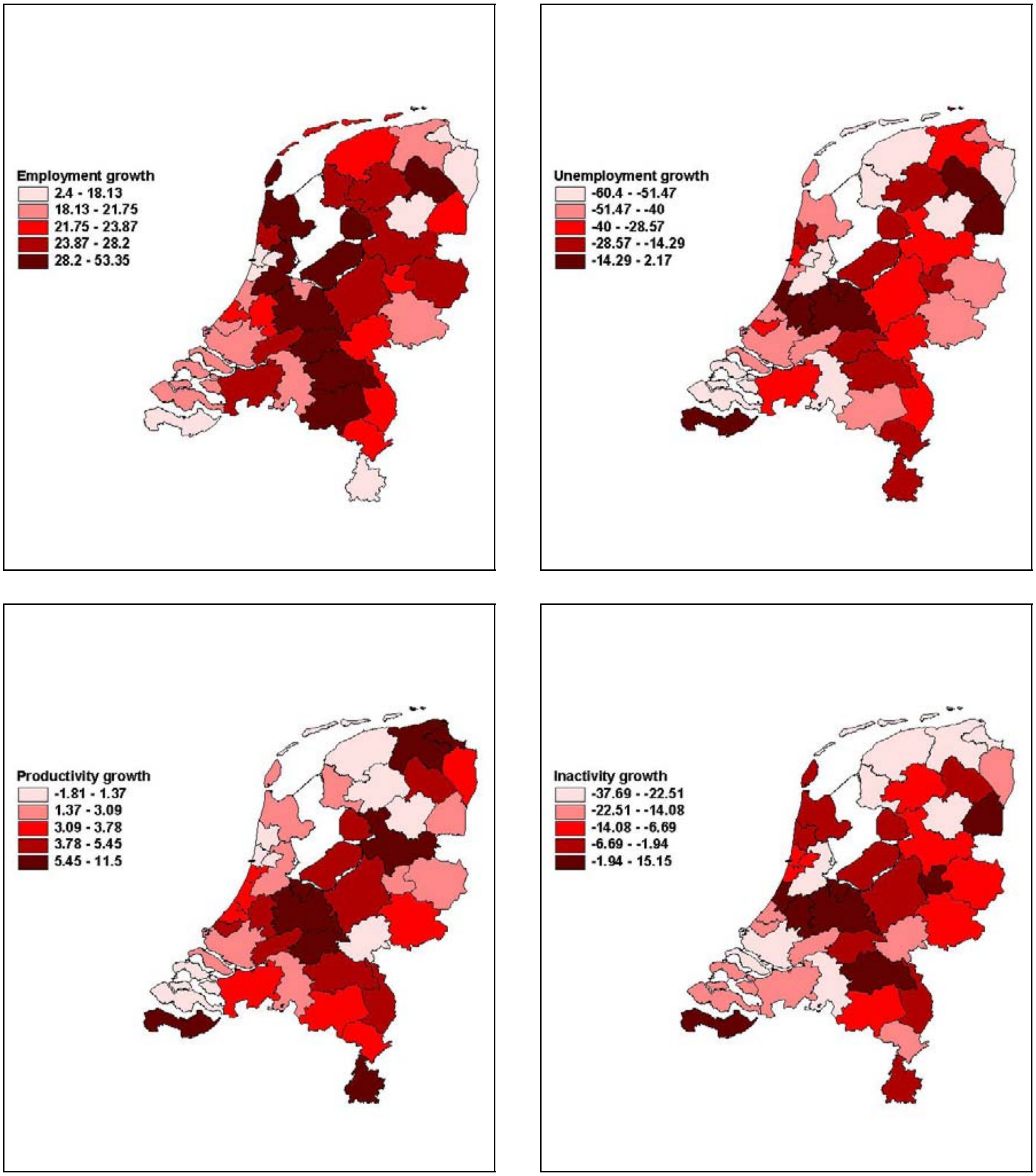


Figure 1. Maps of dependent variables

7.2.2 Independent variables

UNRELATED VARIETY (1996)
RELATED VARIETY (1996)

As explained above, we use entropy statistics to measure sector variety. Marginal variety can be computed at all five digit-levels, meaning the increase in variety when moving from one digit level to the next. As the marginal entropy levels at the three-, four- and five-digit levels are correlated strongly (above 0.7), we chose to compute the marginal increase when moving from the two-digit to the five-digit level instead. This variety indicator we label *related variety* as opposed the two-digit level entropy, which we associate with *unrelated variety* (see figure 2). We will include both types of variety to test whether related variety or unrelated variety have different effects. As argued earlier, we consider related variety to be the indicator for Jacobs externalities because it measures the variety *within* each of two digit classes. We expect the economies arising from variety to be especially strong between sub-sectors, as knowledge spills over primarily between firms selling related products. By contrast, unrelated variety measures the extent to which a region is diversified in very different types of activity. This type of variety is expected to be instrumental in avoiding sticky unemployment. Data on one- to five-digit sectors are based on employment data that stem from the LISA databases on employment in the Netherlands (reworked data from Van Oort 2004).

SPECIALISATION INDUSTRY (1996)
SPECIALISATION DISTRIBUTION (1996)
SPECIALISATION CONSUMER SERVICES (1996)
SPECIALISATION PRODUCER SERVICES (1996)
LOS-INDEX MANUFACTURING (1996)

Localisation economies are associated with the concentration of a particular sector in a region. Often, this type of economies is captured by specialisation indicators (Glaeser *et al.* 1992; Nieuwenhuijsen and van Stel 2000; Van Oort 2004). Furthermore, specialisation measures are useful to tackle the sector-composition effect on employment growth (as in a shift-and-share analysis). For example, in the period considered (1996-2002), one expects regions being specialised in producer services to create more jobs than those specialised in industry (as the negative correlation in table 2 suggests). Using the specialisation measure by Glaeser *et al.* (1992), we distinguish between four types of specialisation, being industrial activities, distribution and transport services, consumer services and producer services (using the LISA data).[11]

We constructed an additional variable called the Los-index (Los 2000) that has not been used hitherto. This index captures the technological relatedness between industrial sectors by computing the similarity between two sectors' input mix from input-output tables. As input mixes reflect production technologies, a high similarity in input mixes of two sectors implies a small 'technological distance' between two sectors, and a high amount of spillovers. Conversely, two industries with very different input mixes are technologically distant, and, consequently, will hardly mutually benefit from spillovers. Technological similarity within a sector is by definition equal to one, as jobs within the same sector are assumed to yield the highest amount of spillovers (underlying the concept of localisation economies). We consider this index to be a better measure of industry specialisation, because (i) it takes into account but whether industrial sectors in a region are technologically related, and (ii) it is not a relative specialisation measure, but it is based on absolute concentration of jobs in particular sectors in a region.

[11] The use of these four main sectors is based on Louter (2000).

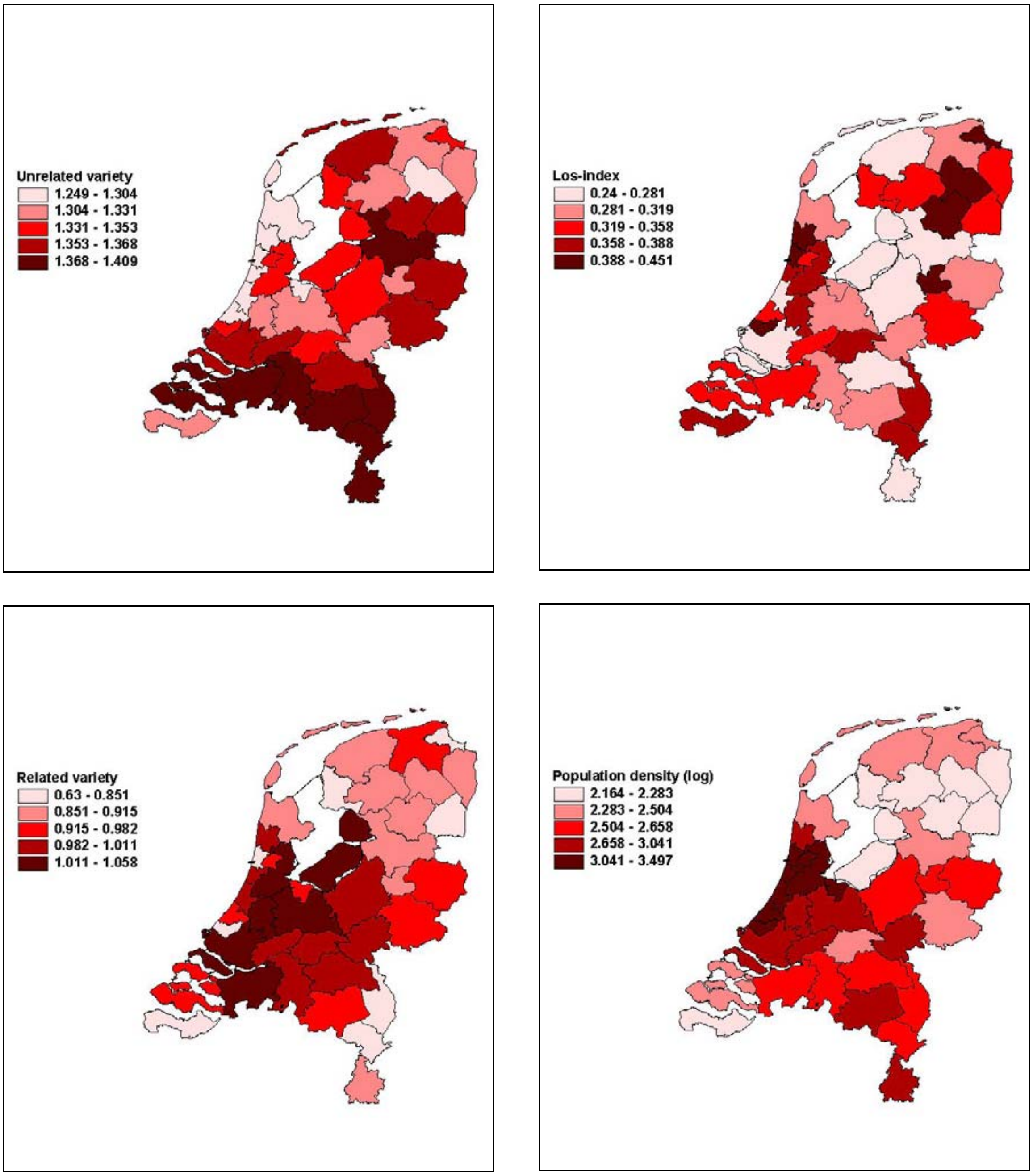


Figure 2. Maps of the four main independent variables

The data on technological similarity are provided by Bart Los from the University of Groningen (Los 2000). We have chosen to apply the measure only to industrial sectors and knowledge intensive service sectors because the concept of knowledge spillovers are known to be strongest in these sectors (including all other services would have substantially lowered the variance in the Los index). The data consists of a matrix of similarity values for each pair of sectors ranging from 0 (no inputs in common) to 1 (all inputs in common). For a region *k*, we multiplied the number of jobs for each pair of sectors. This number is multiplied by the corresponding similarity value between the two sectors. This is repeated for all pairs of sectors. The sum of the pair wise multiplications is finally divided by the maximum possible Los-value (which is obtained if all sectors would have perfect similarity). Let s_{ik} and s_{jk} stand for the number of jobs in sector *i* and *j* respectively, and a_{ij} for the technological similarity value between sector i and j, then the Los index is computed as:

$$Los_k = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} (s_{ik} \cdot s_{jk} \cdot a_{ij})}{\sum_{i=1}^{n} \sum_{j=1}^{n} (s_{ik} \cdot s_{jk})} \qquad (6)$$

This index ranges from the minimum value (1/n) to its maximum value of 1. Note that, as the technological similarity within a sector is by definition equal to one (the diagonal in the similarity matrix), a region which is fully specialised in one sector always acquires the maximum possible value. In all other cases, the Los-value will lie in between the minimum and maximum value (see figure 2).

One could interpret the Los-index as a proxy for *technological clustering*, *i.e.* as indicating the extent to which a regional economy can be characterised as a technological cluster. A value of 1 would indicate the presence of one ideal-type of a *cluster* of technologically perfect related industries (possibly only one industry), in which the amount of localisation economies in a region would be fully maximised.

SPECIALISATION TRADITIONAL MANUFACTURING (1996)

We have constructed an additional specialisation variable regarding the share of traditional manufacturing (using the definition proposed by Louter 2000) to assess whether regions that are 'locked-in' into saturated low-tech sectors experience significantly more problems in creating new jobs.

POPULATION DENSITY (LOG) (1996)

Population density is used as a proximate indicator of urbanisation economies stemming from a large concentration of economic activity *per se i.e.* irrespective of its composition (see also the map in figure 2). The main component of urbanisation economies is the benefits from market size. We have chosen to take the logarithm of this variable reflecting the decreasing marginal benefit from each additional inhabitant in a region. By doing so, we also solve the problem of the positive outlier, which is the agglomeration of The Hague. This region has a density of 3140 inhabitants per square kilometre, which is twice the number of the second dense region (being the region incorporating Haarlem).

7.2.3 Control Variables

WAGE (1996)

Stemming from Statistics Netherlands (CBS) for the year 1996, the average wage levels are used as a control for employment growth differentials due to wage differentials.

INVESTMENT PER FTE (average 1996-2002)

Concern investments in immobile capital goods, excluding houses. The indicator is computed per fte and data are taken from Statistics Netherlands (CBS).

CAPITAL LABOUR RATIO GROWTH (1996-2001)

Capital-labour ratio growth (C-L growth) is computed as percentage growth (1996-2001) and stems from the regional productivity study of Broersma and Oosterhaven (2004). This variable is primarily included to explain productivity growth. Following the production function approach (Solow 1957), a rise in capital over labour will contribute to labour productivity. Alternatively, the C-L growth variable can also be interpreted as process innovation, taken broadly, which is expected to contribute to labour productivity (Kim 1997).

R&D PER FTE (1999)

The data on Research and Development (R&D) refer to the level of investments in R&D in 1999 (as 1996 was not available). Data are taken from Senter (Van Oort 2004).

BUSINESS AREA GROWTH (1996-2002)
DWELLINGS GROWTH (1996-2002)

Newly build residential neighbourhoods and business premises attract economic activity that previously was not present in that location. To control for these potential causes for extreme high differential employment growth we included the growth in business sites in hectares (average 1996-2002) from the IBIS-database (see Van Oort 2002) and the growth of the number of dwellings for the same period taken from Statistics Netherlands (CBS).

COMPETITION (1996)

Following Glaeser *et al.* (1992) and Van Oort (2004), we take into account competition among firms. This indicator tests whether more fierce competition is enhancing employment and/or productivity growth (Porter-thesis) or slowing down employment and productivity growth (Marshall-Arrow-Romer thesis, also known as the MAR-thesis). We computed competition as the number of firms per employee, following Glaeser *et al*. (1992). Critiques on this indicator, mainly stating that it does not measure the level of competition but the ‘average’ industrial organisation structure of sectors in regions, are summarised in Van Oort (2004).

HUMAN CAPITAL (1996)

This knowledge-index based on the highest degree of education of the working labour force in a region progressively weighted for higher degrees of education (polytechnic/academic) and scienific disciplines. This indicator stems from Broersma and Oosterhaven (2004).

RANDSTAD (RS)
INTERMEDIATE ZONE (IZ)
NATIONAL PERIPHERY (PER)

From earlier research (Van Oort 2004) we know that spatial heterogeneity may appear at the level of clusters of regions in the Netherlands, namely whether a region is located in the Randstad (the economic core region of the Netherlands), in the so-called intermediate zone (the regions adjacent to the Randstad region, Gelderland and Noord-Brabant) or in the national periphery. These three variables are used in two ways. First, we use them as dummy variables to test whether the intercept of the estimated model is significantly different in different areas (*i.e.* fixed effects). In this case, we always leave out population density because of the high correlation between population density and the Randstad (.738) and the periphery (-.701). Thus, we capture the concept of urbanisation economies either by including population density or by including dummies. Second, we use these zones to analyse *spatial regimes* using either Ordinary Least Squares (OLS) or Maximum Likelihood (ML) based spatial lag and spatial error models (and including population density as a dependent variable). These models simultaneously (using change-of-slope principles) test for significant differences between the estimated coefficients between the regimes (Anselin 1988).

7.3. Descriptive statistics

Descriptive statistics are given in table 1. The values shown are the values following the definitions presented in the previous section. In the following, we use standardised scores (z-values with average 0 and standard deviation 1) of all variables in order to assess the relative effect of independent variables. Some variables have been corrected[12] for outliers with absolute z-values larger than three. These concern employment growth (Flevoland), related variety (Delfzijl), investment (Flevoland), R&D intensity (Zuidoost Brabant), business area growth (Midden-Brabant), dwellings growth (Flevoland) and the human capital indicator (Groot-Amsterdam).

COROP-maps of the four dependent variables are provided in figure 1. Importantly, we observe two different patterns. Employment growth seems to be concentrated in Flevoland and along the A2 highway connecting Amsterdam via Utrecht to Eindhoven. Most peripheral regions experienced relative low employment growth during our research period. Productivity growth, unemployment growth and inactivity growth, by contrast, are less organised spatially and seems to be distributed almost randomly.

From the correlations presented in table 2, it is clear that many variables are highly correlated. This may cause multicollinearity problems in our regression analysis. For this reason, estimation of a fully specified model including all independent variables is not possible. Instead, we start from theoretically based baseline models in which we include the most relevant variables, which are the indicators related to the different types of agglomeration economies: *UNRELATED VARIETY* (to test for the portfolio effect), *RELATED VARIETY* (to test for Jacobs externalities), the *LOS-INDEX* (to test for localisation economies), and *POPULATION DENSITY* (to control for pure urbanisation economies). Including all these variables allows us to assess the relative effect of different potential sources of agglomeration economies. Note that the correlations between these four variables are all below 0.5.

[12] Corrections are carried out by (1) in a first stage excluding the outlier when computing z-values – allowing variation in the remaining non-outlier observations – and (2) in a second stage incorporating the outliers with a relative high value (3 or minus 3) in the dataset (the outliers do measure reality, and should not be completely excluded from analyses).

The specialisation measures are not introduced in the models, because of the high correlation with the other agglomeration and variety indicators. In particular, *UNRELATED VARIETY* correlates highly with specialisation in industry and traditional manufacturing (+), and with specialisation in consumer services (-), while *RELATED VARIETY* correlates positively with specialisation in distribution and producer services, and negatively with specialisation in industry. Furthermore, most specialisation patterns are correlated with Randstad and periphery (with inverse signs), which reflects the basic spatial structure of producer and consumer services' relative overrepresentation in the Randstad and industrial activities in the national periphery[13], with the intermediate zone indeed being characterised by an 'intermediate' sector structure (see also the respective maps in Appendix II).

As the main control variables, we have chosen to include *INVESTMENT* and *R&D*. In addition, when dealing with productivity growth and unemployment, we included *CAPITAL LABOUR RATIO GROWTH* as a control. There are both theoretical (Solow 1957) and empirical (Broersma and Oosterhaven 2004; Kim 1997) reasons to assume that productivity growth is very sensitive to this ratio as it increases the amount of capital per worker (*e.g.* through ICT investments). Concerning unemployment, an increase in the ratio between capital and labour may indicate labour-saving technological change, and thus, may raise unemployment. Finally, we also included *WAGE* in our baseline model explaining unemployment growth, because regions with higher relative wage levels are expected to experience higher unemployment, *ceteris paribus*. All other variables are added one-by-one to the baseline model to assess whether the specification of the model improves. If so, these variables are shown in the results.

7.4 Factor analysis

In addition to the descriptive statistics we applied factor analysis to uncover the correlated structures underlying our main dependent variables. By doing so, we get a better insight in the qualitative differences in terms of variety, urbanisation and specialisation across regions in The Netherlands. Thus, we use factor analysis here purely as a descriptive tool. Factor analysis is used to remove the overlap between the different variables. What remains after the analysis, are two or more factors that group the variables with highest similarity. We included nine variables in the factor analysis: the four main independent variables (related variety, unrelated variety, Los-index and population density) and five specialisation measures (industry, producer services, distributions services, consumer services, and traditional manufacturing) as to obtain two factors. Factor scores are shown in Figure 3 (individual scores of regions on each factor can be found in table 6). Regions with a relative high score for the factor 1 can be characterised as urban regions with producer services and related variety. As can be seen from the map, high scores on factor one are mainly in the Randstad, especially in the G4-regions, while low scores are found in the peripheral regions.[14] Regions scoring high on factor 2 can be described as industry and distribution regions with unrelated variety. Factor 2 is highly present in the so-called Betuwe-corridor (distribution axis from Rijnmond to Germany) and in the industry-oriented regions of Noord-Brabant. Note that the Los-index has for both factors a low and negative value, and therefore is of little importance for the factor determination.

[13] Especially Delfzijl, IJmond and Zeeuws-Vlaanderen.

[14] G4-regions are the COROP regions with the four largest Dutch cities (Agglomeratie 's Gravenhage, Groot-Amsterdam, Rijnmond and Utrecht).

	Factor 1	Factor 2
Unrelated variety	-.336	.766
Related variety	.774	.254
Los-index	-.310	-.166
Log population density	.715	-.199
Specialisation industry	-.872	.365
Specialisation distribution	.320	.714
Specialisation producer services	.868	-.098
Specialisation consumer services	.122	-.854
Specialisation traditional manufacturing	-.688	.488

(Extraction method: principal component analysis. Rotation method: varimax with Kaiser normalization)

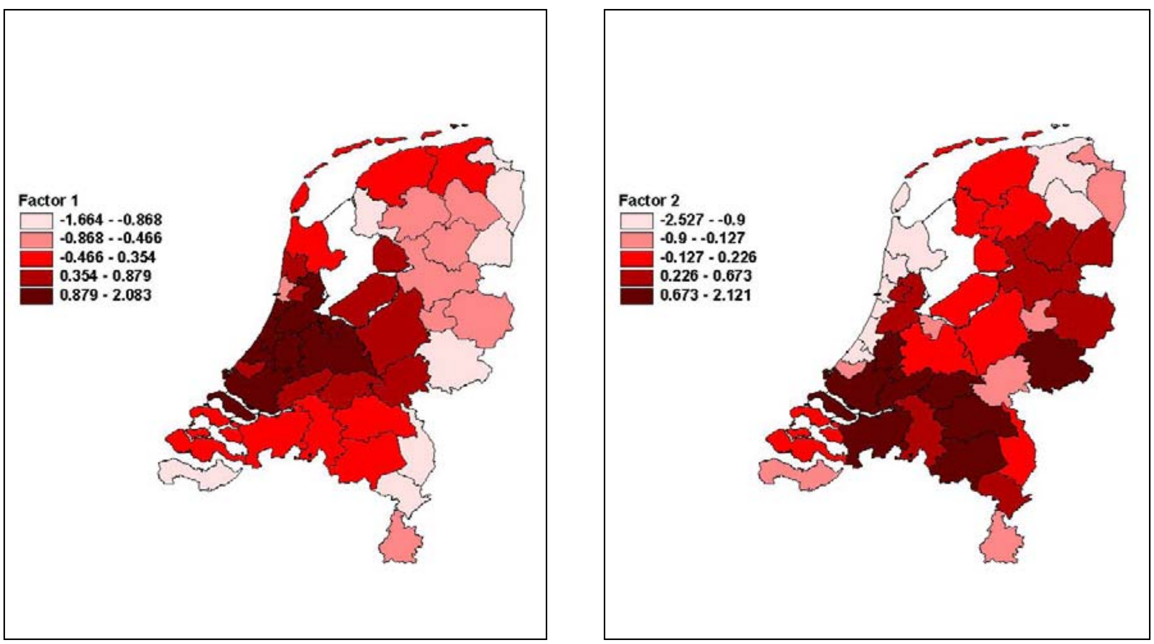


Figure 3. Maps of factor scores

7.5 Regression results

7.5.1 Results for employment growth

Table 3 provides the results for *EMPLOYMENT GROWTH* as the dependent variable. Model 1 specifies the OLS baseline model. From the results, it can be concluded that our main hypothesis is **confirmed**: related variety as an indicator for Jacobs externalities is indeed positively and significantly related to employment growth.[15] Since we used z-values, the results also show that related variety contributes most to employment growth. Furthermore, investment as a control variable has the expected sign. Interestingly, population density has no significant effect on employment growth suggesting that it is not urbanisation *per se* that contributes to job creation, but related variety (which correlates with population density, see table 2). Put it differently, cities do not create jobs 'automatically'. Rather, related variety is responsible for job creation, which is often, but not necessarily, highest in cities.

Models 1a and 1b test for the robustness of model 1, by substituting the dependent variable, employment growth during the period 1996-2002, by the same variable for different periods (1997-2002) and (1996-2001). The results show that model 1 is robust in the sense that the same variables are significant (and of the same sign) in models 1a and 1b.

Using model 1, we added, one-by-one, all other variables. None of these variables additionally turned out to be significantly related to employment growth except for the average wage level (model 2) and dwellings growth (model 3). In both cases, investment was no longer significant, yet the significance and sign of related variety proved to be robust. Model 2 suggests that employment has been created in high-wage areas. This is contradictory to the traditional expectation that low wage levels attract investment, and by doing so, enhance employment growth. This outcome may reflect the higher human capital levels in high-wage regions (although our human capital variable did not prove to be significant when added to the baseline model). High wages may also have acted as a trigger to migrate, and by doing so, raise employment/supply of labour (compare Broersma and Van Dijk 2002). This is akin to the core mechanism explaining agglomeration in models of the new economic geography. Note that including the wage variable renders population density significant and negative (probably due to relatively high correlation between wages and population density). Model 3 is relevant as it shows the effect of the growth of dwellings on employment growth. This relates to the issue of whether investments in dwellings attract jobs ('work follows living') or *vice versa* ('living follows work'). Employment growth is partly dependent on dwellings as new inhabitants attract new employment in most service sectors.[16]

We also tested whether employment growth is spatially autocorrelated, *i.e.* whether fast (slowly) growing regions are neighbours of other fast (slowly) growing regions. This is done by computing the Lagrange multiplier for the error term and for the spatial lag of the dependent variable (see Model 1, 2 and 3). Exploratory spatial analysis using Spacestat estimation software (Anselin 1988) revealed that a simple contiguity matrix of adjacency between the 40 COROP-regions best captures the spurious spatial dependence between regional scores.[17] The dependence is spurious because the COROP-level turned out to be a robust measurement level in spatial statistical terms: no variation between regional indicators can significantly be attributed to spatial

[15] We also estimated the baseline model without *UNRELATED VARIETY* and without *RELATED VARIETY*, while including five-digit entropy instead as one comprehensive 'total variety' indicator. This did not alter the results.

[16] Clearly, this relates to an endogeneity problem.

[17] We also tested for the sensitivity for higher order contiguity spatial dependence and for first- and second order inverse distance weights using physical distances (kilometres) – and none of these spatial weight formulations captured spatial dependence significantly better.

correlation. In all employment growth models presented in table 3, the LM-test statistics indeed presented no significant indications for spatial lag or spatial error specifications of the models (all p-values are well above 0.10), which implies that the model structure and model fit do not gain from spatial error or spatial lag specifications.

Although spatial adjacency is not a characteristic of our growth and variety models, spatial dependence on an even higher spatial level was also hypothesised. We thus also ran OLS regressions for different spatial regimes. Model 4 shows the results for regime analysis distinguishing the Randstad from the rest of The Netherlands, and Model 5 shows the results for Randstad, intermediate zone and national periphery regimes (see Appendix II for the definition of these zones). In both cases, the Chow-Wald test indicates that the differences between zones are not significant. However, the effect of individual variables can be significantly different in different zones (as indicated by a grey shading in the table). Only investment levels show a different relation with employment growth in different regimes with its effect being significantly positive in the Randstad, while not significant in the rest of the country.

Finally, spatial dependence can occur in the independent variables of the model. Therefore we repeated the specification in model 1 using the window-average (WA) values of the independent variables. WA-values are the average of the value of a COROP region and all its neighbouring regions.[18] Thus, in a specification with WA-variables, independent variables are measured at the supra-regional level of the region itself and its neighbouring regions, thus taking into account the effects of nearby regions on a region's growth (*e.g.*, demand effects, crowding out, spillovers, etc.). From the specification including the WA-variables in Model 6 it can be concluded that only related variety positively affects employment growth using WA-variables, while localisation economies as indicated by the LOS-index now have a significant negative effect. This result reinforces our conclusion that, as hypothesised, related variety is a main driver of employment.

7.5.2 Results for productivity growth

Table 4 provides in a similar manner as table 3 the results for *PRODUCTIVITY GROWTH* as the dependent variable. Model 1 specifies the OLS baseline model, which corresponds to the baseline model for employment growth plus C-L growth. Investment, R&D and C-L growth are all significant and positively related to regional productivity growth, as expected. Related variety is also significant, but negatively related to productivity growth. This means that whereas related variety contributed to employment growth, it slows down productivity growth. Our main hypothesis concerning productivity growth – localisation economies enhancing productivity growth – is **not confirmed**. The Los-index does not contribute to productivity growth neither does the inclusion of any of the specialisation variables (not shown).

Models 1a and 1b again test for the robustness of model 1, by substituting the dependent variable, productivity growth during the period 1996-2001, by the same variable for different periods (1997-2001) and (1996-2000). Model 1 is not entirely robust for changes in the period of observation as investment and related variety are significant either model 1a or model 1b, but not

[18] We used the first-order contiguity matrix for calculating WA-values in Spacestat (Anselin 1988). It is important to note though that the window average of entropy values (used to indicate unrelated and related variety) and the Los-index cannot be computed as the average of a region and its neighbours, because these indices reflect a qualitative state of the economy rather than a quantitative value. Two regions with the same sector structure (say, 50% shipbuilding and 50% textiles) have the same entropy value (H=1), but regions with the same entropy value (H=1) do not necessarily share the same sector structure, but only the distribution (for example, one region can have 50% shipbuilding and 50% textiles and another region 50% shipbuilding and 50% transport equipment). When distributions are aggregated across regions, the window average entropy is to be computed from the newly obtained frequency distribution at the supra-regional level. The window average values for unrelated variety, related variety and the Los-index are given in the form of maps as well.

in both. Conclusions about these two variables should therefore be drawn with care. The variables R&D and C-L growth show robustness in the sense that their sign and significance remained unchanged. Again, using model 1, we added, one-by-one, all other variables. None of these variables turned out to be significantly related to productivity growth (not shown), while the variables that were significant in Model 1 remain robust.

We tested whether productivity growth is spatially autocorrelated by interpreting again the Lagrange multiplier test statistics for a spatial error term and for the spatial lag of the dependent variable (again using a first-order contiguity matrix). The Lagrange multiplier value for spatial lag is significant at the 10% level (0.068), which means the model specification can be improved by including a spatial lag of the dependent variable, which is the average productivity growth in a region's neighbouring regions. Model 2 shows the results of the spatial lag model. Interestingly, the spatial lag of productivity growth (W_productivity growth) is significant, yet negative. This means that there is an inverse relationship between productivity growth in a region and its neighbouring regions: regions surrounded by low productivity growth tend to have high productivity growth and *vice versa*. This result underlines that the choice of COROP as the unit of analysis is justified as no positive relations can be found at the supra-regional level.

The spatial lag specification is continued in Models 3 and 4 where we introduce the spatial regimes. The Chow-Wald statistic shows that only the spatial regimes as specified in Model 4 are significantly different from one another. Thus, productivity growth is explained differently in different areas in The Netherlands. In particular, population density affects productivity growth negatively in the Randstad (suggesting diseconomies of agglomeration), and positively in the periphery (suggesting economies of agglomeration).[19] Finally, the window average specification of the baseline model shows that R&D and C-L growth also remain positive in that specification.

Concluding, most specifications show that the main drivers of productivity growth are the 'usual suspects' of R&D and C-L growth, both commonly associated with process innovation. The effects of population density and investment levels are specific for different spatial regimes. Importantly, the spatial-lag results show negative spatial autocorrelation with neighbouring regions, which supports the choice of COROP regions as the relevant spatial level of analysis.

7.5.3 Results for unemployment growth

Table 5 provides the results for *UNEMPLOYMENT GROWTH* and *INACTIVITY GROWTH* as dependent variables. Model 1 specifies the OLS baseline model, which is equal to the baseline model for productivity but including wage as an additional control variable. From the results it can be concluded that our main hypothesis concerning unemployment growth – unrelated variety is negatively related to unemployment growth – is **confirmed**. This means that regions with higher unrelated variety experience lower rates of unemployment growth. Furthermore, we find a negative significant relation between urbanisation economies and unemployment growth. This can be explained by the fact that regions with high population densities are also regions where unemployed have more job opportunities within commuting range (compare Broersma and Van Dijk 2002). Urbanisation economies, therefore, provide a safeguard against high unemployment growth. We also find that regions with relative high R&D expenditures per fte and C-L growth experience higher unemployment growth, which suggests that some part of innovative activity is labour-saving. Finally, we find the expected effect of wages on unemployment.

Models 1a and 1b test for the robustness of model 1, by substituting unemployment growth during the period 1996-2002, by the same variable for different periods (1997-2002) and (1996-2001). Model 1c provides an additional robustness check by using *INACTIVITY GROWTH* (including physically disabled in the inactivity-definition besides unemployed persons) as an

[19] COROP regions in the periphery with relatively high population density are all regions in Limburg, Twente and Deventer.

alternative unemployment measure for the same period. The results on robustness show that the baseline model is not entirely robust for changes in the period of observation in particular with regard to wage and C-L growth. Note that the unrelated variety, which is of main interest to our analysis of portfolio effects, is significant in model 1c as well as in particular spatial regimes. In some specifications, however, related variety seems to take over the effect of unrelated variety. As for the regressions on employment growth and productivity growth, we used the baseline model 1 to add the other dependent variables one-by-one. None of these variables proved to be significantly related to unemployment growth.

It is of no help to include a spatial error or spatial lag specification of the dependent variable: the LM-test statistics do not suggest so. In spite of the missing spatial autocorrelation we find that distinguishing between spatial regimes is informative. From the Chow-Wald statistic, it is clear that the distinction between Randstad and not-Randstad as two spatial regimes is significant. This is especially due to differences in the effect of population density (highly negative in the Randstad regions and not significant outside the Randstad), which suggests that unemployed in the Randstad profit from the high density of potential jobs in their region. Further differentiating between Randstad, intermediate zone and national periphery does not yield significant results. Finally, the window average specification of the baseline model shows that, when assuming neighbouring regions affect a region's unemployment, related variety, population density and investment prove to counter-act unemployment, while high wages and technological clustering enhance unemployment.

Concluding, though evidence has been found that unrelated variety counter-acts unemployment as portfolio theory predicts, in some specifications this effect seems to be replaced by related variety. This can be understood from the fact that related variety *also* acts as a portfolio strategy, but less effectively so (as demand cycles are more correlated between sub-sectors below the two-digit level). Another important result is the dampening effect of population density on unemployment growth in the Randstad. The effects of control variables are not entirely robust, although the negative effect of high wages on unemployment is, as expected, significant in most model specifications.

8. Conclusions and policy implications

The goal of our study has been to analyse the effects of variety on regional economic growth. The main contribution has been to distinguish between unrelated variety and related variety. Unrelated variety is measured at the sector level (two-digit), while related variety is measured at the sub-sector level (within two digit classes). We found that the two variables had very different effects on productivity, employment and unemployment. Previous studies measured variety only in terms of what we have called unrelated variety, and therefore ignored the important effects of related variety (Glaeser *et al*. 1992; Nieuwenhuijsen and Van Stel 2000; Acs 2002). The results of these studies should therefore be regarded as less reliable in this respect.[20]

We associated related variety with Jacobs-type externalities arising from spillovers between sectors stimulating employment creation (hypothesis 2), and unrelated variety with a portfolio that prevents regions from experiencing shocks in unemployment (hypothesis 3). We did not focus only on the effects of related variety and unrelated variety on regional development, but also on localisation economies stemming from regional specialisation and on urbanisation economies stemming from population density. In particular, we expected that localisation

[20] Even though we feel that progress has been made in this study by distinguishing between variety at different levels of sectoral aggregation, the results remain sensitive to the given Standard Industry Classification that traditionally overemphasises industrial sectors over service sectors. Future studies could attempt to make an alternative sectoral aggregation scheme.

economies, as present in specialised technological clusters, would primarily enhance productivity growth (hypothesis 1). Using the variables related variety, unrelated variety, localisation economies and urbanisation economies, our study analysed all possible sources of agglomeration economies at the regional level (COROP). Control variables including investment, R&D, growth in the capital-labour ratio, human capital, and wage level were also taken into account.

The empirical results show that related variety indeed enhances employment growth (hypothesis 2), while other type of agglomeration economies are not significant. Knowing that related variety is mainly present in densely populated areas, and given that population density is not significantly affecting employment growth, we can conclude that related variety in cities is responsible for job creation and not urban density in itself.[21] This outcome is also in line with evolutionary economics and urban lifecycle theory that predict new employment stemming from product innovation and new firm creation, to emerge in diversified cities, while labour-saving productivity growth is more likely to be realised by large established firms located in more rural areas.

We also found that unrelated variety is indeed negatively related to unemployment growth meaning that the presence of unrelated sectors in a region acts as a portfolio against unemployment shocks (hypothesis 3). Another result has been that population density dampens unemployment growth in the Randstad. Higher wages, as expected, enhance unemployment growth. Using statistical robustness techniques, the results on unemployment were shown not to be entirely robust. In some specifications related variety took over the effect of unrelated variety, which may suggest that related variety may also have a portfolio effect, though possibly to a lesser extent.

Concerning productivity growth, we obtain more ‘classical’ results with investment, R&D and C-L growth being the drivers behind productivity increases. Hypothesis 1 could not be supported because all our indicators for localisation economies were not significantly affecting productivity growth. Further analysis on productivity growth using three spatial regimes (Randstad, intermediate zone, national periphery), showed that population density has an effect on productivity growth, but differently so in different regimes: population density affects productivity growth negatively in the Randstad (suggesting diseconomies of agglomeration) and positively in the national periphery (suggesting economies of agglomeration).

From our study, and given statistical error, it follows that employment policy should stimulate related variety of any sort. Related variety is primarily present in urban areas, which provides a rationale to formulate employment policy in urban settings. However, one would be mistaken to think that agglomeration enhances job creation *per se*, because population density has no significant effect on employment growth. Rather, policy could try to actively promote the development of related sectors within a region, for example, by stimulating high-end niches in existing sectors and spin-offs firms of existing firms.

Productivity policy, by contrast, could still rely on traditional generic instruments stimulating R&D and investment, and on policies increasing the capital-labour ratio (*e.g.*, ICT policy). These productivity determinants can be the objective of national policy, due to their generic nature, even though we found that productivity growth is spatially differentiated at the COROP level. The spatial pattern of productivity growth is better understood as regional-specific outcomes of national policy (and international market forces).

Given the rather mixed results on unemployment, implications for policy are more speculative. Regional economies with unrelated sectors indeed experienced less unemployment growth. In principle, policy could take into account the beneficial effects of unrelated variety. However, this should not be taken to mean that we advise regions to develop more unrelated variety, because the development of a completely new sector is a highly risky endeavour. Stimulating the supply

[21] Delft and The Hague are two examples of cities with a low level of related variety relative to cities with similar population densities.

of human capital and university research can be more effective in this context, as these will raise the probability of an endogenous generation of new sectors in the long run, without the necessity to place one's bets on a specific sector *ex ante*.

The role of cities seems to be different in different spatial regimes: cities in the Randstad have a high degree of related variety and typical job-creating poles (services mainly). At the same time, the high population density in the Randstad acts as an unemployment safeguard. Cities in the periphery, by contrast, primarily support capital-intensive sectors and hereby productivity growth. Thus, policy makers should recognise that the Randstad is most dynamic in the creation of new jobs and new (service) industries, but also that productivity growth still relies heavily on the industrialised regions outside the Randstad. Both dynamics are necessary for long-term economic development, yet policies supporting the one dynamic need not be supportive of the other, and *vice versa*.

Concerning the recent debate on regional economic policy in The Netherlands, two observations can be made. A first observation, in line with a recent study by Raspe *et al.* (2004) on knowledge and economic growth at the municipality level, holds that the dynamics of economic growth indeed take place at a low level of spatial aggregation (COROP or even lower). This does not imply that economic policy should therefore be regionalised *per se*, but it suggests that the effects of national economic policy may have profound differential effects at the COROP level. In this context, the recent proposal to pin down a number of quite large 'knowledge regions', such as the Aachen-Leuven-Eindhoven region or the Wageningen-Arnhem-Nijmegen-Enschede region, is questionable (Ministerie van Economische Zaken 2004).

A second observation concerns the recurrent tendency of policy makers to select *a priori* particular sectors and/or particular regions as their policy objective. Advocates of such a 'picking-the-winner' policy argue that promising developments in an economy tend to cluster in specific sectors and in specific regions. Others argue that structural change in an economy, though extremely important in itself, should be encouraged by generic policy because the regional outcomes of structural change are fundamentally uncertain. Policies based on supporting related variety, which we found to be the driving force behind employment growth, may take an intermediate position. On the one hand, the risk of selecting wrong activities is reduced because one takes existing regional competences as building blocks to broaden the economic base of the region. On the other hand, such a policy could still acknowledge the fact that generic technologies (like ICT) may have a huge and pervasive impact on economic development in many regions due to the many potential fields of application. A regional policy based on related variety combines the advantages of specialisation and variety, and is to be supplemented by national policies on generic technologies.

References

Abernathy, W.J., and J. Utterback (1978), Patterns of industrial innovation. *Technology Review* 50, pp. 41-47.

Acs, Z. (2002), *Innovation and the Growth of Cities*. Cheltenham and Northampton MA: Edward Elgar.

Anselin, L. (1988), *Spatial Econometrics: methods and models*. Dordrecht: Kluwer.

Attaran, M. (1985), Industrial diversity and economic performance in U.S. areas. *The Annals of Regional Science* 20, pp. 44-54.

Audretsch, D.B., and M.P. Feldman (1996a), R&D spillovers and the geography of innovation and production. *American Economic Review* 86(3), pp. 630-640.

Audretsch, D.B., and M.P. Feldman (1996b), Innovative clusters and the industry life cycle. *The Review of Industrial Organization* 11, pp. 253-273.

Baldwin, J.T., and W.M. Brown (2004), Regional manufacturing employment volatility in Canada: the effects of specialisation and trade. *Papers in Regional Science* 83(3), pp. 519-541.

Bhattacharya, R. (2003), Sources of variation in regional economies. *Annals of Regional Science* 37, pp. 291-302.

Bleichrodt, H., P.J. Louter and W.F. Sleegers (1992), *Jonge bedrijven in Nederland*, EGI-onderzoekspublikatie 02, Rotterdam: Erasmus Universiteit.

Bongaerts, D., F. Cörvers and M. Hensen (2004), *The Delineation and Coherence of Functional and Administrative Regions in the Netherlands.* The Hague: Ministry of Economic Affairs.

Boschma, R.A., and K. Frenken (2003), Evolutionary economics and industry location. *Review of Regional Research* 23(2), pp. 183-200.

Boschma, R.A., and J.G. Lambooy (1999), Evolutionary Economics and Economic Geography. *Journal of Evolutionary Economics* 9, pp. 411-429.

Boudeville, J.R. (1966), *Problems of Regional Economic Planning*, Edinburgh: University Press.

Brakman, S., H. Garretsen and C. van Marrewijk (2001), *An Introduction to Geographical Economics*. Cambridge: Cambridge University Press.

Braczyk, H.J., P. Cooke and H. Heidenreich, (1998) (eds.), *Regional Innovation Systems. The role of governances in a globalized world*. London: University College London Press.

Breschi, S., and F. Lissoni (2001), Knowledge spillovers and local innovation systems: A critical survey. *Industrial and Corporate Change* 10, pp. 975-1005.

Breschi, S., and F. Lissoni (2003), Mobility and social networks: Localised knowledge spillovers revisited. *CESPRI Working Paper* 142, www.cespri.it

Brewer, H.L. (1985), Measures of diversification: predictors of regional economic instability. *Journal of Regional Science* 25, pp. 463-470.

Broersma, L., and J. van Dijk (2002), Regional labour market dynamics in the Netherlands. *Papers in Regional Science* 81(3), pp. 343-364.

Broersma, L., and J. Oosterhaven (2004), *Regionale Arbeidsproductiviteit: Niveau, groei en verklaring*. The Hague: Ministry of Economic Affairs.

Brouwer, E., H. Budil-Nadvornikova and A. Kleinknecht (1999), Are urban agglomerations a better breeding place for product innovation? An analysis of new product announcements. *Regional Studies* 33, pp. 541-49.

Carlton, D.W. (1982), The location and employment choices of new firms: an econometric model with discrete and continuous endogenous variables. *The Review of Economics and Statistics* 27, pp. 440-49.

Chamberlin, E.J. (1933), *The Theory of Monopolistic Competition*. Cambridge Mass.: Harvard University Press.

Chapman, K., and D.F. Walker (1991), *Industrial Location. Principles and policies*. Oxford: Blackwell.

Dissart, J.C. (2003), Regional economic diversity and regional economic stability: research results and agenda. *International Regional Science Review* 26, pp. 423-446.

Dixit, A.K., and J.E. Stiglitz (1977), Monopolistic competition and optimum product diversity. *American Economic Review* 67, pp. 297-308.

Dumais, G., G. Ellison and E.L. Glaeser (2002), Geographic concentration as a dynamic process. *The Review of Economics and Statistics* 84, pp. 193-204.

Duranton, G., and D. Puga (2000), Diversity and specialisation in cities: Why, where and when does it matter? *Urban Studies* 37(3), pp. 533-555.

Eaton, B.C., and R.G. Lipsey (1975), The principle of minimum differentiation reconsidered: some new developments in the theory of spatial competition. *Review of Economic Studies* 42, pp. 27-49.

Feldman, M.P., and D.B. Audretsch (1999), Innovation in cities: Science-based diversity, specialization and localized competition. *European Economic Review* 43(2), pp. 409-429.

Feser, E.J. (2002), Tracing the sources of local external economies. *Urban Studies* 39, pp. 2485-2506.

Fingleton, B. (2003), Externalities, economic geography and spatial econometrics: conceptual and modelling developments. *International Regional Science Review* 26, pp. 197-207.

Freeman, C. (1987). *Technology policy and economic performance: lessons from Japan*. London: Pinter.

Frenken, K. (2005), Entropy statistics and information theory, in H. Hanusch and A. Pyka (eds.), *The Elgar Companion to Neo-Schumpeterian Economics.* Cheltenham: Edward Elgar, forthcoming, http://econ.geog.uu.nl/frenken/frenken.html.

Frenken, K., P.P. Saviotti and M. Trommetter (1999), Variety and niche creation in aircraft, helicopters, motorcycles and microcomputers. *Research Policy* 28, pp. 469-488.

Fujita, M., and J.F. Thisse (2002), *Economics of Agglomeration. Cities, industrial location and regional growth.* Cambridge: University Press.

Glaeser, E.L., H. Kallal, J. Scheinkman and A. Shleifer (1992), Growth in cities. *Journal of Political Economy* 100, pp. 1126-1152.

Gordon, I.R., and P. McCann (2000), Industrial clusters: complexes, agglomeration and/or social networks? *Urban Studies* 37, pp. 513-532.

Griliches, Z. (1992), The search of R&D spillovers. *Scandinavian Journal of Economics* 94, pp. 29-47.

Harrison, B., M.R. Kelley and J. Gant (1997), Innovative firm behavior and local milieu: exploring the intersection of agglomeration, firm effects, and technological change. *Economic Geography*, 72, pp. 233-58.

Haug, P. (2004), Diversifikation und regionale Wirtschafts- und Beschäftigungsentwicklung. Eine empirische Analyse für ausgewählte deutsche Gebiete. *Review of Regional Research* 24(2), pp. 177-195.

Helsey, R.W., and W.C. Strange (2002), Innovation and input sharing. *Journal of Urban Economics* 51, pp. 25-45.

Henderson, J.V. (1974), The sizes and types of cities. *The American Economic Review*, 64, pp. 640-56.

Henderson, J.V. (1997), Externalities and industrial development. *Journal of Urban Economics*, 42, pp. 449-70.

Henderson, J.V. (2003), Marshall's scale economies. *Journal of Urban Economic*s 53, pp. 1-28.

Henderson, J.V., A. Kuncoro and M. Turner (1995), Industrial development in cities. *Journal of Political Economy* 103, pp. 1067-1085.

Hoover, E.M. (1948), *The location of economic activity*. New York: McGraw Hill.

Hotelling, H. (1929), Stability in competition. *Economic Journal* 39, pp. 41-57.

Hymer, S.H. (1976), *The International Operations of National Firms*. Cambridge: The MIT Press.

Isard, W. (1956), *General theory: social, political, economic and regional with particular reference to decision-making analysis*. Cambridge, MA: MIT Press.

Izraeli, O., and K.J Murphy (2003), The effect of industrial diversity on state unemployment rate and per capita income. *The Annals of Regional Science* 37, pp. 1-14.

Jacobs, J. (1969), *The Economy of Cities*. New York: Vintage.

Jacquemin, A.P., and C.H. Berry (1979), Entropy measure of diversification and corporate growth. *Journal of Industrial Economics* 27(4), pp. 359-369.

Jaffe, A.B. (1986), Technological opportunity and spillovers of R&D. *American Economic Review* 76, pp. 984-1001.

Jaffe, A.B. (1989), Real effects of academic research. *American Economic Review* 79, pp. 957-970.

Jaffe, A.B., M. Trajtenberg and R. Henderson (1993), Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108(3), pp. 577-598.

Kim, S.J. (1987), Diversity in urban labour markets and agglomeration economies. *Papers in Regional Science* 62, pp. 57-70.

Kim, S.J. (1997), *Productivity of cities*. Aldershot: Ashgate.

Klepper, S. (1997), Industry lifecycles. *Industrial and Corporate Change* 6, pp. 145-182.

Knaap, T. (2004), *Models on Economic Geography. Dynamics, estimation and policy evaluation*. SOM Research School, University of Groningen (dissertation).

Kort, J.R. (1981), Regional economic instability and industrial diversification in the US. *Land Economics* 57, pp. 596-608.

Krugman, P.R. (1991), Increasing returns and economic geography. *Journal of Political Economy* 99(3), pp. 483-499.

Krugman, P.R. (1993), Lessons of Massachusetts for EMU, in: F. Torres and F. Giavazzi (eds.), *Adjustment and Growth in the European Monetary Union* (Cambridge: Cambridge University Press), pp. 241-269.

Kuznets, S. (1965), *Economic Growth and Structure*. New York: Norton.

Lancaster, K.J. (1979), *Variety, Equity and Efficiency*. New York: Columbia University Press.

Laursen, K. (2000), Do export and technological specialisation patterns co-evolve in terms of convergence or divergence? Evidence from 19 OECD countries, 1971-1991. *Journal of Evolutionary Economics* 10 (4), pp. 415-436.

Los, B. (2000), The empirical performance of a new inter-industry technology spillover measure, pp. 118-151 in: P.P. Saviotti and B. Nooteboom (eds.), *Technology and Knowledge.* Cheltenham: Edward Elgar.

Louter, P.J. (2000), *Rangen en Standen in Gemeenteland. De economische vitaliteit van gemeenten en stadsgewesten.* Delft: TNO-Inro.

Lucas, R.E. (1988), On the mechanics of economic development *Journal of Monetary Economics* 22(1), pp. 3-42.

Markusen, R.L. (1985), *Profit Cycles, Oligopoly and Regional Development*. London: MIT Press.

Ministerie van Economische Zaken (2004), *Pieken in de Delta. Gebiedsgerichte economische perspectieven*. Den Haag: Ministerie van Economische Zaken.

Montgomery, C.A. (1994), Corporate diversification. *Journal of Economic Perspectives* 8(3), pp. 163-178.

Montobbio, F. (2002), An evolutionary model of industrial growth and structural change. *Structural Change and Economic Dynamics* 13(4), pp. 387-414.

Moomaw, R.L. (1998), Agglomeration economies: are they exaggerated by industrial aggregation? *Regional Science and Urban Economics* 28, pp. 199-211.

Murata, Y. (2003), Product diversity, taste heterogeneity, and geographic distribution of economic activities. *Journal of Urban Economics*, pp. 126-144.

Nelson, R.R. (1995a), Recent evolutionary theorizing about economic change. *Journal of Economic Literature* 33, pp. 48-90.

Nelson, R.R. (1995b), Co-evolution of industry structure, technology and supporting institutions, and the making of comparative advantage. *International Journal of the Economics of Business* 2(2), pp. 171-184.

Nieuwenhuijsen, H., and A. van Stel (2000), *Kennis-spillovers en Economische Groei.* Zoetermeer: EIM Research Report 0007/N.

Van Oort, F.G. (2004), *Urban Growth and Innovation. Spatially bounded externalities in the Netherlands.* Aldershot: Ashgate.

Van Oort, F.G., and O.A.L.C. Atzema (2004), On the conceptualisation of agglomeration economies: the case of new firm formation in the Dutch ICT-sector. *The Annals of Regional Science* 38(2), pp. 263-290.

Van der Panne, G. (2004), *Entrepreneurship and localized knowledge spillovers. Evidence from 398 innovators*. Dissertation. Delft: Technical University.

Parr, J.B. (2001), On the regional dimensions of Rostow's theory of growth. *Review of Urban and Regional Development Studies* 13, pp. 2-19.

Parr, J.B. (2002), Agglomeration economies: ambiguities and confusions. *Environment and Planning A* 34, pp. 717-731.

Pasinetti, L.L. (1981), *Structural Change and Economic Growth*. Cambridge: Cambridge University Press.

Pasinetti, L.L. (1993), *Structural Economic Dynamics*. Cambridge: Cambridge University Press.

Peneder, M. (2001), *Entrepreneurial Competition and Industrial Location*. Cheltenham: Edward Elgar.

Perroux, F. (1950). Economic space: theory and applications. *Quarterly Journal of Economics* 64(1), pp. 89-104.

Phelps, N.A., R.J. Fallon and C.L. Williams (2001), Small firms, borrowed size and the urban-rural shift. *Regional Studies* 35, pp. 613-624.

Quigly, J.M. (1998), Urban diversity and economic growth. *Journal of Economic Perspectives* 12, pp. 127-138.

Raspe, O., F. van Oort and P. de Bruijn (2004), *Kennis op de Kaart, Ruimtelijke patronen in de kenniseconomie*. Rotterdam: NAi Uitgevers.

Richardson, H.W. (1973), *The Economics of Urban Size.* Westmead: Saxon House.

Rigby, D.L., and J. Essletzbichler (2002), Agglomeration economies and productivity differences in US cities. *Journal of Economic Geography* 2, pp. 407-432.

Romer, P.M. (1986), Increasing returns and long-run growth, *Journal of Political Economy*, 94, pp. 1002-1037.

Romer, P.M. (1990), Endogenous technological change. *Journal of Political Economy*, 98, pp. 71-102.

Saviotti, P.P. (1996), *Technological Evolution, Variety and the Economy*. Cheltenham & Brookfield: Edward Elgar.

Saviotti, P.P. (2001), Variety, growth and demand. *Journal of Evolutionary Economics* 11 (1), pp. 119-142.

Saviotti, P.P., and A. Pyka (2004a), Economic development by the creation of new sectors. *Journal of Evolutionary Economics* 14(1), pp. 1-35.

Saviotti, P.P., and A. Pyka (2004b), Economic development, variety and employment creation. *Revue Economique* 55(6), pp. 1023-1059.

Shannon, C.E. (1948), A mathematical theory of communication. *Bell System Technical Journal* 27, pp. 379-423 and pp. 623-656.

Solow, R.M. (1957), Technical change and the aggregate production function. *Review of Economics and Statistics*, 39, pp. 312-320.

Storper, M., and R. Walker (1989), *The Capitalist Imperative. Territory, technology and industrial growth*. Oxford: Basil Blackwell.

Teece, D.J., G. Pisano and A. Shuen (1997), Dynamic capabilities and strategic management. *Strategic Management Journal* 18, pp. 509-534.

Theil, H. (1967), *Economics and Information Theory.* Amsterdam: North-Holland.

Theil, H. (1972), *Statistical Decomposition Analysis.* Amsterdam: North-Holland.

Thompson, W.R. (1968), Internal and external factors in the development of urban economics, in: H.S. Perloff and L. Wingo (eds.), *Issues in Urban Economics.* Baltimore: John Hopkins Press.

Vernon, R. (1966), International investment and international trade in the product lifecycle. *Quarterly Journal of Economics* 80, pp. 190-207.

Wasylenko, M.J., and R.A. Erickson (1978), On measuring economic diversification: a comment on the entropy-measure. *Land Economics* 54, pp. 106-109.

Weitzman, M.L. (1998), Recombinant growth. *Quarterly Journal of Economics* 113, pp. 331-360.

Zajdenweber, D. (1972), Une application de la théorie de l'information à l'économie: la mesure de la concentration. *Revue d'Economie Politique* 82, pp. 486-510.

	N	Minimum	Maximum	Mean	Std. Dev.
Employment growth CBS (%)	40	2.395	53.352	22.888	8.316
Productivity growth RUG (%)	40	-1.810	11.501	3.831	2.901
Unemployment growth CBS (%)	40	-.604	.022	-.346	.169
Inactivity growth CBS (%)	40	-.377	.152	-.108	.127
Unrelated variety LISA (two-digit)	40	1.249	1.409	1.341	.037
Related variety LISA (two- to five-digit)	40	.630	1.058	.933	.088
Los-index RUG	40	.240	.451	.340	.058
Log population density CBS	40	2.164	3.497	2.655	.357
Specialisation industry LISA	40	39.479	179.014	114.453	36.512
Specialisation distribution LISA	40	73.843	142.902	97.877	15.407
Specialisation producer services LISA	40	47.377	151.965	87.208	25.571
Specialisation consumer services LISA	40	76.385	124.530	100.275	11.601
Specialisation traditional manufacturing LISA	40	33.224	210.126	115.423	49.050
Wage CBS	40	28266.979	39840.510	32595.878	2517.512
Average investment per fte CBS	40	9058.064	27912.600	13136.561	2924.819
Capital-labour ratio growth RUG (%)	40	-12.927	19.009	2.987	6.188
R&D level per fte Senter	40	229.802	1969.259	641.889	353.011
Business areas growth RPB (index)	40	6.170	688.280	114.747	119.609
Dwellings growth CBS (%)	40	-6.061	19.513	-.035	3.668
Competition level CBS	40	72.793	165.027	110.020	21.514
Human capital RUG	40	.000	1.210	.179	.272
Randstad dummy	40	.000	1.000	.325	
Intermediate zone dummy	40	.000	1.000	.225	
National periphery dummy	40	.000	1.000	.450	

Table 1. Descriptive statistics of dependent and independent variables

	empl. growth	prod. growth	unempl. growth	inact. growth	un-related variety	related variety	Los-index	log pop. density	spec ind.	spec distr.	spec prod. services	spec cons. services	spec trad. manuf.	wage	invest.	growth c-l ratio	R&D	bus. areas growth	dwel. growth	comp.	human capital	Rand-stad	inter-mediate zone	national peri-phery
employment growth	1																							
productivity growth	-.065	1																						
unemployment growth	.228	.460**	1																					
inactivity growth	.273	.340*	.814**	1																				
unrelated variety	.101	.009	-.223	-.194	1																			
related variety	.483**	-.178	.035	.118	-.046	1																		
Los-index	-.401*	-.144	-.020	-.108	-.206	-.487**	1																	
log population density	-.100	-.044	-.014	.064	-.336*	.405**	-.034	1																
spec industry	-.320*	.091	-.047	.022	.570**	-.610**	.254	-.527**	1															
spec distribution	.221	-.051	.020	-.052	.215	.323*	.091	-.022	-.101	1														
spec producer services	.246	-.118	-.048	-.014	-.307	.542**	-.190	.662**	-.763**	.073	1													
spec consumer services	.029	.021	.080	.021	-.581**	.067	-.214	.116	-.508**	-.639**	.061	1												
spec trad. manufacturing	.019	-.081	-.191	-.029	.622**	-.234	-.138	-.597**	.764**	-.049	-.643**	-.361*	1											
wage	.628**	.164	.178	.228	-.026	.472**	-.373*	.429**	-.504**	.236	.627**	-.065	-.385*	1										
investment	.328*	.224	.010	-.068	.053	-.215	.017	-.434**	.039	.200	-.012	-.178	.012	.179	1									
growth capital-labour ratio	-.016	.588**	.240	.023	-.122	.278	-.291	.221	-.351*	.104	.255	.141	-.272	.185	-.043	1								
R&D	.034	.247	.204	.328*	.220	-.066	-.038	.191	.275	-.226	-.080	-.104	.020	.092	-.184	-.065	1							
business areas growth	.166	-.147	-.170	-.199	.262	.261	-.053	.125	-.035	-.020	.246	-.139	-.005	.152	-.185	-.100	.185	1						
dwellings growth	.810**	-.146	.144	.207	.159	.334*	-.263	-.160	-.151	.108	.154	-.019	.092	.346*	.366*	-.236	-.001	.310	1					
Competition	.229	-.023	.050	.037	.177	-.178	.084	-.637**	.269	.312*	-.481**	-.157	.434**	-.291	.334*	-.280	-.139	-.025	.268	1				
human capital	.133	.020	-.057	-.153	.040	.400*	-.324*	.427**	-.464**	.016	.555**	.099	-.392*	.511**	-.092	.250	.079	.305	.000	-.552**	1			
randstad dummy	-.003	-.024	.076	.120	-.395*	.338*	.009	.738**	-.533**	.212	.682**	-.057	-.634**	.435**	-.043	.230	.029	.020	-.011	-.401*	.272	1		
intermediate zone dummy	.247	-.127	.072	.070	.216	.345*	-.055	.007	.004	.060	.059	-.094	.117	.090	-.140	-.111	.120	.268	.292	-.023	.033	-.374*	1	
periphery dummy	-.205	.130	-.132	-.172	.191	-.608**	.038	-.701**	.499**	-.250	-.692**	.132	.499**	-.485**	.158	-.124	-.128	-.244	-.235	.397*	-.283	-.628**	-.487**	1

** Correlation is significant at the 0.01 level (two-tailed) * Correlation is significant at the 0.05 level (two-tailed).

Table 2. Correlation matrix

	(1) OLS	(1a) OLS 1996-2001	(1b) OLS 1997-2002	(2) OLS	(3) OLS	(4) OLS		(5) OLS			(6) OLS (WA-variables)
						RS	Not-RS	RS	IZ	Per	
CONSTANT	0.081 (0.581)	0.074 (0.530)	0.072 (0.495)	0.106 (0.9451)	0.018 (0.177)	-0.685 (0.832)	0.232 (0.975)	-0.685 (-0.822)	-0.098 (-0.047)	0.027 (0.053)	0.094 (0.609)
UNRELATED VARIETY	-0.040 (-0.248)	-0.030 (-0.189)	-0.112 (-0.666)	-0.099 (-0.765)	-0.088 (-0.751)	0.022 (0.069)	-0.220 (-1.014)	0.022 (0.068)	-0.854 (-1.399)	-0.630* (-1.744)	0.071 (0.345)
RELATED VARIETY	0.627** (3.836)	0.532** (3.297)	0.569** (3.363)	0.460** (3.394)	0.310** (2.364)	0.689 (1.324)	0.571** (2.686)	0.689 (1.308)	1.474 (0.523)	-0.094 (-0.219)	0.520* (2.620)
LOS-INDEX	-0.131 (-0.779)	-0.249 (-1.498)	-0.152 (-0.870)	0.027 (0.191)	-0.122 (-1.004)	0.080 (0.196)	-0.174 (-0.884)	0.080 (0.194)	0.046 (0.053)	-0.514 (-1.545)	-0.487** (-2.804)
POPULATION DENSITY	-0.263 (-1.400)	-0.207 (-1.120)	-0.240 (-1.236)	-0.654** (-3.772)	-0.130 (-0.944)	0.366 (0.617)	-0.211 (-0.584)	0.366 (0.610)	-1.485 (-0.660)	-0.134 (-0.244)	-0.085 (-1.870)
INVESTMENT	0.361** (2.903)	0.291** (2.371)	0.328** (2.550)	0.087 (0.744)	0.130 (1.312)	0.794** (3.087)	0.005 (0.981)	0.794** (3.038)	0.162 (0.115)	-0.626 (-1.526)	-0.044 (-0.568)
R&D	0.194 (1.443)	0.147 (1.111)	0.214 (1.538)	0.151 (1.404)	0.115 (1.171)	0.319 (0.874)	0.159 (1.030)	0.319 (0.863)	0.512 (0.832)	0.047 (0.201)	-0.011 (-0.142)
WAGE				0.726** (4.469)							
DWELLINGS GROWTH					0.548** (5.587)						
R2	0.502	0.480	0.446	0.694	0.748	0.624		0.718			0.446
ADJ. R2	0.412	0.385	0.346	0.627	0.693	0.436		0.421			0.345
MAX. LIKELIHOOD	-47.698			-37.999	-34.081	-42.093		-36.323			-49.839
LAGRANGE MULTIPLIER (ERROR)	1.871 (0.171)			2.014 (0.156)	0.208 (0.648)	0.431 (0.512)		0.131 (0.718)			
LAGRANGE MULTIPLIER (LAG)	1.597 (0.206)			1.669 (0.196)	0.091 (0.763)	2.687 (0.101)		3.196 (0.074)			
CHOW-WALD TEST						1.201 (0.337)		1.040 (0.459)			

t-values in parentheses (except for Lagrange multipliers test statistics, where p-values are shown)
RS stands for Randstad area ; IZ for intermediate zone ; Per for national periphery ; WA for window-average variables (Anselin 1988).
** Significant at the 0.05-level * Significant at the 0.10-level

Table 3. Dependent variable: *EMPLOYMENT GROWTH*

	(1) OLS	(1a) OLS 1996-2000	(1b) OLS 1997-2001	(2) Spatial lag	(3) Spatial lag		(4) Spatial lag			(5) OLS (WA-variables)
					RS	Not-RS	RS	IZ	Per	
CONSTANT	-0.075 (-0.678)	-0.072 (-0.641)	-0.071 (-0.583)	-0.106 (-1.136)	-0.060 (-0.113)	-0.014 (-0.096)	-0.066 (-0.147)	0.837 (0.438)	0.183 (0.682)	0.038 (0.254)
UNRELATED VARIETY	-0.044 (-0.342)	0.004 (0.035)	0.038 (0.272)	0.007 (0.070)	-0.487** (-2.521)	0.144 (1.077)	-0.493** (-2.988)	0.005 (0.013)	0.033 (0.173)	-0.045 (-0.223)
RELATED VARIETY	-0.300** (-2.335)	-0.290** (-2.228)	-0.133 (-0.937)	-0.288** (-2.685)	-0.224 (-0.481)	-0.329** (-2.539)	-0.231 (-0.582)	-1.293 (-0.527)	-0.435** (-1.995)	-0.317 (-1.661)
LOS-INDEX	-0.112 (-0.822)	0.015 (0.108)	0.042 (0.277)	-0.117 (-1.026)	0.152 (0.619)	-0.144 (-1.183)	0.138 (0.657)	-0.563 (-0.831)	-0.136 (-0.798)	0.061 (0.335)
POPULATION DENSITY	-0.040 (-0.272)	-0.094 (-0.624)	-0.035 (-0.212)	-0.070 (-0.561)	-0.522 (-1.394)	0.224 (0.947)	-0.534* (-1.670)	-0.236 (-0.102)	0.571* (1.709)	0.007 (0.149)
INVESTMENT	0.178* (1.825)	0.132 (1.338)	0.273** (2.529)	0.193** (2.353)	0.198 (1.284)	0.333** (2.615)	0.211 (1.593)	0.882 (0.765)	0.278 (1.365)	0.088 (1.017)
R&D	0.282** (2.665)	0.318** (2.970)	0.267** (2.286)	0.281** (3.173)	0.206 (0.678)	0.202** (2.177)	0.203 (0.782)	-0.160 (-0.319)	0.223* (1.720)	0.136* (1.865)
CAPITAL-LABOUR RATIO GROWTH	0.684** (5.704)	0.692** (5.696)	0.629** (4.740)	0.733** (7.311)	0.921** (3.654)	0.643** (4.489)	0.942** (4.349)	0.572 (0.607)	0.545** (3.686)	0.188** (2.335)
W_PRODUCTIVITY GROWTH				-0.379** (-2.325)	-0.428** (-2.585)		-0.516** (-3.137)			
R2	0.611	0.601	0.524	0.639	0.714		0.788			0.350
ADJ. R2	0.526	0.514	0.420	0.661	0.740		0.811			0.208
MAX. LIKELIHOOD	-37.361			-35.262	-30.154		-24.189			-47.631
LAGRANGE MULTIPLIER (ERROR)	1.197 (0.274)									
LAGRANGE MULTIPLIER (LAG)	3.331 (0.068)									
LIKELIHOOD RATIO TEST				4.197 (0.040)	5.127 (0.024)		6.191 (0.013)			
CHOW-WALD TEST					11.710 (0.165)		30.633 (0.015)			

t-values in parentheses (except for Lagrange multiplier test statistics, where p-values are shown)
** Significant at the 0.05-level * Significant at the 0.10-level

Table 4. Dependent variable: *PRODUCTIVITY GROWTH*

	(1) OLS	(1a) OLS 1996-2001	(1b) OLS 1997-2002	(1c) OLS (incl. disabled)	(2) OLS		(3) OLS			(4) OLS (WA-variables)
					RS	Not-RS	RS	IZ	Per	
CONSTANT	-0.008 (-0.056)	-0.050 (-0.344)	-0.003 (-0.021)	-0.010 (-0.064)	1.193 (1.245)	0.272 (1.196)	1.193 (1.112)	0.501 (0.112)	-0.217 (-0.376)	-0.337* (-2.025)
UNRELATED VARIETY	-0.392** (-2.241)	-0.101 (-0.602)	-0.383** (-2.234)	-0.414** (-2.375)	-0.685** (-2.168)	-0.293 (-1.408)	-0.685* (-1.937)	-0.435 (-0.528)	-0.070 (-0.144)	-0.145 (-0.733)
RELATED VARIETY	0.012 (0.064)	-0.387** (-2.211)	-0.033 (-0.183)	0.077 (0.423)	-0.066 (-0.083)	-0.395* (-1.848)	-0.066 (-0.074)	-0.367 (-0.067)	-0.350 (-0.660)	-0.419* (-1.910)
LOS-INDEX	0.130 (0.673)	-0.155 (-0.840)	0.123 (0.649)	0.018 (-0.092)	-0.299 (-0.732)	0.495** (2.192)	-0.299 (-0.654)	0.387 (0.218)	0.321 (0.661)	0.399** (2.266)
POPULATION DENSITY	-0.521** (-2.209)	-0.416* (-1.837)	-0.588** (-2.542)	-0.427* (-1.814)	-1.684** (-2.126)	0.012 (0.032)	-1.684* (-1.899)	-0.627 (-0.117)	-0.824 (-1.012)	-0.224** (-3.441)
INVESTMENT	-0.148 (-0.940)	-0.075 (-0.495)	-0.200 (-1.300)	-0.153 (-0.978)	-0.522 (-1.420)	-0.368 (-1.626)	-0.522 (-1.268)	-0.130 (-0.045)	-0.280 (-0.642)	-0.241** (-2.531)
R&D	0.299** (2.062)	0.275* (1.980)	0.271* (1.910)	0.380** (2.629)	0.507 (1.039)	0.019 (0.124)	0.507 (0.928)	-0.129 (-0.122)	0.370 (1.370)	-0.002 (-0.024)
WAGE	0.383* (1.751)	0.270 (1.286)	0.541** (2.523)	0.348 (1.598)	0.389 (1.106)	0.977** (3.505)	0.389 (0.988)	0.641 (0.675)	1.278** (2.290)	0.409** (3.949)
CAPITAL-LABOUR RATIO GROWTH	0.286* (1.743)	0.155 (0.989)	0.227 (1.413)	-0.007 (-0.041)	0.138 (0.341)	0.141 (0.605)	0.138 (0.305)	0.038 (0.015)	0.364 (0.951)	0.113 (1.424)
R2	0.298	0.354	0.324	0.301	0.636		0.731			0.420
ADJ. R2	0.117	0.187	0.150	0.121	0.355		0.192			0.270
MAX. LIKELIHOOD	-49.179				-36.023		-30.019			-45.368
LAGRANGE MULTIPLIER (ERROR)	0.741 (0.389)				0.294 (0.588)		0.000 (0.992)			
LAGRANGE MULTIPLIER (LAG)	0.998 (0.318)				1.390 (0.238)		0.611 (0.434)			
CHOW-WALD TEST					2.275 (0.056)		1.160 (0.399)			

t-values in parentheses (except for Lagrange multiplier test statistics, where p-values are shown)
** Significant at the 0.05-level * Significant at the 0.10-level

Table 5. Dependent variable: *UNEMPLOYMENT GROWTH*

COROP	COROP name	Unrelated variety	Related variety	Los-index	Pop. density	Factor 1	Factor 2
1	Oost-Groningen	1.325	0.897	0.355	2.258	-1.368	-0.380
2	Delfzijl e.o.	1.333	**0.630**	**0.451**	2.305	**-1.662**	-0.568
3	Overig Groningen	1.312	0.957	0.289	2.453	0.144	-1.010
4	Noord-Friesland	1.358	0.913	0.277	2.292	-0.250	-0.114
5	Zuidwest-Friesland	1.348	0.851	0.336	**2.207**	-1.091	0.137
6	Zuidoost-Friesland	1.326	0.915	0.334	2.238	-0.637	0.008
7	Noord-Drenthe	**1.268**	0.906	0.412	2.262	-0.478	**-1.965**
8	Zuidoost-Drenthe	1.364	**0.808**	0.330	2.283	-1.488	0.280
9	Zuidwest-Drenthe	1.368	0.858	0.421	**2.167**	-0.823	0.503
10	Noord-Overijssel	1.370	0.889	0.272	2.344	-0.475	0.419
11	Zuidwest-Overijssel	1.328	0.882	**0.447**	2.522	-0.804	-0.643
12	Twente	1.356	0.977	0.319	2.615	-0.466	0.373
13	Veluwe	1.352	0.983	**0.240**	2.526	0.379	0.065
14	Achterhoek	1.364	0.965	0.330	2.387	-0.868	0.866
15	Arnhem/Nijmegen	1.330	0.993	0.304	2.857	0.541	-0.868
16	Betuwe	1.350	1.008	0.388	2.483	0.379	**2.121**
17	Utrecht	1.321	**1.058**	0.286	2.897	**1.722**	0.065
18	Kop Noord-Holland	1.298	0.915	0.295	2.504	0.354	-0.900
19	Alkmaar e.o.	1.302	1.011	0.408	2.888	0.879	-1.026
20	IJmond	**1.285**	0.818	**0.449**	3.078	-0.576	-0.969
21	Haarlem e.o.	1.289	1.005	0.369	**3.201**	0.911	**-1.790**
22	Zaanstreek	1.333	0.948	0.358	3.118	0.588	0.673
23	Groot-Amsterdam	1.347	1.024	0.375	**3.199**	**2.083**	0.550
24	Gooi	1.304	0.982	0.317	3.106	1.089	-0.710
25	Leiden e.o.	**1.286**	0.993	0.266	3.196	0.970	-1.060
26	Aggl. 's-Gravenhage	1.249	0.927	0.334	**3.497**	1.411	**-2.527**
27	Delft e.o.	1.349	0.830	0.438	3.080	0.365	-0.349
28	Groene Hart	1.313	**1.037**	0.367	2.802	1.104	0.801
29	Rijnmond	1.365	**1.033**	0.264	3.041	**1.760**	**1.476**
30	Zuidoost-Zuid-Holland	1.367	0.992	0.349	2.889	0.822	**2.026**
31	Zeeuwsch-Vlaanderen	1.331	**0.786**	0.368	**2.164**	-1.378	-0.399
32	Overig Zeeland	1.371	0.937	0.358	2.391	-0.347	0.081
33	West-Brabant	1.387	1.017	0.358	2.658	0.074	1.074
34	Midden-Brabant	1.385	0.994	0.289	2.658	-0.087	0.629
35	Noordoost-Brabant	1.366	1.011	0.281	2.643	0.158	1.081
36	Zuidoost-Brabant	**1.401**	0.982	0.293	2.701	0.182	1.184
37	Noord-Limburg	1.386	0.847	0.383	2.505	**-1.543**	0.179
38	Midden-Limburg	**1.409**	0.828	0.373	2.525	**-1.664**	0.585
39	Zuid-Limburg	**1.390**	0.887	**0.263**	2.979	-0.642	-0.127
40	Flevoland	1.353	1.017	**0.258**	2.281	0.734	0.226

N.B. Bold and dark grey: top three scoring regions. Bold and light grey: the worst scoring regions.

Table 6. Scores per COROP on main independent variables and factors

Appendix I: The entropy decomposition theorem

One of the most powerful and attractive properties of entropy statistics is the way in which problems of aggregation and disaggregation are handled (Theil 1972: 20-22; Zajdenweber 1972; Frenken 2005). This is due to the property of additivity of the entropy formula.

Let E_i stand again for an event, and let there be n events E_1, ..., E_n with probabilities $p_1, ..., p_n$. Assume that all events can be aggregated into a smaller number of sets of events S_1, ..., S_G in such a way that each event exclusively falls under one set S_g, where $g=1,...,G$. The probability that event falling under S_g occurs is obtained by summation:

$$P_g = \sum_{i \in S_g} p_i \qquad \text{(A.1)}$$

The entropy at the level of sets of events is:

$$H_0 = \sum_{g=1}^{G} P_g \log_2 \left(\frac{1}{P_g} \right) \qquad \text{(A.2)}$$

H_0 is called the between-group entropy. The entropy decomposition theorem specifies the relationship between the between-group entropy H_0 at the level of sets and the entropy H at the level of events as defined in (2). Write entropy H as:

$$H = \sum_{i=1}^{n} p_i \log_2 \left(\frac{1}{p_i} \right) = \sum_{g=1}^{G} \sum_{i \in S_g} p_i \log_2 \left(\frac{1}{p_i} \right)$$

$$= \sum_{g=1}^{G} P_g \sum_{i \in S_g} \frac{p_i}{P_g} \left(\log_2 \left(\frac{1}{P_g} \right) + \log_2 \left(\frac{P_g}{p_i} \right) \right)$$

$$= \sum_{g=1}^{G} P_g \left(\sum_{i \in S_g} \frac{p_i}{P_g} \right) \log_2 \left(\frac{1}{P_g} \right) + \sum_{g=1}^{G} P_g \left(\sum_{i \in S_g} \frac{p_i}{P_g} \log_2 \left(\frac{P_g}{p_i} \right) \right)$$

$$= \sum_{g=1}^{G} P_g \log_2 \left(\frac{1}{P_g} \right) + \sum_{g=1}^{G} P_g \left(\sum_{i \in S_g} \frac{p_i}{P_g} \log_2 \left(\frac{1}{p_i / P_g} \right) \right)$$

The first right-hand term in the last line is H_0. Hence:

$$H = H_0 + \sum_{g=1}^{G} P_g H_g \qquad \text{(A.3)}$$

where:

$$H_g = \sum_{i \in S_g} \frac{p_i}{P_g} \log_2 \left(\frac{1}{p_i / P_g} \right) \quad g = 1,...,G \qquad \text{(A.4)}$$

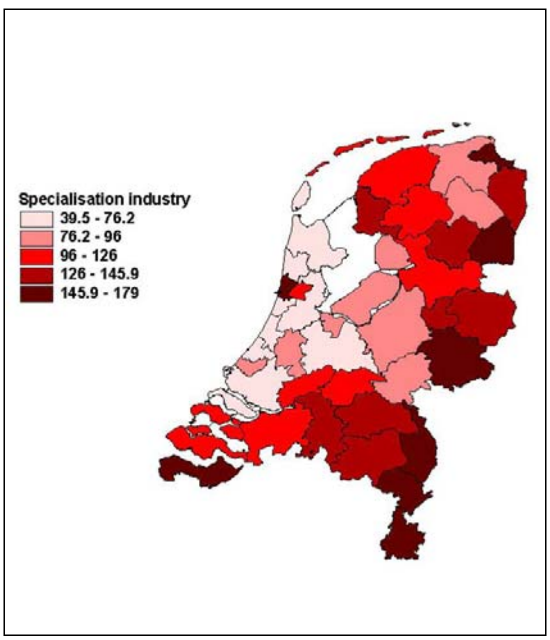
The probability p_i/P_g, $i \in S_g$ is the conditional probability of E_i given knowledge that one of the events falling under S_g is bound to occur. H_g thus stands for the entropy within the set S_g and the term $\sum P_g H_g$ in (9) is the *average within-group entropy*. Entropy thus equals the between-group entropy plus the average within-group entropy. Two properties of this relationship follow (Theil 1972: 22):

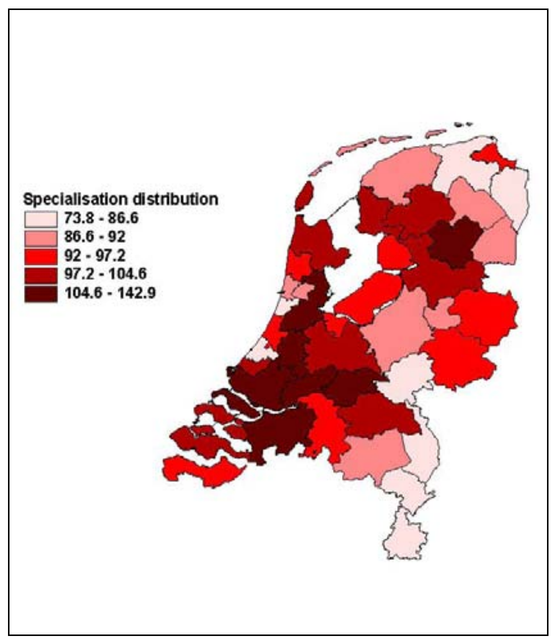
(i) $H \geq H_0$ because both P_g and H_g are nonnegative. It means that after grouping there cannot be more entropy (uncertainty) than there was before grouping.

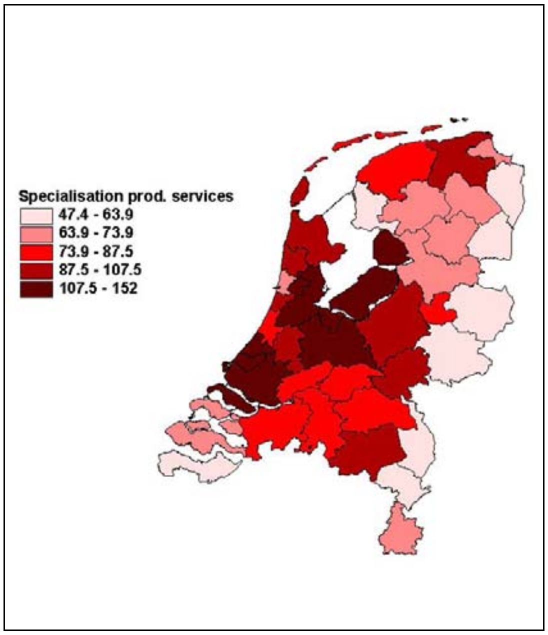
(ii) $H = H_0$ if and only if the term $\sum P_g H_g = 0$ and $\sum P_g H_g = 0$ if and only if $H_g = 0$ for each set S_g. It means that entropy equals between-group entropy if and only if the grouping is such that there is at most one event with nonzero probability.

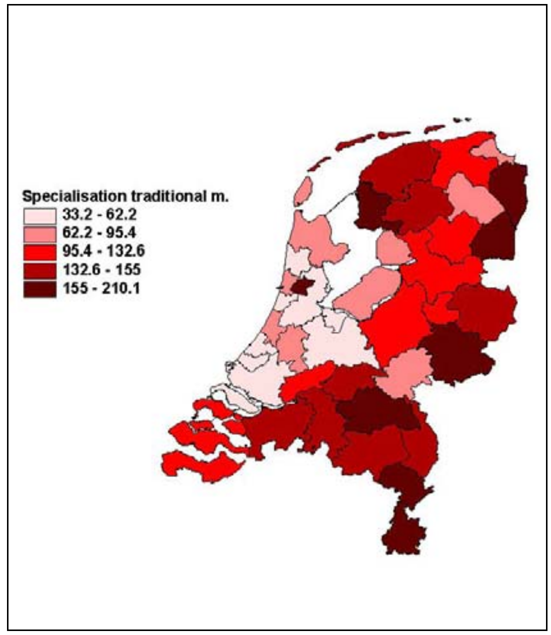
In informational terms, the decomposition theorem has the following interpretation. Consider the first message that one of the sets of events occurred. Its expected information content is H_0. Consider the subsequent message that one of the events falling under this set occurred. Its expected information content is H_g. The total information content becomes $H_0 + \sum P_g H_g$. Applications of the decomposition theorem will be discussed in the third and fourth section.

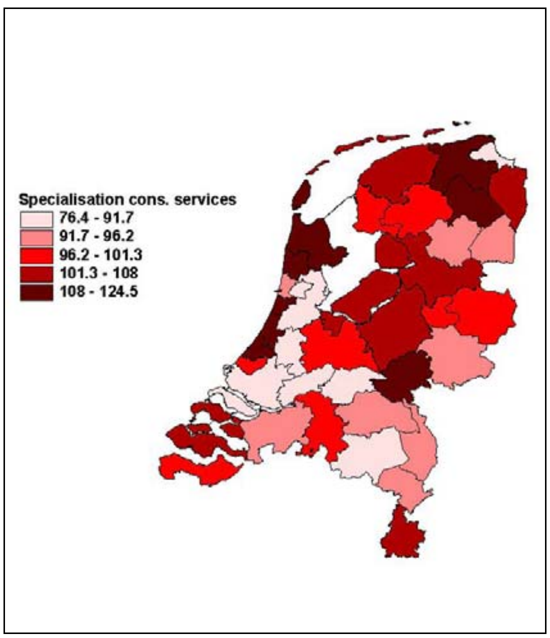
Appendix II: Maps

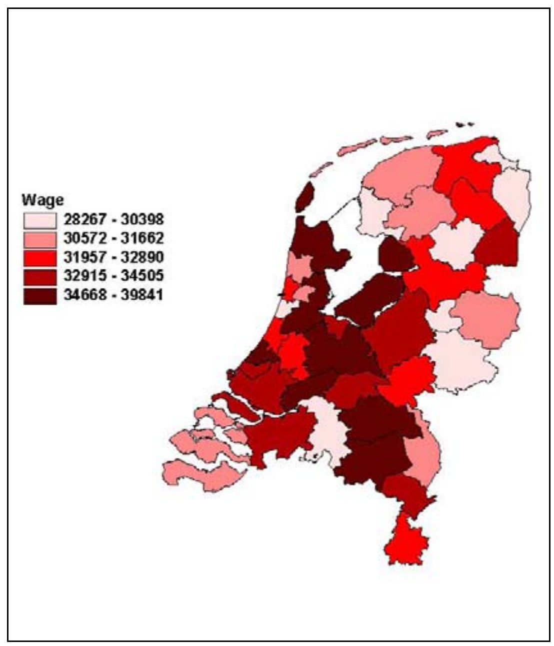


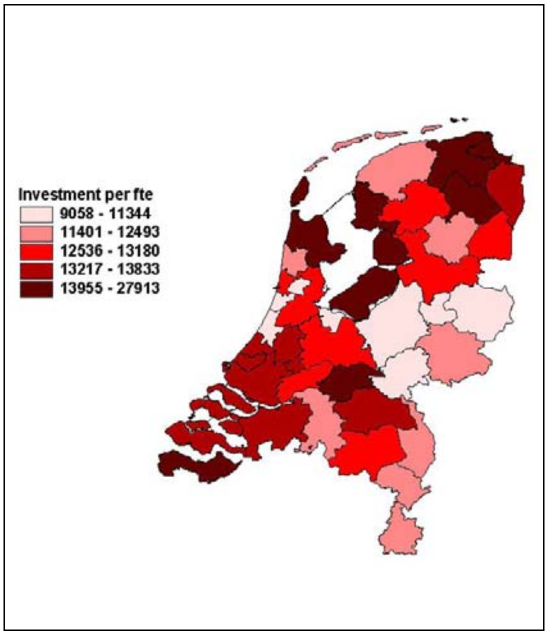


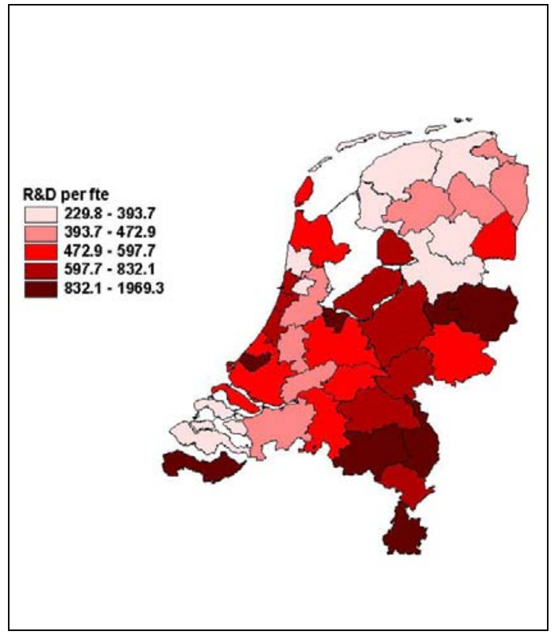

Specialisation prod. services
47.4 - 63.9
63.9 - 73.9
73.9 - 87.5
87.5 - 107.5
107.5 - 152

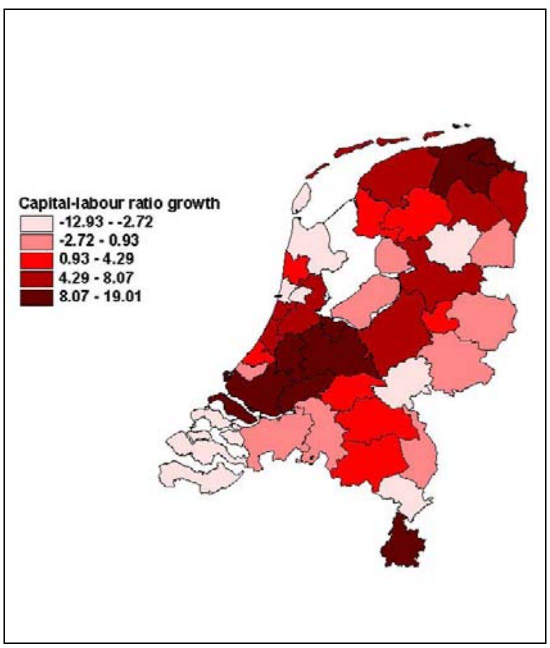

Specialisation traditional m.
33.2 - 62.2
62.2 - 95.4
95.4 - 132.6
132.6 - 155
155 - 210.1

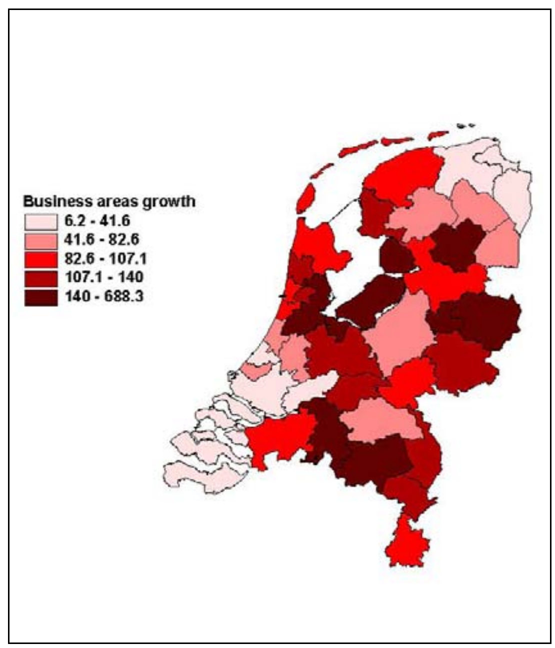

Specialisation cons. services
76.4 - 91.7
91.7 - 96.2
96.2 - 101.3
101.3 - 108
108 - 124.5

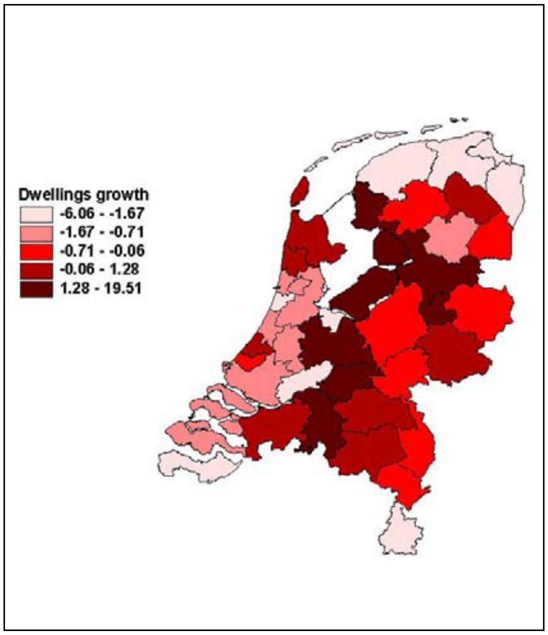

Wage
28267 - 30398
30572 - 31662
31957 - 32890
32915 - 34505
34668 - 39841

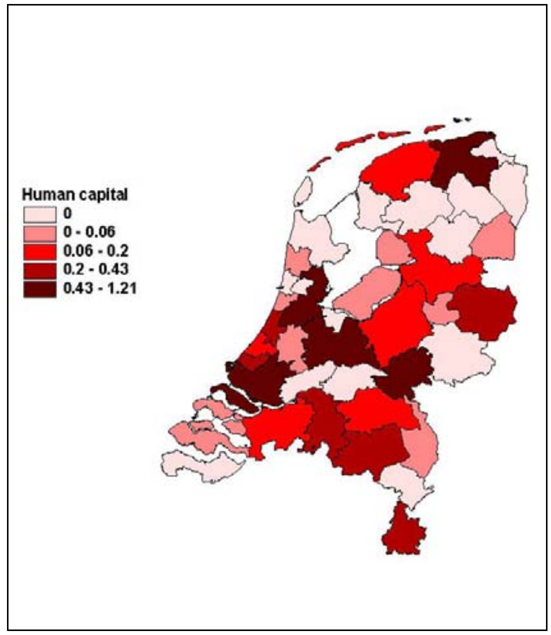

Investment per fte
9058 - 11344
11401 - 12493
12536 - 13180
13217 - 13833
13955 - 27913

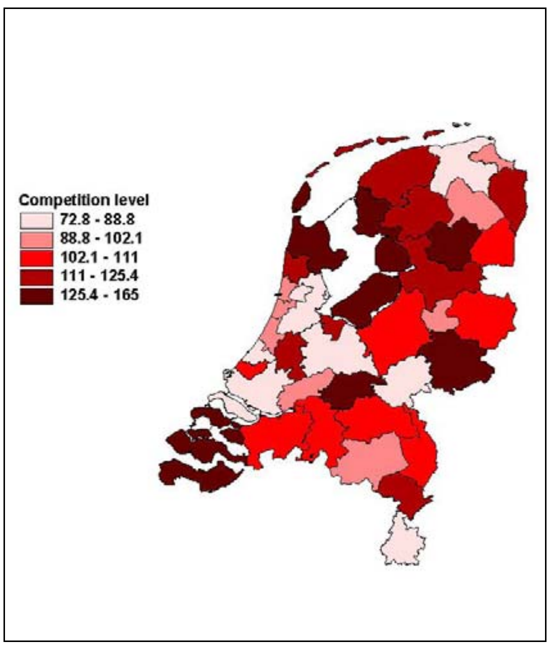

R&D per fte
229.8 - 393.7
393.7 - 472.9
472.9 - 597.7
597.7 - 832.1
832.1 - 1969.3

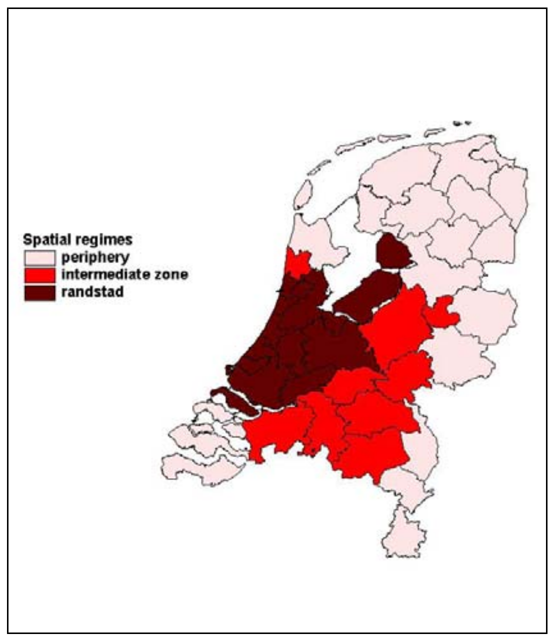

Capital-labour ratio growth
-12.93 - -2.72
-2.72 - 0.93
0.93 - 4.29
4.29 - 8.07
8.07 - 19.01

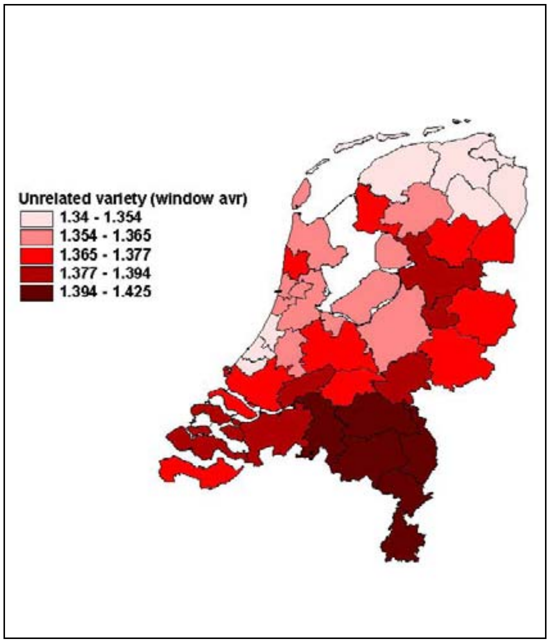

Business areas growth
6.2 - 41.6
41.6 - 82.6
82.6 - 107.1
107.1 - 140
140 - 688.3

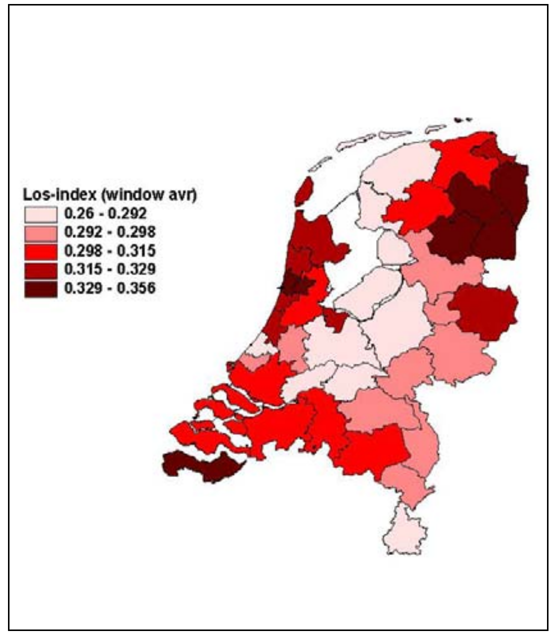

Dwellings growth
-6.06 - -1.67
-1.67 - -0.71
-0.71 - -0.06
-0.06 - 1.28
1.28 - 19.51

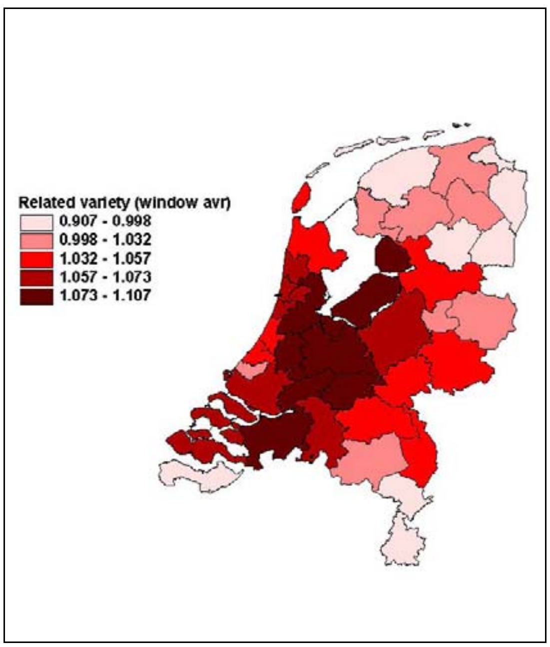

Human capital
0
0 - 0.06
0.06 - 0.2
0.2 - 0.43
0.43 - 1.21


Competition level
72.8 - 88.8
88.8 - 102.1
102.1 - 111
111 - 125.4
125.4 - 165


Spatial regimes
periphery
intermediate zone
randstad


Unrelated variety (window avr)
1.34 - 1.354
1.354 - 1.365
1.365 - 1.377
1.377 - 1.394
1.394 - 1.425


Los-index (window avr)
0.26 - 0.292
0.292 - 0.298
0.298 - 0.315
0.315 - 0.329
0.329 - 0.356


Related variety (window avr)
0.907 - 0.998
0.998 - 1.032
1.032 - 1.057
1.057 - 1.073
1.073 - 1.107