Online working papers 15/2010

Robert Hagfors and Jouko Kajanoja

Welfare states and social sustainability

An application of SEM and SOM in a virtuous circle environment

The social Insurance Institution of Finland (Kela), Research Department

Authors

Robert Hagfors, D.Soc.Sc, docent, Chief Economist
The Social Insurance Institution, Research Department, Finland
firstname.lastname@kela.fi

Jouko Kajanoja, D.Soc.Sc, docent
firstname.lastname@emmapalvelut.fi

Presented at the 8th Annual ESPAnet Conference, 2–4 September 2010, Budapest.

Helsinki 2010

Contents

1 Introduction

The welfare state is the focus of a lively public debate. In part, this debate is motivated by concerns over the sustainability of public finances given the expected demographic development and the openness of national economies. How, then, might the welfare state and the social policy implemented prove themselves sustainable under the financial pressures?

One of the current topics of debate is whether civic activity could make up for such cuts in social expenditure as may become necessary as a result of tax decreases and the targeting of public expenditure to investments aimed at boosting productivity. Are there any signs of civic activity that would compensate for retrenchments of public social policy? How does social policy influence social capital, and how does social capital, in turn, influence the sustainability of development? In this study, we will first construct the interrelationships referred to above, and then test them by means of empirical comparisons.

A central concept of the first part of this study is the virtuous circle. We use it here to refer to the propensity of actions with positive effects to function in a self-reinforcing manner. In the welfare state context, this means that to the extent that a welfare effort has a positive effect on well-being and the reduction of inequality, the citizens are willing to maintain and even to intensify such an effort.

Our theory was inspired by a research result suggesting that trust between citizens and public-spiritedness are strongest in the Nordic countries and weakest in countries where the level of social security is lowest. We study in a comparative perspective the hypothesis that the risks should be covered by the welfare state. This hypothesis argues that the welfare state increases equality between people by covering risks and by equalising opportunities and the income distribution. The core of our hypothesis is that the equality created by the welfare state is positively interrelated with generalised trust between the people and with bridging social capital. This interrelation diminishes risks and it is interrelated with the well-being of the people. We combine variables in a way that produces a continuous process chain. For this purpose we use several indicators from developed OECD countries and latent variable structural equation modelling (SEM).

The aim of the second part of this study is to assess the appropriateness of a new method for comparative welfare research. We approach here the classification and clustering of welfare states by means of a novel method consisting of an algorithm for large high-dimensional data sets. This method, known as the Self-Organizing Map (SOM), is in the literature also referred to as the Kohonen map after its developer. (Kohonen 1982, 1990 and 2001; Kaski and Kohonen 1995; Oja and Kaski 1999.) It was developed for purposes of data analysis and information visualization. This neural network algorithm can cluster cases characterized by multidimensional data on a two-dimensional output space, where the central dependencies of the data are represented on a two-dimensional grid or map. The learning is carried out by iterative regressions. The neurons of the map (node) represent a model. The models become ordered so that similar models move closer to each other while dissimilar models move away from each other as the iterations proceed. The Self-Organizing Map is applied to test the sustainability of the welfare state models using the indicators revealed by the empirical estimation of the virtuous circle. In a way we are approaching the dependent variable problem by replacing the unilateral variable with a multilateral set of indicators. In the clustering process (SOM-Ward-clusters) it is found that the Nordic welfare state model remains coherent while the other welfare state models are dispersed. This emphasizes the relation between the idea of the virtuous circle, neural networks and the social sustainability of the welfare state.

2 The Virtuous Circle

> ”Quite obviously a circular relationship between less poverty, more food, improved health and higher working capacity would sustain a cumulative process upward instead of downward.” (Gunnar Myrdal, 1957, 12.)

A central concept of the first part of this study is the virtuous circle. We use it here to refer to the propensity of actions with positive effects to function in a self-reinforcing manner. In the welfare state context, this means that to the extent that a welfare effort has a positive effect on well-being and the reduction of inequality, the citizens are willing to maintain and even to intensify such an effort.

Though not always referred to as the virtuous circle, this idea has antecedents in the literature. Studying third-world poverty, Gunnar Myrdal (1957) laid out a hypothesis of circular causation (Chapter II, The principle of circular and cumulative causation). He describes the phenomenon of the vicious circle, where poverty produces morbidity, which engenders poverty. While not an original idea as such, Myrdal discovered that the direction of causality could be reversed. This upward spiral is well represented in the dictum describing the American dream, according to which “nothing succeeds like success”. Yet the spiral can also lead downward: “Nothing fails like failure.” (Myrdal, 1957, 12.)

Olli Kangas and Joakim Palme (2009) see the virtuous circle as being underpinned by certain institutional characteristics. According to them the uniform state structure in the Nordic countries, especially in Sweden, laid the groundwork for the flexible co-operation of the central and local governments, where the production of public services was based on decision-making at the local level. This possibility for local decision-making legitimized the public sector. In addition, the possibility to participate in the local decision-making process strengthened the feeling of coherence among citizens. This is the way in which the virtuous circle between democracy and social policy was shaped in the Nordic countries.

Manuel Castells and Pekka Himanen (2002) represent another kind of approach. They use Finland as an example of the Information Society with a virtuous circle of the information economy and welfare state at its core. The financial basis for the welfare state is created within sectors of the information economy. Public-sector education, health and social services guarantee a well educated and healthy workforce, and because of this connection support for the welfare state is strong. While it was initially the case that the welfare state relied on the development of an information society, the globalization process has now led to a situation where it is social policy that creates the necessary preconditions for the development of the information society

Circular causation is also found in Rothstein and Uslaner (2005, 44), but in their construct it serves to create equality and general well-being. In the welfare state, an equal distribution of resources and opportunities leads to increasing general well-being, which contributes to the practice of a universalistic social policy, which strengthens generalized trust. This chain of causation contains feedback effects. Rothstein and Uslaner place the decrease in equality at the beginning of the chain of causation.

Taking the Nordic countries as an example, Bo Rothstein identified social capital as the central starting point to the circulation process (Rothstein, 2008). There existed, at the beginning, only a small amount of social capital, maybe just a little more than in other countries. That was, however, enough to start the feedback process. From the starting point of a small amount of social capital a universal social policy emerged, which then increased social capital so that in time it was possible to deepen and widen the existing social policy systems, which for their part strengthened social trust and so on. On the other hand, Rothstein and Uslaner considered it quite possible that the virtuous circle was set in motion by a universal social insurance system constructed at the upper levels of government.

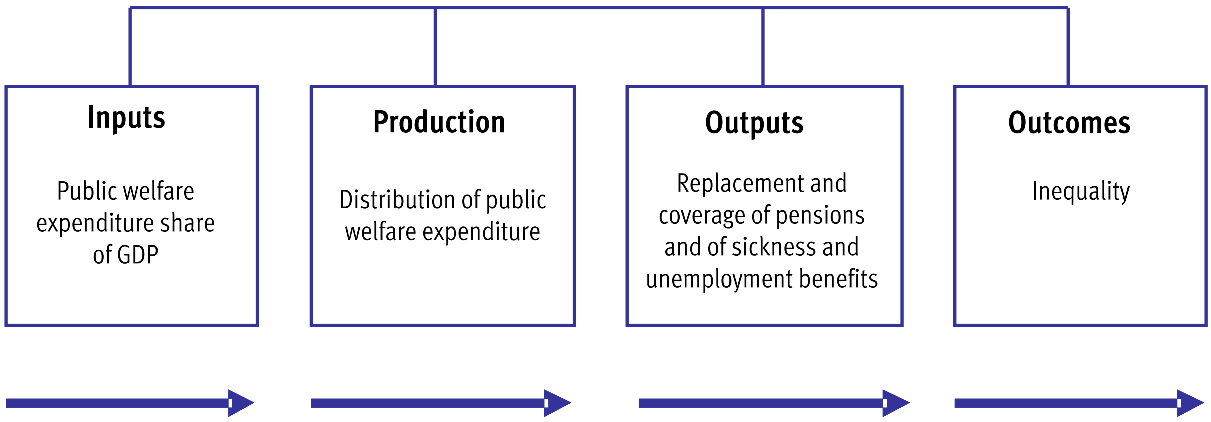
Wim van Oorschot and Ellen Finsveen of the University of Tilburg have constructed a circle where the factors affecting each other comprise both the differences in economic and cultural backgrounds and the differences in social capital. They assume that the welfare state affects economic and cultural characteristics separately by levelling off differences in each. The researchers investigated causal relations by analysing time series data. According to them the results concerning causal relations remained uncertain partly because of the narrowness of the data. (van Oorschot and Finsveen 2008.)

According to John Dowling and Yap Chin-Fang the principle of the virtuous circle has also been discovered in economic growth research, where the increasing accumulation of human capital is related to the behaviour of the altruistic older generation. Inheritances left to descendants together with increasing longevity strengthen children's motivation to seek greater educational attainment. Thus the expenditure on health, which leads to greater life expectancy, can produce a virtuous circle between economic growth, education and increasing life expectancy. The virtuous circle may emerge through trust exhibited towards other people. When there is more trust and social capital in a society, new possibilities are born, which create an upward rising spiral akin to a virtuous circle of increasing flexibility and social cohesion. (Dowling and Chin-Fang 2007, 169, 177, 277.)

The theory of a virtuous circle employed in this study combines the ideas presented above. We start off with the model of welfare production. From there, we proceed from input factors through production processes to outputs and finally to outcomes. Designed by Hill and Bramley (1986), the model was adapted by Mitchell (1991) to the purposes of comparative welfare state study. Basically the model of welfare production is a linearly forward moving chain process with the public welfare inputs as an initial stage. These inputs are distributed to citizens in different ways and they may differ according to coverage and generosity. The outcomes of the process vary, examples being a decrease in poverty or inequality. The model of welfare production is presented schematically in figure 1 (p. 9).

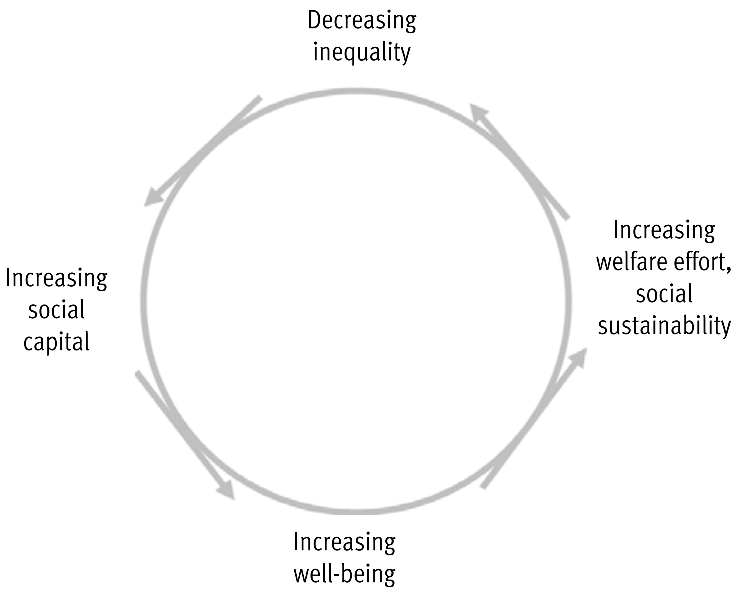
The idea of a circle appears when we consider that the chain process not only begins and ends, but that after the last stage there will be a feedback effect to the input stage and the process begins all over again. The circle process contains four stages: (1) welfare effort or the practiced social policy, (2) inequality, (3) social capital and (4) well-being. We flesh out and complement the idea of the virtuous circle by proposing that social policy inputs reduce inequality, which strengthens social capital, which increases the population's well-being. If the population finds a social policy input to have increased their well-being, this input gains the population's support. The virtuous circle and the process are presented in figure 2 (p. 9).

Figure 1. Schematic representation of the model of welfare production.



Source: Modified from Figure 10.1 in Mitchell 1991, 156.

Figure 2. The virtuous circle.



3 The virtuous circle analysed with a structural equation model

3.1 Countries and data

In the empirical analysis of the virtuous circle there are, in addition to the choice of statistical method, two central questions to be answered. The first concerns the choice of countries and the second the operationalization of the variables in the virtuous circle.

We are using data from 23 OECD countries. This means that we exclude countries that potentially could be in the so called social trap where the virtuous circle cannot start to operate, for instance because of lack of trust. In the meta-regime classification of Abu Sharkh and Gough (2010) we thus concentrate on the original welfare state regime framework instead of two other meta-welfare regimes: an informal security regime and an insecurity regime. This also restricts the choice of variables. Another question with relevance for the choice of countries is the time period. This is evident in the study by Denis Bouget, where he describes the variation in the number of countries participating in the European Community and in the European Union. (Bouget 2009.) We are using data from one cross section, which naturally limits the analysis.

The countries included in the study are: Australia (AUS), Austria (AUT), Belgium (BE), Canada (CAN), Germany (DE), Denmark (DK), Greece (EL), Spain (ES), Finland (FI), France (FR), Iceland (IC), Ireland (IE), Italy (IT), Japan (JAP), Luxembourg (LU), the Netherlands (NL), Norway (NO), New Zealand (NZL), Portugal (PT), Sweden (SE), Switzerland (SWI), the United Kingdom (UK), and the United States (US).

When thinking about the choice of variables to be included in the virtuous circle analysis, it is of course important to account for the multidimensional character of welfare states. John Hudson and Stefan Kühner note that the choice of indicators and the choice of method can have a real bearing on findings. They approach the dependent variable problem by stating that the data used to measure the welfare effort are important, but that there is a need to account for both the productive and protective elements of welfare. (Hudson and Kühner 2010.) Here we might find some similarities with the model of welfare production, which lies behind the virtuous circle.

Abu Sharkh and Gough define in their cluster analysis welfare regimes as combinations of institutions and welfare outcomes. The institutions comprise the resources and programmes that serve to enhance welfare and security. The welfare outcomes are the final welfare conditions of the citizens. (Abu Sharkh and Gough 2010, 31.) Here, again, we can find some elements of the virtuous circle.

We now operationalize the variables of the virtuous circle presented in figure 2. The variables of the circle are not observable; that is, we do not have direct data about them. These kinds of variables are called latent variables. There exist statistical methods which allow the latent variables to be presented as a combination of several indicators and to be used in statistical analysis. The four latent variables of the virtuous circle and the corresponding eighteen indicators are presented in table. Welfare effort consists of five indicators, inequality of five, and social capital and well-being each of four indicators. The indicators are defined partly on the basis of the availability of data and partly on the basis of the results of several trials and errors. (Hagfors and Kajanoja 2009a, 2009b and 2010.)

Table. The latent variables and indicators of the virtuous circle.

Latent variable	Indicator	
Welfare effort	X1	Decommodification
	X2	Public education / GDP
	X3	Public welfare services / GDP
	X4	Social transfers / GDP
	X5	Active labour market policy / GDP
Inequality	X6	Human Poverty Index
	X7	Global Gender Gap Index
	X8	Gini Index
	X9	Income mobility
	X10	Restrictions on freedom of choice
Social capital	X11	Trust in institutions
	X12	Generalized trust
	X13	Passive participation
	X14	Active participation
Well-being	X15	Life expectancy
	X16	Satisfaction with life
	X17	Gross enrollment rate
	X18	GDP per capita (PPP)

3.2 Welfare effort

The first variable in the virtuous circle is the welfare effort; i.e., the allocation of public resources to social policy. We use five indicator variables (X1–X5).

The first variable used to analyse the composition of the welfare effort of the individual countries is the index of decommodification (X1). Developed by Esping-Andersen (1990), the index was modified and recalculated by L. Scruggs (see Scruggs and Allan 2005), who has published an index of decommodification for the years 1971–2002. Scruggs himself refers to the index as an "indicator of generosity". The index describes the degree of universality and the replacement rate of social security systems. Here, we use the term "index of decommodification". It describes the level of income security afforded by pensions and by sickness and unemployment benefits. The replacement rate refers to the ratio of these benefits to earnings and coverage to the share of the population entitled to them. Decommodification is a representation of the quality of life available under a social security system to persons outside the labour market. We use 2002 figures, which are available for 18 countries.

The ratio of public expenditure on education to GDP (X2) represents the amount of resources allocated to human capital, i.e., the increase of knowledge and, indirectly, productivity through the public finances. The data have been collected by the UNESCO Institute for Statistics.

The amount of economic resources dedicated to the implementation of social policy is naturally of great import to society. We begin with the calculations of net social expenditure made by W. Adema under the auspices of the OECD (Adema and Ladaique 2005) (X3–X4). These calculations are based on the understanding that total social expenditure is exclusive of the flows of money returned to the public sector in the form of taxes and charges. This is an important distinction because income transfers may or may not – depending on the country – be treated as taxable income. Spending on public services is included in net social expenditure. Adema's calculations accommodate also private social expenditure, which varies significantly by country. This study focuses primarily on public-sector social protection. Hence, we exclude private social expenditure from the analysis and refer only to net public social expenditure.

The ratio of the public expenditure on active labour market policies to GDP is derived from the Statistical Annex to the OECD's Employment Outlook data (OECD 2006) (X5). Country-specific variation is considerable. Denmark, the Netherlands and Sweden stand out with a particularly high level of expenditure on active labour market policies.

3.3 Inequality

Here we use five indicator variables (X6–X10). Our first indicator of inequality is the United Nations' Human Poverty Index (HPI) (X6). It is also referred to as the deprivation index. HPI scores are available for 18 of the countries included in this study. The HPI index consists of four dimensions, the first of which is the share of those in a cohort who do not survive to age 60. A second factor is the ratio of those who are functionally illiterate to the population at large. A third factor is the ratio of the long-term unemployed to the labour force. A fourth component of the indicator is the population share of those with an income less than 50 percent of the median income. According to the index, the lowest levels of deprivation are seen in the Nordic countries and in the Netherlands. The score for Italy raises some doubts as to its accuracy.

Our second indicator of inequality is the World Economic Forum's Global Gender Gap Index (GGG) (X7). It provides a look at inequality from the viewpoint of gender equality. The index comprises four basic pillars, which measure the relative positions of women and men on the basis of subindexes related to economic participation and opportunity, education, political empowerment and health. (Hausman et al. 2007.)

Traditional income inequality is represented by the harmonized Gini Index (X8), which is based on a database maintained by the Luxembourg Income Study.

The fourth indicator (X9) describes the differences in the extent to which one's inherited position determines one's future progress. We use the intergenerational income elasticity (β), which represents the degree to which sons' earnings depend on their fathers' earnings. Income elasticity is derived from a model where the sons' level of earnings is explained by their fathers' level of earnings in a bilogarithmic regression. Elasticity data is available for 12 countries. Jäntti et al. (2006) calculate elasticity scores for 6 of them. Their data have been standardised and are therefore significantly better for cross-country comparisons than meta-analyses performed with the use of a variety of methods and data. The income elasticity scores calculated for the Nordic countries, the United Kingdom and the United States are based on uniform data. Because of the heterogeneity of the data, comparisons between the countries should be made with some caution (Corak 2004).

We call our fifth indicator of inequality the degree of restriction on freedom of choice (X10). Based on people's subjective views of their status, it is similar to the Power Distance Index of G. Hofstede, which represents the ability of individuals to influence the decision-making concerning themselves, mainly in relation to their supervisors and managers. Unfortunately, Hofstede's indexes are based on data which are outdated for the purposes of this study (see Hofstede 2001).

The indicator used here consists of two parts derived from the basic data for the World Value Survey 1999/2001 (Inglehart et al. 2002). First, we have selected from the data a question which asks respondents to estimate, on a scale of 1–10, their ability to influence the course of their lives. The second question concerns the respondents' view of their ability to influence the decision-making that affects their own work. The scale is again 1–10. In the index we use, the share of those who answered 1–6 to each of the questions is multiplied with itself and then squared (geometric mean). In other words, we examine the share of those who consider their ability to influence the course of their lives to be relatively limited. We choose the geometric mean because we do not expect the answers to be distributed normally. The degree of freedom of choice is equal to $(N_1 + N_2)^{0.5}$, where N1 is the share of those who answered 1–6 to the first question and N_2 the share of those who answered 1–6 to the second question. According to the indicator, residents of the Nordic countries experienced the least restrictions on their freedom of choice. The Netherlands was level with the Nordic countries. High scores were seen in France and the Mediterranean countries (Spain, Italy and Portugal).

3.4 Social capital

Social capital is the third latent variable in the virtuous circle. A framework linking social capital to inequality makes an important contribution to the debate on social capital. (The problems associated with the uneven distributions in the welfare state and social capital have been analysed by van Oorschot and Finsveen (2008).) The mainstream conception of social capital sees it as an integrative force which adds to the common good. Judging from the literature, social capital is a kind of terminological umbrella for a variety of different approaches. (On discussion of definitions and measurement, see Dowling and Hill-Fang 2007, 251–300, Svendsen and Svendsen 2009 and Kajanoja 2009.)

The indicator variables used here have been obtained from the data of the European Value Survey (EVS) and the World Value Survey (WVS). Trust and citizen networks (X13, X14) have been chosen as the measures of social capital. Generalized trust in other citizens (X12) and trust in public institutions (X11) are indicators of the former. Both are derived from WVS. Citizens' answers to questions about their trust in six key institutions were used to calculate a summation index using Likert scales. Data for 14 countries have been obtained from table 5 of van Oorschot, Arts and Gelissen (2006, 159). It is extended by the addition of variables of trust for 9 countries, which we have constructed on the basis of the WVS data. Our decision to concentrate on networked (bridging) social capital and generalized trust is supported by the key writings developing the concept of social capital (Putnam 1993 and 2000; Woolcock 2000). The indicator is a summation index with a range of [1,2] according to citizens' reported trust in others.

3.5 Well-being

Moving on along the virtuous circle, we now come to indicators of well-being (X15–X18). Four variables are used here. The range of possible indicators is extensive (see e.g. Vogel and Wolf 2004 and Saari and Sailas 2006; for a review of welfare indicators see Boarini et al. 2006 and Stiglitz, Sen and Fitoussi 2009).

The indicator of life expectancy (X15) used in the empirical calculation is obtained from Society at a Glance, a compendium of indicators compiled by the OECD.

Satisfaction with life (X16) is a subjective indicator of well-being which is based on data from the 2001 wave of the World Value Survey (Inglehart 2004). The respondents were asked about general satisfaction with their lives. The answers were scored on a scale of 1–10, where 1 was 'not at all satisfied' and 10 'fully satisfied'. Here we have constructed an index representing the number of respondents who answered 8, 9 or 10. Again, the indicator represents life satisfaction or happiness. The highest proportion of those satisfied with their lives was seen in Denmark, Iceland and Finland, and the lowest in Japan and in the Mediterranean countries.

Gross enrolment rate (X17) describes the share of those who get an education related to those entitled to education. The enrolment rate is a gross figure in which the total number of persons in education is compared to the number persons in the relevant age groups who are entitled in principle. This means that because there are always adult students and those who spend an exceptionally long time in education the enrolment rate may be over 100.

The last of our well-being indicators (X18) is the ordinary GDP per capita corrected with purchasing power parities.

3.6 The structural equation model

Structural equation models (SEMs) are a useful tool for analysing variables that do not lend themselves to direct observation. Two examples of such latent variables are intelligence and fairness. In the virtuous-circle context, latent variables include the welfare effort, inequality, social capital and well-being. The relations of these variables form the hypothetical part of the SEM. The effects may be direct or indirect. Structural equation models can be used to calculate the dependencies between variables, which may be uni- or bidirectional. Unidirectional models are recursive; they do not contain feedback loops. When relations are bidirectional, we speak of non-recursive models. The relation between the latent variables and the observed indicator variables constitute the measurement model, in which the empirical measurements are performed. Regression analyses and factor analyses are performed simultaneously. For details about structural equation models, see Arbuckle 2006, Byrne 2001 and Maruyama 1998. Figure 3 presents the virtuous circle within an SEM framework.

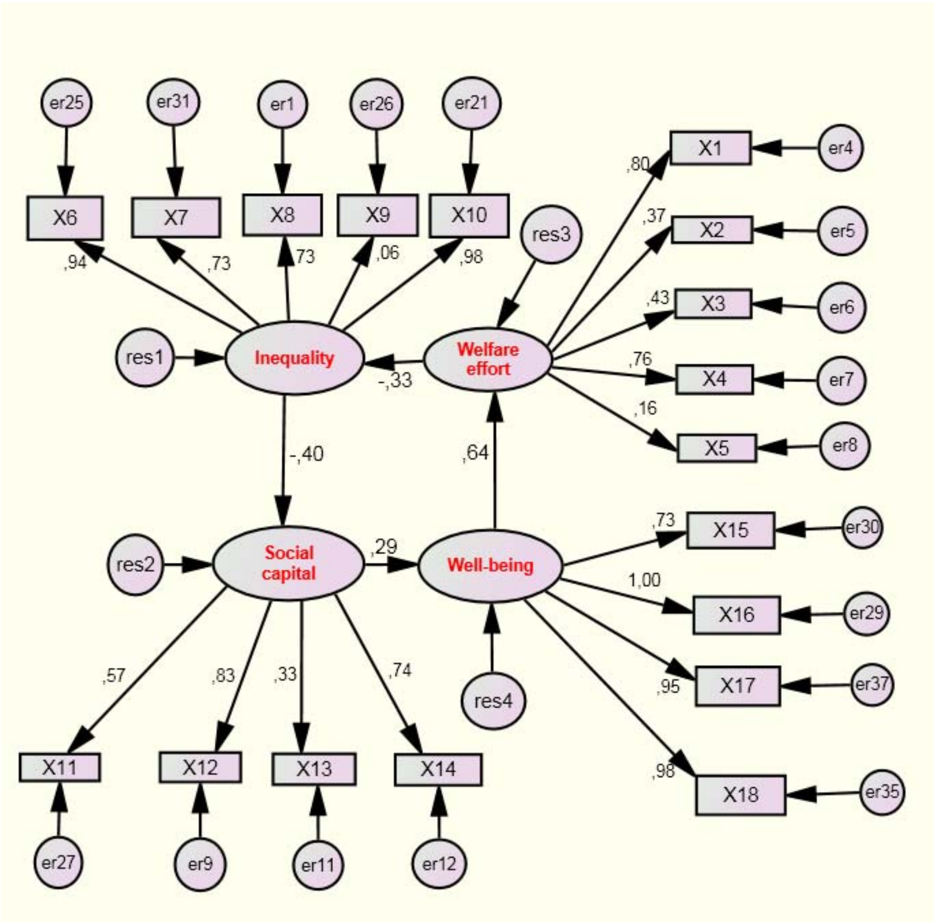
In the middle we see a structural equation model consisting of four latent variables, with four separate regressions. In the hypothetical model, the effects are presented as unidirectional arrows. The errors and residuals represent factors which have not been explicitly defined. The quantity of variables and the parameters estimated determine the degrees of freedom in the model. They, in turn, are relevant to the identifiability of the model. When dealing with a simultaneous non-recursive model, the stability of the model is a necessary condition for a solution to exist. The coefficients must fulfil a set of stability conditions which are analysed by means of a stability index. Statistical figures describing the goodness of fit of the model can be used to perform comparisons between model alternatives.

The model in figure 3 was estimated using the AMOS software. The figure shows the standardized values of the parameter estimates. They allow us to estimate the direct and indirect effects between the variables as well as the total effects of the variables. Maximum Likelihood estimation was used. In summary, we can note that the model is non-recursive. The model's degrees of freedom are 154, and the estimation yielded the minimum solution. It is a condition of stability that the index of stability should receive a value in the range of [−1,1]. Here the index value is 0.155. The value of the log likelihood function is 3174.1. To achieve identifiability, the variances of the error terms have been normalized in the standard way. The effects on the virtuous circle can be examined by reference to the standardized coefficients of figure 3. We are interested mainly in how the welfare effort impacts inequality and, indirectly via the change in inequality, the social capital. A fuller picture of the causes and effects in the virtuous circle is gained by looking at the effect of social capital on well-being and that of well-being on the welfare effort. This method helps us to identify a causal chain in which the welfare effort, through other latent variables, also has an indirect effect on itself, which was one of the starting premises of the virtuous-circle hypothesis.

In figure 3 direct effects are symbolized by unidirectional arrows. Indirect effects are derived by multiplying the direct effects among each other. The total effect is the sum of direct and indirect effects. The results show the welfare effort as having a negative impact on inequality (coefficient –0.33). This conforms to our hypothesis. Inequality also has a negative impact on social capital (–0.40). The indirect effect of the welfare effort on social capital, via reduced inequality, is positive (0.132). Since the welfare effort, as postulated in our model, has no direct effect on social capital, but rather an indirect effect via inequality, it is the total effect at the same time.

As we move along the circle from social capital towards the latent variable of well-being, we receive a positive coefficient of 0.29 for the regression coefficient. The indirect effect of the welfare effort and, at the same time, its total effect on well-being is 0.038. The final stage of the circle measures the relation between well-being and the welfare effort. A positive regression weight of 0.64 is derived for the direct effect. Finally, based on the coefficients we can calculate the effect of the welfare effort on itself, which effect is an indirect one and intermediated by inequality, social capital and well-being. At 0.025, the coefficient implies a positive feedback loop, which keeps up momentum in the circle. This is an important result both in terms of the virtuous-circle hypothesis and of social sustainability.

Figure 3. The virtuous circle. Standardized coefficients.



Having now defined the relevant indicators and estimated the parameters for a well functioning virtuous circle we move on to study how this kind of information can be used in the comparative welfare state research. The question then is how can welfare states be described according to their virtuous circle characteristics? Is it possible to find classes or typologies using the virtuous circle as a base for classification? We approach this question using neural computation methods and the self organizing map.

4 The Self-Organizing Map

> ”In practice people have applied many methods long before any mathematical theory existed for them and even if none may exist at all. Think about walking; theoretically we know that we could not walk at all unless there existed gravity and friction, by virtue of which we can kick the globe and the other creatures on it in the opposite direction. People and animals, however, have always walked without knowing this theory.” (Teuvo Kohonen 2001, XI–XII.)

The principles according to which neural networks work have their origins in an analogy involving the architecture of the parts of the human brain and its ability to learn. It consists of neurons (also called nodes or cells) which can take impulse signals, process them and send outputs to other neurons. As in the case of biological neurons certain impulses make neurons in different parts of the network react by firing. The neurons are connected to each other and the connections have specific weights. The training of the network proceeds by feeding the network with input data, comparing the obtained output value with the actual correct output values, which are known in advance, calculating the error and feeding it backwards to the net by changing the weights to correspond better to the actual value. The input data is fed again to the network and new output values are calculated using the re-adjusted weights. After repetition of this procedure the weights will adjust so that they produce the actual output value. The structure of the weights is stabilized and the learning process completed. Now the trained network (in statistical terms, the estimated model) is ready for use, for instance, to classify new cases or to produce forecasts for time series. In the case of neural networks the question is one of computational intelligence. The learning means that the weight structure can be altered by experience.

One example is diagnosis-making and the so called expert systems that have been developed for it. Such systems comprise large computer programs utilizing the experiences gathered by doctors in making decisions on the basis of symptoms. On the other hand, in the neural network approach the network is trained using real diagnoses and information on the corresponding symptoms. By training the network the dependencies between diagnoses and symptoms are learned from real cases. The trained network can be used in making new diagnoses on the basis of symptoms.

The origins of neural networks can be traced back to the 1790s. However, it was not until the beginning of the 1940s that neurobiologists and statisticians formulated mathematically what kind of calculations a network of binary-valued neurons is capable of performing. In the 1950s and 1960s this field was called connectionism and the main processing element was called the perceptron. The perceptron was the first neural network with ability to learn.

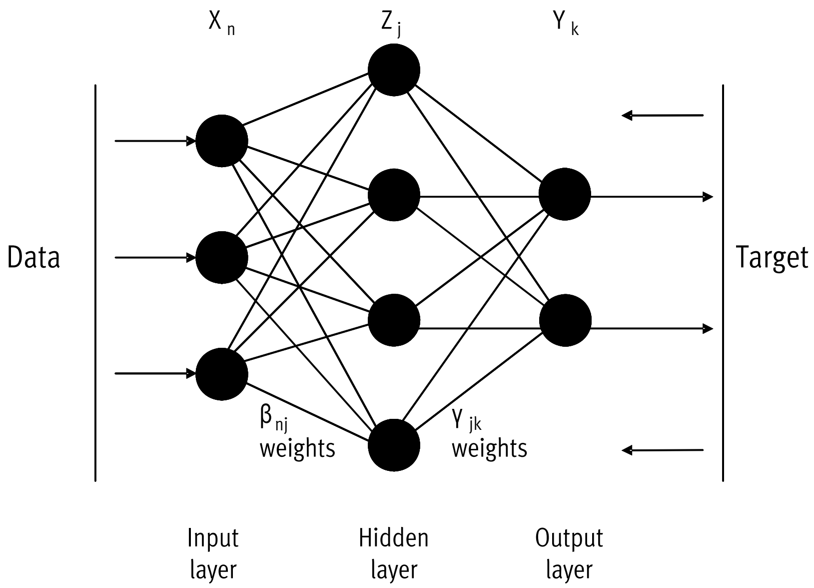
In the simple network there are two layers, the input layer and the output layer. The target value is known. The node of the output layer receives the inputs multiplied by a corresponding weight. The output is calculated using weighted summation and an error is calculated by comparing it to the target value. The weights are then changed using the difference between the current and desired outputs, which difference the process aims to minimize.

At the end of 1960s it was found that the perceptron could not distinguish patterns that were not linearly separable. This means that they could not solve problems with nonlinear relationships between inputs and outputs. This limitation led to the shifting of research resources to a rival school of thought, called artificial intelligence, which contains such areas of study as expert systems. A few researchers continued to work with neural networks. One of them was Teuvo Kohonen at the Helsinki University of Technology. His research on the theory of self-organization, associative memories, neural networks and pattern recognition had a great effect on reawakening interest in neurocomputing research at the beginning of the 1980s. The problems with nonlinearities were solved when a new learning model was developed. The most commonly used modern neural network learning rule is called backpropagation. (Smith1996, 20–26.)

There are numerous variants of the backpropagation model, but generally the process assumes a layer of input nodes, a middle layer of hidden nodes, and a layer of output nodes. The weighted inputs are summed, processed and output to the next layer of neurons. In the learning process the weights are adjusted according to input data by simulation until convergence with the term ”learning” or ”training”, corresponding to the statistical term ”estimation”. (Glossary of neural terms, Kohonen 2001, 373–401.)

In figure 4 the net is composed of an input layer, a hidden layer and an output layer. Here we have input nodes x_n and the weights β_{nj} between input nodes and nodes z_j. The nodes of the hidden layer process the weighted inputs and the nodes y_k of the output layer get these as inputs multiplied by weights γ_{jk}. The output from nodes in the output layer is compared to the values of the target and the differences for correcting weights are calculated. The process then turns in the other direction. Using the corrected weights the desired output to the previous layer is calculated and the weights are corrected. Every node contains a summation function for inputs and an activation function (transfer function) for outputs. One proceeds in this manner until the updates of the weights are smaller than some cut-off criterion.

Figure 4. A multi-layer perceptron and backpropagation learning algorithm.



The neural computation can contain supervised learning nets or unsupervised learning nets. (Garson 1998, 37–58). In the supervised learning nets the correct targets are known. This is not the case with unsupervised learning nets. The nets cluster input examples according to similarity and the process of unsupervised learning is an algorithm of compressed information without reference to expected correct classifications. The network uses a kind of competitive method where the nodes compete with each other according to given criteria to be the best representative node for the examples given in the input data. Unsupervised learning is therefore also called competitive learning. One of the best-known unsupervised neural network techniques is the Self-Organizing Map invented by Academician Teuvo Kohonen, which we are using in this study for purposes of welfare state classification.

There are three subsets among the soft computing technologies that have been used in classifications. These are fuzzy logic, genetic algorithms and neural networks. The term fuzzy logic is in a broad sense synonymous with fuzzy set theory, which has been developed for comparative welfare state research by Ragin (2000) and applied among others by Kvist (1999, 2000 and 2002) and Hudson and Kühner (2010). The advantage of fuzzy logic is its ability to approximate reasoning, its disadvantages being that the construction of fuzzy membership functions and rules is difficult and the method lacks an effective learning capability. The genetic algorithms are the most commonly used subset of evolutionary computing, the other main subsets being evolutionary programming and evolutionary strategy. The genetic algorithm method has been used for classification purposes in insurance-related applications. Its advantages are the availability of systematic random search and derivative-free optimization, while its disadvantages include the difficulty of tuning and lack of convergence criteria. On the other hand, neural networks have the advantages of adaptation, learning and approximation, but the disadvantages of slow convergence speed and the so-called black box data processing structure. (Shapiro 2002, 115–131.)[1] Empirical applications are numerous.[2]

The clustering of welfare states is made by means of a method invented by Teuvo Kohonen. The central dependencies of the data are represented on a kind of two dimensional grid or map. The learning is carried out by iterative regressions. The data input values are connected to each node of the map and the models become ordered so that similar models move closer to each other while dissimilar models move away from each other as the iterations proceed.

Presented more formally the self organizing map consists of a grid of units, or “neurons”. On the grid each unit represents a model, which consists of a vector of features. In the map the observations will be presented optimally with a subset of models. In the course of this process similar models are ordered close to each other and dissimilar ones far from each other. The computation is made by a sequential weighted regression. The winner index c is identified in the process using the following condition:

(1) $\forall i, \; ||x(t) - m_c(t)|| \leq ||x(t) - m_i(t)||.$

1 The following nine useful properties of neural networks have been summarized (Garson 1998, 15): Nonlinearity, input-output mapping, adaptivity, evidential response, contextual information, fault tolerance, suitability for very large scale implementation, uniformity of analysis and design, and neurobiological analogy. Neural networks differ from conventional statistical computing in terms of their massive parallelism, high interconnectivity, simple processing, distributed presentation, high fault tolerance, collective computation, and the ability of self-organization. Neural network models are universal, nonparametric and robust.

2 Neural networks have been applied in several fields including engineering sciences, economics and business studies, sociology, political science, and psychology. There are, of course, numerous other more specific applications within these fields, such as measurement, process control, robotics and various types of analysis performed as part of industrial processes. Neural networks have been used in time series forecasting, financial analysis, experimental physics, chemistry and medicine. Neural networks are applied to fields of speech recognition and machine vision, just to mention a few applications. (Garson 1998, 111–148. See also Kohonen 2001, 327–328.)

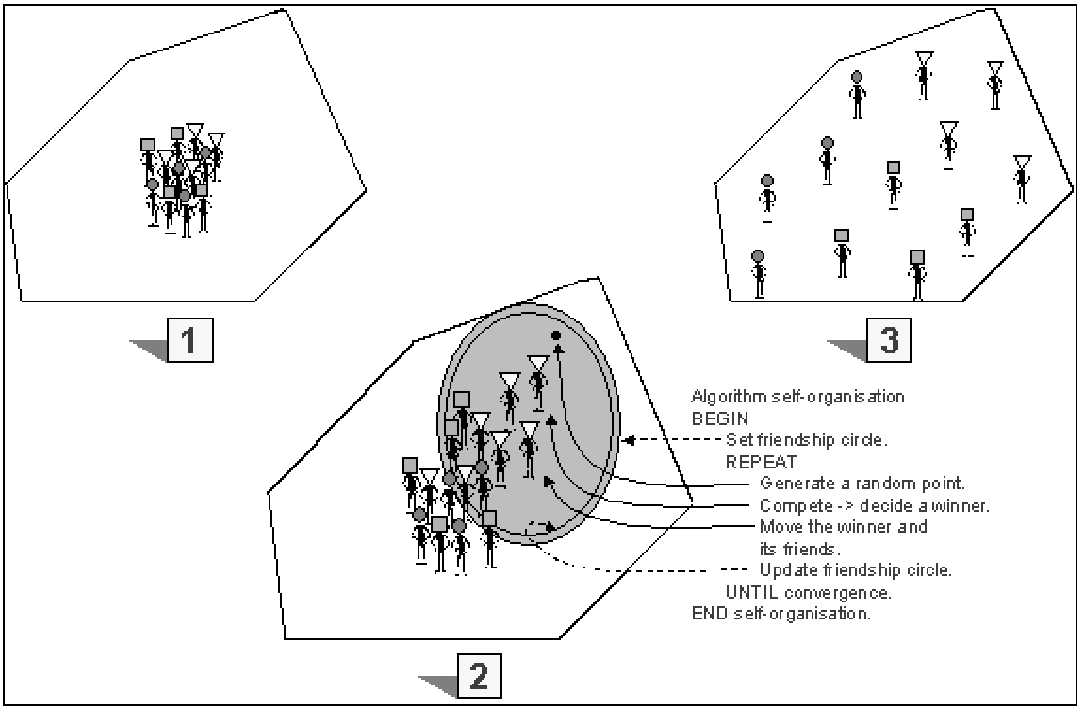
In (1) m is a model vector and x(t) are samples. Next the model centred around node c is updated in the following way:

(2) $m_i (t + 1) = m_i (t) + h_{c(x), i} (x(t) – m_i (t))$.

In equation (2) $h_{c(x), i}$ is the neighbourhood function, which is a function of the distance between the ith and cth nodes. It decreases during the iterative process. We shall here mention only the main areas of application of the Kohonen map (Kohonen 2001, 358–369).[3]

The self-organization procedure is presented in figure 5 in an illustrative way. There we have a group of people at the centre of a polygon. Each person is a friend of someone else with the degree of friendship between them indicated by the similarity of their heads. Here a triangle and square are considered to have greater similarity than a triangle and a circle. We have two objectives: first to make uniform the distribution of space within the polygon and, second, to make the distance between two persons proportional to their similarity.

Figure 5. Self-organizing algorithm schematically presented. Panel 1: Initial crowd. Panel 2: Process. Panel 3: Result.



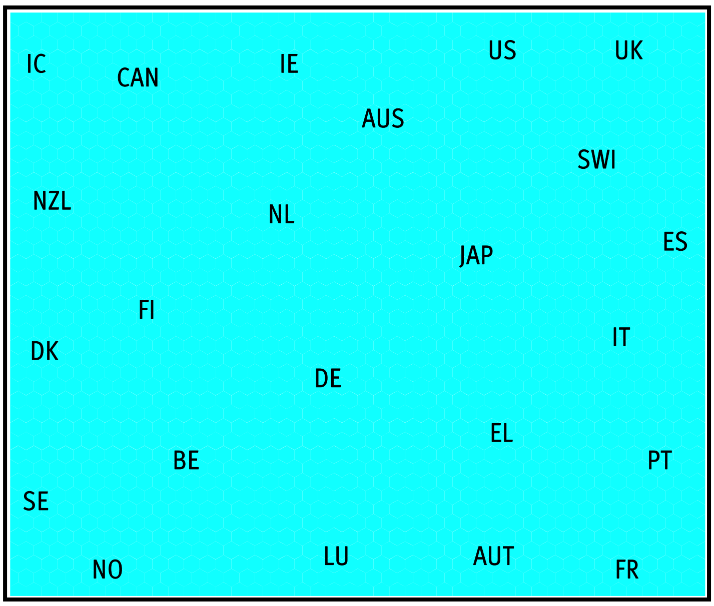
Source: Adapted from Hemani and Postula (1999) figures 2a, 2b and 3.

3 Teuvo Kohonen's work on the Self-Organizing Map has provided the inspiration for thousands of scientific papers around the world. In them the Kohonen map is analysed and applied to data analysis and pattern recognition problems. (See Kohonen 2001, 403–486; Oja and Kaski 1999.) An interesting application is the Kohonen map of welfare and poverty in the world (Kaski and Kohonen 1995). Other fields of application include: Machine vision and image analysis; Optical character and script reading; Speech analysis and recognition; Acoustic and musical studies; Signal processing and radar measurements; Telecommunications; Industrial and other real-world measurements; Process control; Robotics; Electronic-circuit design; Physics; Chemistry; Biomedical applications without image processing; Neurophysiological research; Data processing and analysis; Linguistic and AI problems; Mathematical and other theoretical problems. The last mentioned area involves the hybridization of the Self-Organizing Map with other networks such as genetic algorithms and fuzzy logic.

Following Hemani and Postula (1999, 233) we can present the self-organizing process in five steps: 1. Set process parameters. First set friendship circle to a large value which gradually decreases with each iteration. It controls how many friends a person has at each iteration. As the circle decreases more distant friends are excluded. 2. Generate a random point in the polygon. 3. Set up a competition. The winner is the person closest to the random point. 4. Move the winner and his friends in the circle towards the random point. 5. Decrease the friendship circle until the circle is small enough to have no friends in it. In panel 3 of figure 5 we have the result, where persons are uniformly distributed because everyone competes for points which are generated on equal terms. Clustering results because the winner in moving himself also moves some of his friends towards the target point. Iterations reinforce clustering.

As an illustration of how the self-organizing mapping method functions as a tool for classifying welfare states, we present here an application for clustering 23 OECD countries. Data consists of the eighteen indicators which were used in the formation of the virtuous circle. The software package used here is Viscovery SOMine 5.0. The dimensions of the map were chosen so that there are altogether 2000 artificial neurons or nodes in the map. There are then 2000 possible locations where the 23 countries may be placed during the process. The approach here is an extension of the research by Hagfors (2004) and Hagfors and Kangas (2004), both of which studied exclusively the structure of social protection financing. We present a visualization of the self-organizing process on a Kohonen map in figure 6. We present also its component planes according to each indicator.

Figure 6. The Kohonen map of 23 OECD countries.



The positions of the countries on the map in figure 6 are the result of an iterative process. We can say that for instance Belgium is more similar to Germany than to Spain and that Australia is more similar to the Netherlands than to Sweden. The positions are, however, the result of a process where all relations affect 18 dimensions. We cannot make classifications without a clustering method. Next we approach the welfare state classification using the SOM-Ward cluster method, which can utilize the local ordering information of the map.

5 Comparative robustness of regime typologies

Considering the abundance of typologies and regimes it comes as no surprise that some authors have tired of the classifications and have considered the modelling business (an 'academic industry' in the words of Abrahamsson, 1999) to be irrelevant to the study of European welfare policy (Baldwin, 1996). However this may be, one of the background factors for our study here is the observation that real welfare states are not pure types but hybrid cases, and that the issue of ideal-typical welfare states remains unsolved because of an inadequate theoretical background or deficiencies in the methods used for comparative research. (Arts and Gelissen 2002.) The problem of cluster analysis is the need to fit the cases into mutually exclusive subgroups when the conceptualization of policy regimes is ideal-typical. (See Shalev 2003.) If the analysis is carried further towards factor analysis the requirement of linear combinations of variables becomes crucial. The assumption of linear relationships between the variables may not, however, stand the test of empirical analysis.

The notion of an ideal type is important and associated with the question of the two methodological strategies for comparative welfare state research: quantitative, variable-oriented research and qualitative, case-oriented research. The differences of these approaches are described in Ragin (1987 and 1994). His method of comparative social research (QCA) is based on the application of Boolean algebra and binary data. However, Ragin's method has the disadvantage that it requires data to be in binary format. Its usefulness is limited by its qualitative nature, which means that QCA variables cannot be converted into binary form without loss of information. An improvement in this respect is to use fuzzy logic or fuzzy set theory in comparisons. Fuzzy logic was formulated in the mid-1960's and is one of the three principal components of the soft computing technologies that have been used also in classifications. The two others are neural networks and genetic algorithms.

Ragin (2000) has developed fuzzy set methods for comparative studies in social sciences and Kvist (1999, 2000 and 2002) applied this theory to assess the conformity of the Nordic countries to ideal types and to models of unemployment compensation of the Northern European policy. In fuzzy set theory, partial membership allows membership scores consisting of three, five or seven values or located within a continuous interval between 0 and 1. The correspondence between concepts and fuzzy set membership scores is important. Also the way in which the attributes or aspects are transformed into sets is important because the combinations of the sets establish the corners of the property space which defines the ideal types. In fuzzy set theory there are 2^k ideal types, where k is the number of aspects. The aspects or concepts may come from theoretical bases, other knowledge or the nature of the data. A recent study by Hudson and Kühner (2010) concludes that the fuzzy-set ideal type analysis offers considerable advantages over more traditional, statistically rooted analysis (z-scores, hierarchical cluster analysis or factor analysis).

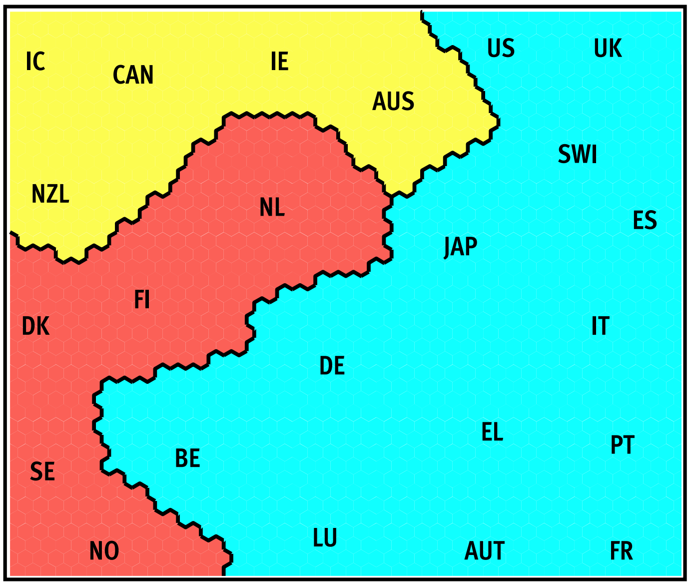
The weakness of the fuzzy-set ideal type method is that as the number of constitutive aspects increases, so does the number of ideal types and some types do not exist empirically at all or are irrelevant in some other sense. Another problem besides that of empty cells is the arbitrariness of the scoring of indicators, which depends on the researcher's knowledge of the indicator. The better that knowledge is, the more accurate the scoring. Another weakness in fuzzy-set analysis is the need for minimum requirements in classification. (Shalev 2003.)

Like fuzzy logic systems, neural computing is a soft computation technology but it has a feature which fuzzy logic lacks, a form of artificial intelligence, the ability to learn complex patterns and trends from data. Neural networks are also alternatives to traditional statistical methods when the relationships are highly nonlinear and data is large, noisy and imperfect. (Smith 1996; Garson 1998; A glossary of "neural" terms in Kohonen 2001.) For instance, traditional statistical methods perform very badly when asked to recognize a face, a task which presents no problems for humans. Neural networks are capable of pattern recognition partly because the structure of a neural network imitates the architecture of the human brain. In our exercise we use the mapping ability of neural networks in order to map a large number of inputs to a small number of output classes. In this procedure an unsupervised learning approach is used to cluster 23 OECD countries into classes according to the 18 indicators of the virtuous circle. The clusters are defined completely from data and no model for clustering need be defined a priori.

We are using the SOM-Ward clustering method, which combines the local ordering information of the map with Ward's classical hierarchical cluster algorithm. At first each individual node forms a separate cluster. In each step of the algorithm two clusters are merged, namely those with minimal distance according to a distance measure, the SOM-Ward distance. It takes into account the Ward distances as well as the positioning of two clusters in the map picture by defining that the distance of clusters that are not adjacent is always infinity. In this way the process of merging clusters will be limited to topologically neighbouring clusters.

For each cluster count an indicator is computed, which is a heuristic quality measure for each cluster count. It helps to find an initial clustering. When the indicator is high for a particular cluster count, the clustering may be seen as natural for the map. If the cluster count indicator has a low value, the clustering is artificial. The peaks of the cluster indicators show interesting clusterings. One of them is the clustering in figure 7, where the number of clusters is three. We can find analogies with Esping-Andersen's three worlds of welfare capitalism.

Figure 7. Three SOM-Ward clusters in the Kohonen map.

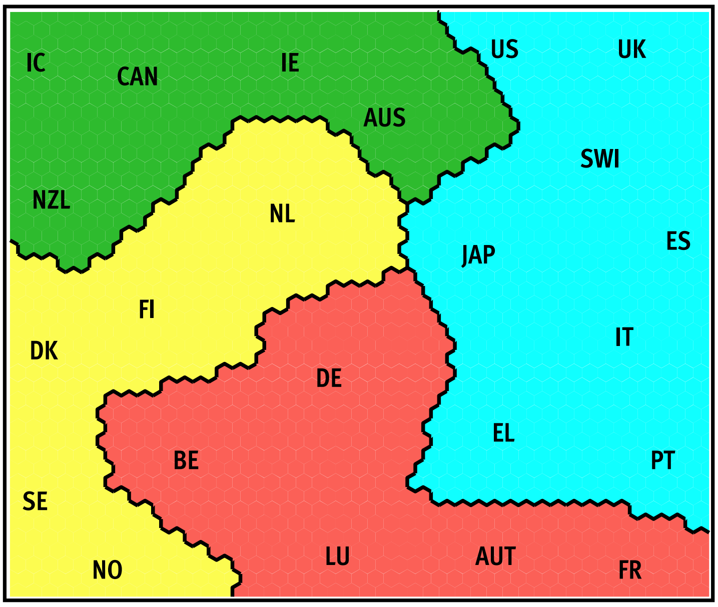


While the original 'three worlds' typology was based rather one-sidedly on income maintenance, the clustering in figure 7 is constructed using the similarities and differences of models according to 18 indicators of the virtuous circle. We are here closer to the conception of Esping-Andersen, according to which "we must weigh the relative importance of different, possibly conflicting attribute" and "no regime, let alone country, is pure". (Esping-Andersen, 1999, 88.)

Looking at figure 7 we can see that the Nordic countries clearly cluster together. While in earlier research the Netherlands have been considered to possess attributes from several regimes, here it is clustered with the four Nordic countries. There is another cluster consisting of five countries including the Oceanian countries, Canada, Ireland and Iceland. It is difficult to identify this cluster as an Esping-Andersen's welfare regime. Also the third cluster includes countries both from liberal and conservative-continental regimes of welfare capitalism.

Esping-Andersen makes the following comments on the number of models: "There will always be slippery or ambiguous cases, and one programme does not define a regime. The real problem is how to deal with systematic deviants. The issue here is whether a three-way typology adequately exhausts the variance. If there are cases that follow a wholly different underlying logic, we would have to construct yet another, separate ideal-type – a fourth 'world of welfare capitalism'." (Esping-Andersen, 1999, 88.) This is what we have done in figure 8.

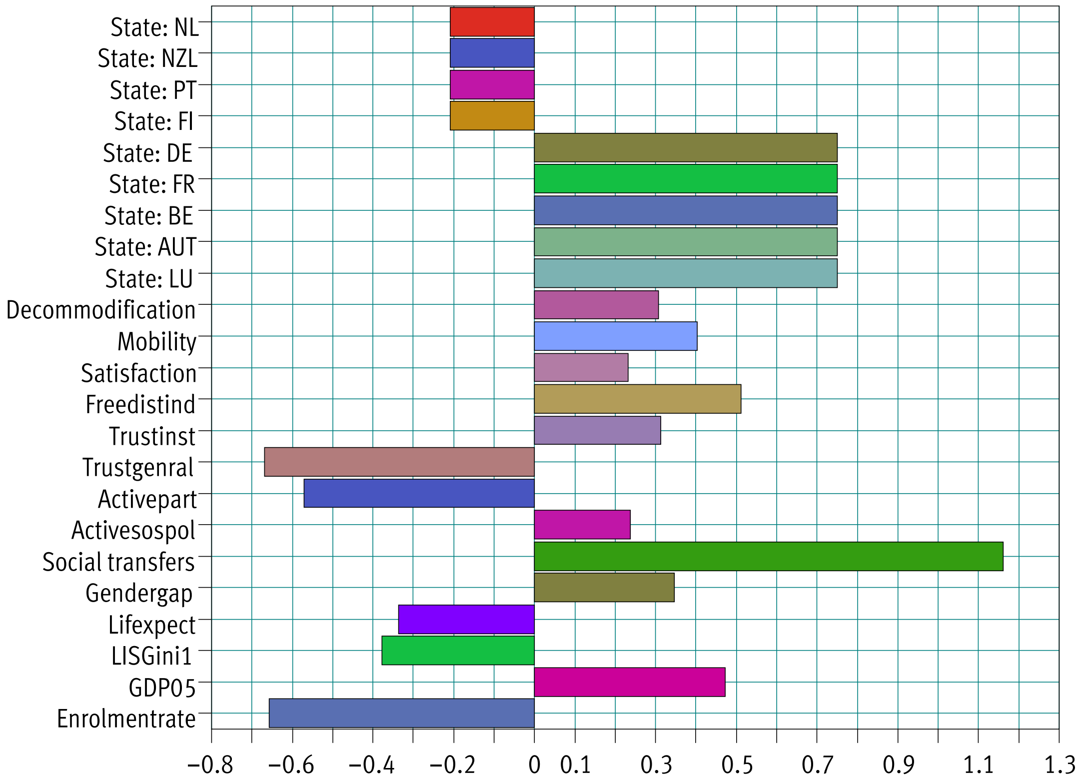
Figure 8. Four worlds of welfare capitalism.



Now we can identify four virtuous circle regimes. The big cluster of thirteen states in figure 7 has been divided into two clusters. The first consists of central European countries, while the second consists of Mediterranean countries, Japan, the United States, the United Kingdom and Switzerland. If we refer to the new clusters as the continental-corporative regime and the liberal regime, respectively, then the Nordic regime comes close to the Three World regimes. Except that we now have a fourth regime including both Oceanian countries. We call this regime the Antipodean regime. Before we continue the clustering further, we might take a closer look at the structure of the virtuous circle or the group profiles of the four clusters: the continental-corporative regime, the Nordic regime, the Antipodean regime and the fourth world regime. This is done on the level of averages in figures 9–12 and more specifically at the country level in figures 13–16.

The figures feature a bar chart for each regime. Each bar displays the deviation of the mean of the range from the mean of the entire data set, where the unit consists of the standard deviations of the entire data set. Only those attributes are displayed whose confidence exceeds the preferred confidence level. In a two-sided T-test the 95% confidence level was selected.

Figure 9. The characteristics of the Continental-Corporative virtuous circle regime.



In the virtuous circle it was found that when social effort is high, inequality is low and, as a result, social capital is at a high level. This will raise the level of well-being, and finally leads to support for welfare policy. In figure 9 we can look at the indicators of the continental-corporative cluster from this point of view. The level of decommodification exceeds by a small margin the average level, as does the level of participation in active labour market policy. On the other hand, the level of social transfers is at a very high level indicating the type of social security system where public services are at an average level. Income differences are smaller than average. On the other hand, there exists relatively little generational mobility, less freedom of choice than on average and more gender inequality than on average. There is some trust in institutions, but trust in other people is lacking. People do not participate actively and though the income level is above average, enrolment in public education is at a low level. In this cluster people's satisfication with life is above average.

Figure 10. Characteristics of the Nordic virtuous circle regime.

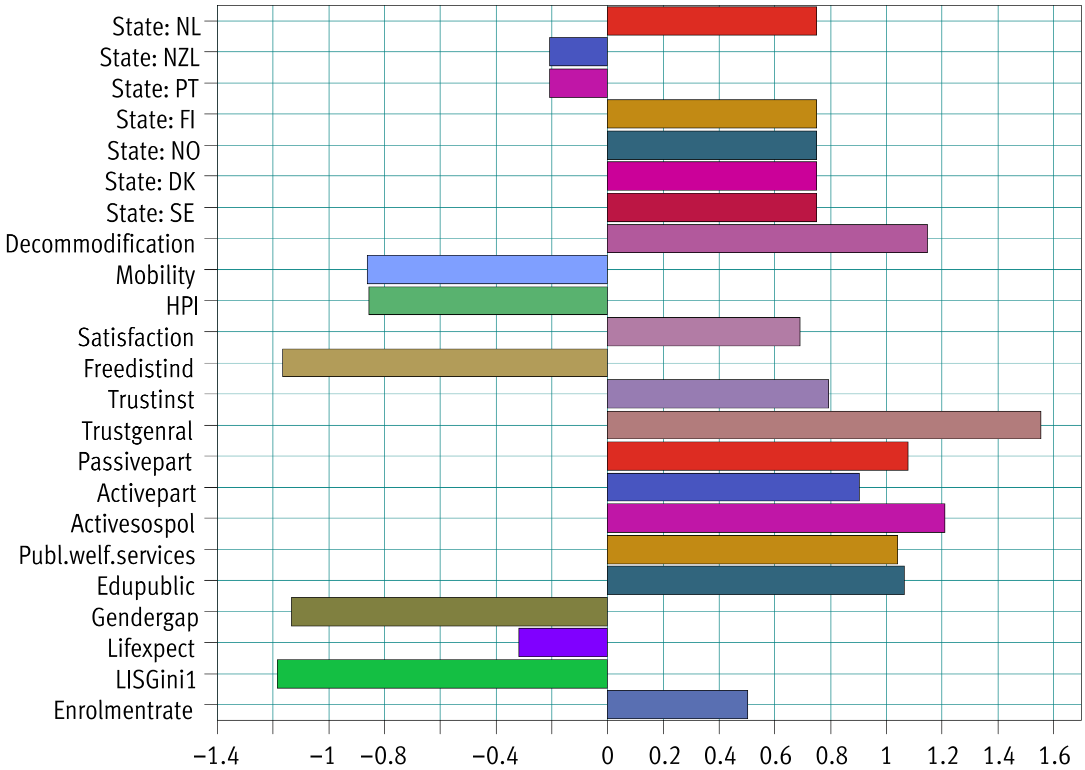


Figure 11. Characteristics of the Antipodean virtuous circle regime.

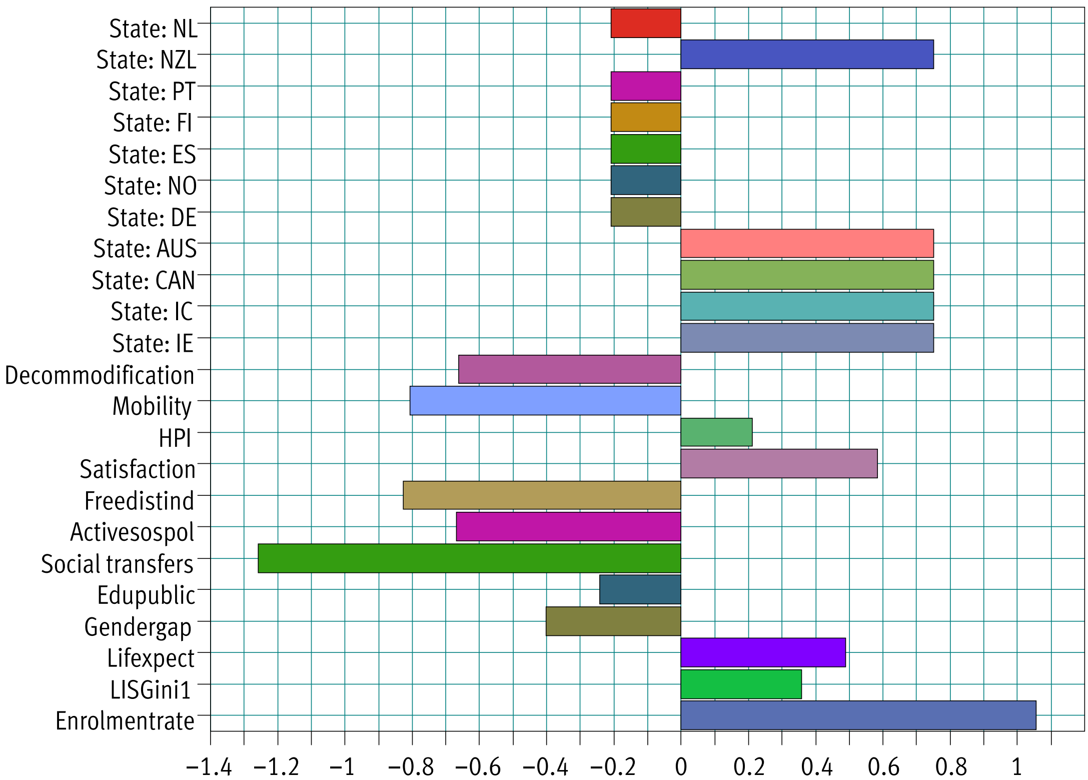
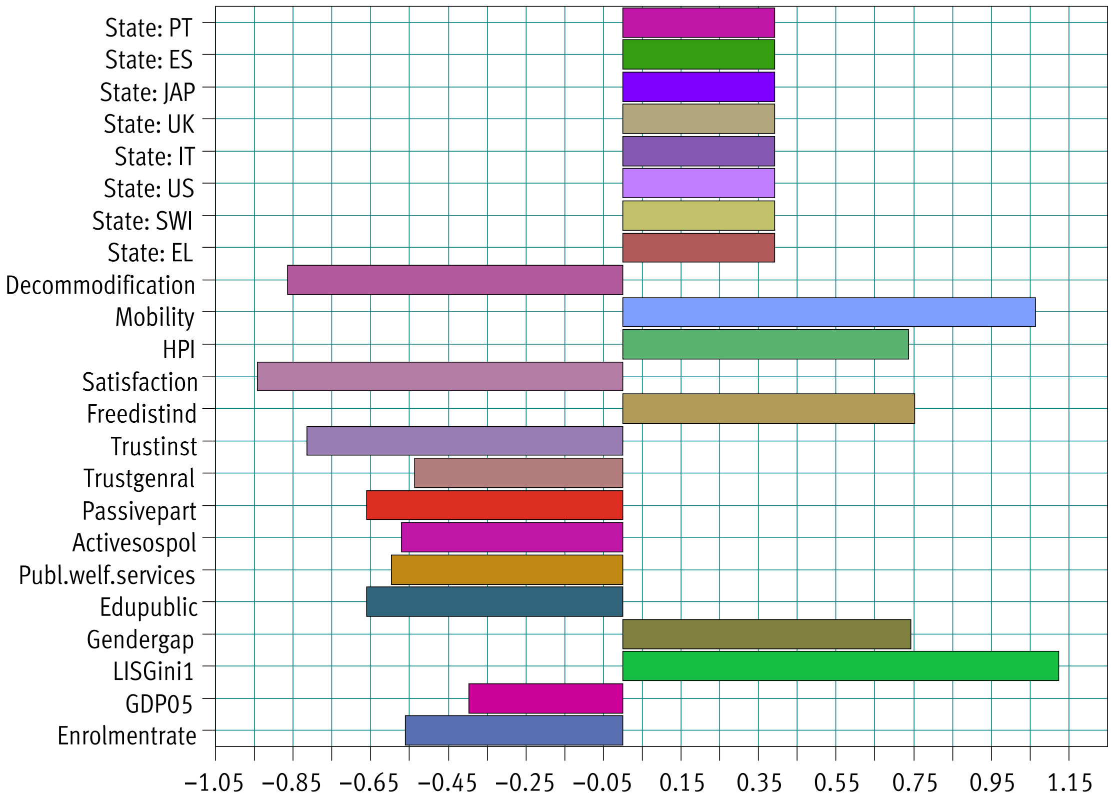


Figure 12. Characteristics of the Fourth world virtuous circle regime.



In figure 10, representing the Nordic regime, we see a different picture. Welfare effort is strong. Decommodification, activating labour market policy, public welfare services and expenditure on public education are all at a high level. Gender mobility is relatively unfettered, the human poverty index is low, people have possibilities to make decisions concerning their lives, there exists relatively high gender equality and income differences are low. Social capital is high, people trust both institutions and other people. The participation rate is high and people are relatively satisfied with life.

In the Antipodean regime, portrayed in figure 11, we find a low level of decommodification and of activating labour market policy and especially low social transfers. Also the level of public education is below average. Income differences are a little higher than on average, but there exists social mobility, freedom of choice and the gender gap is smaller than average. Participation in public education is high and people are relatively satisfied with life, and life expectancies are somewhat longer than average.

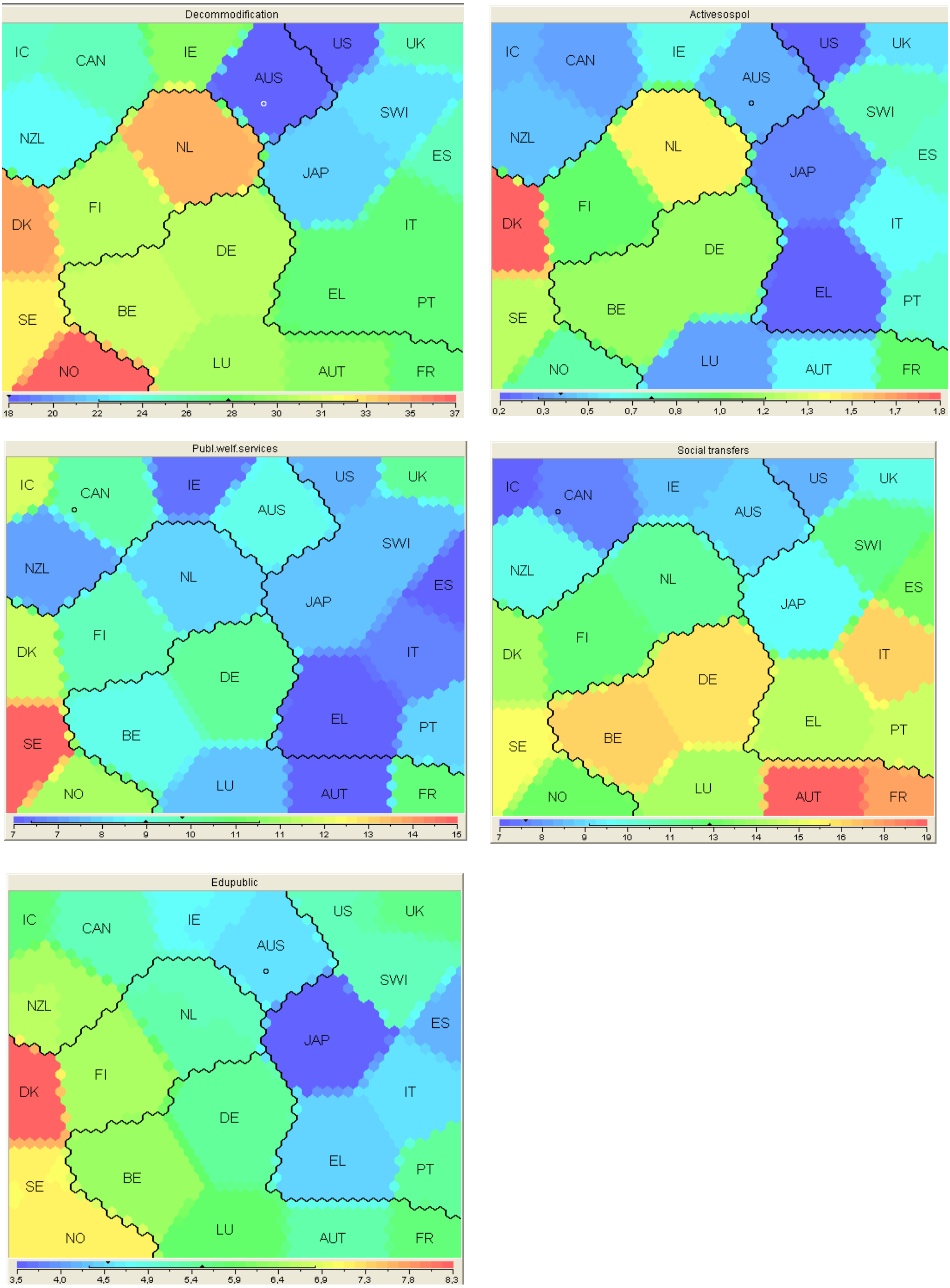
Finally the fourth relatively heterogenous cluster in figure 12 reveals a low decommodification rate, only some activating labour market policy, a low level of public services, and small inputs in public education. In this case there exists little income mobility, the human poverty index is higher than average, gender inequality is high and income differences are the highest of all of the clusters. People's trust in institutions and other people is low. The income level is lower than average as is participation in public education. In this cluster dissatisfaction with life is the highest of all clusters.

Having now described the characteristics of the clusters at an average level we take a closer look at the country level. For this purpose we shall use the figures from 13 to 16, where the Kohonen map is "sliced" into component planes in order to see how the value of an individual characteristic variable contributes to the node of the Kohonen layer. The component planes are converging simultaneously in the iterations of the learning process and the locations of the indicator value cells are located here correspondingly to the nodes of the Kohonen layer. Each plane represents the value of one component of the characteristics vector at each node of the map, using the colour scale representation where blue represents a low value and red a high value.

In figure 13 we have the indicators of the latent variable welfare effort in five panels. The first panel shows how the value of the decommodification index is distributed on the Kohonen map and in the four clusters. From the colour scale beneath the panel we can see that the highest values are in the Nordic cluster, especially in Norway, Denmark, and the Netherlands. Especially low values are found in Australia in the Antipodean cluster and in the United States in the fourth world cluster. In the second panel we can see the expenditure share of activating labour market policy. Here the Nordic countries (especially Denmark) and the Netherlands are at a high level. A very low level may be seen in the United States, Japan and Greece. In Belgium and Germany this indicator is at a relatively high level.

The following two panels present the component planes of the main elements of social protection, public welfare services and social transfers. The differences between the systems are clearly seen in the high value of public services in Sweden as well as in the high value of social transfers in Austria. The last panel of figure 13 presents the share of public education expenditure, which is high not only in Denmark but also in Sweden and Norway. This particular kind of welfare effort is very low in Japan, pointing to the dominant role of private education there. All in all, the panels confirm the earlier observation that the Nordic cluster is the leader in terms of the level of welfare efforts, but that there exists variation also inside that cluster.

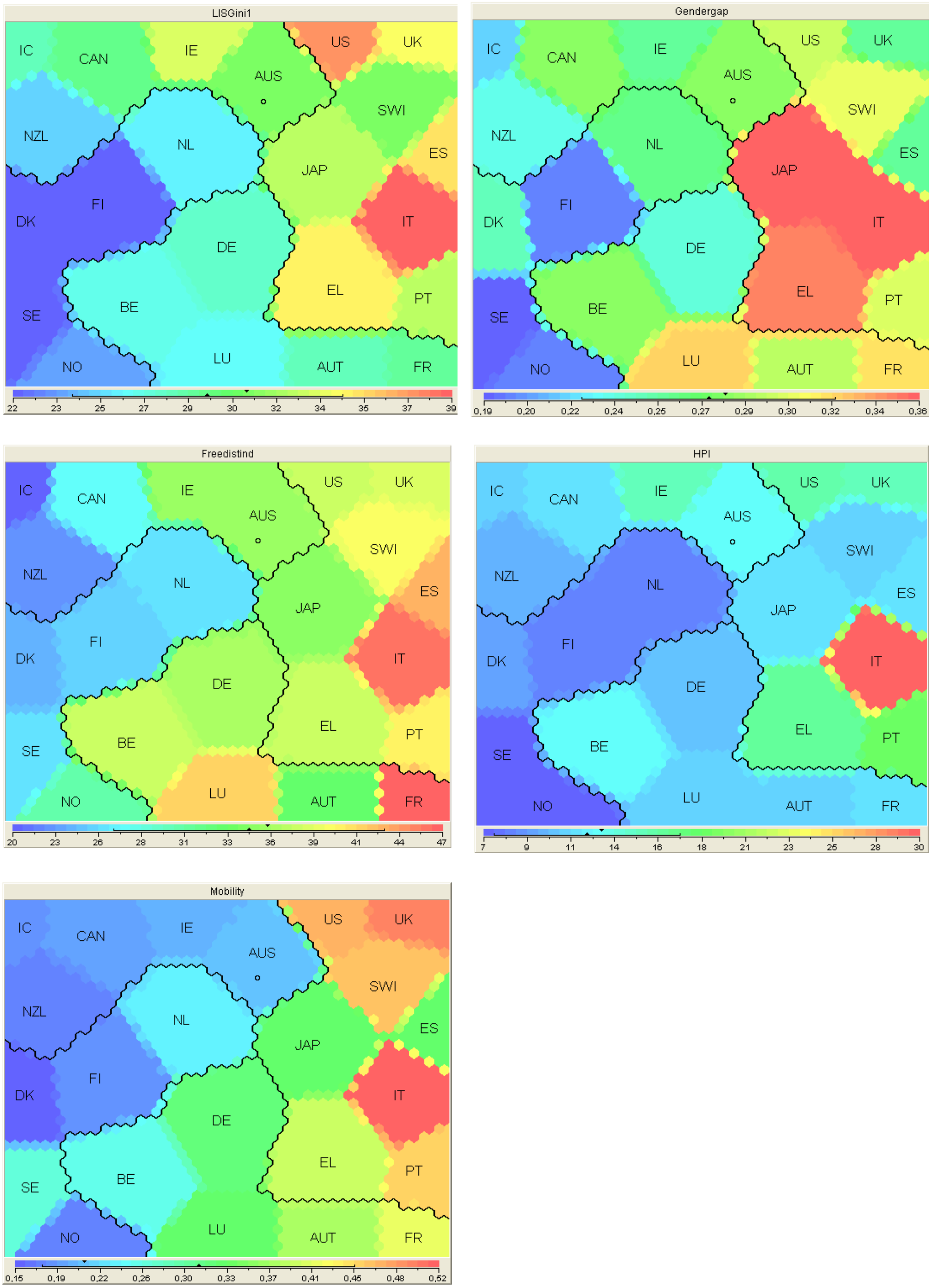
Figure 13. The component planes of the latent variable Welfare effort.



In figure 14 the indicators of the latent variable inequality are presented in five panels of the Kohonen map. In the first panel we have the Gini index values of the countries. It seems that the lowest Gini values are found in the Nordic cluster, the values being somewhat higher in Norway and the Netherlands than in the other three countries. The largest income differences are found in Italy and in the United States as well as in Greece. In the Antipodean regime the differences are not as great in New Zealand as in the rest of the countries of that cluster. Income differences according to the Gini index in the corporative-conservative cluster are largely at the average level. Like income differences, the gender gap index in the second panel shows great variability within clusters and countries. The highest gender differences exist in the fourth world cluster of our classification. Italy and Japan are leaders, but Greece is close behind them. Again the lowest values can be found in the Nordic cluster, with Sweden, Denmark and Finland showing small gender differences.

In the remaining three panels we have the distribution of the possibilities to affect decisions affecting one's own life, the human poverty index consisting of several indicators and income mobility, representing the possibility to move on the income scale. The Nordic cluster seems to afford the best possibilities for citizens to make decisions concerning their lives. In the Antipodean cluster this is true of Iceland and New Zealand, and to a certain extent also of Canada. The least possibility to influence one's life is found in the Mediterranean countries, especially in Italy. In the Corporative cluster the least possibility to influence one's life is seen in France. In the next panel the value of human poverty index is distributed rather evenly. However, the Nordic cluster has lower values than average. One notable exception is Italy, where the value of the index is high. This may be the result of one of the component indicators of the index. The last panel presents the income mobility of citizens. In the Nordic and Antipodean clusters the indicator of income mobility is low, which means high mobility – perhaps to a somewhat lesser extent in Sweden and the Netherlands. Low income mobility is found in Italy as well as in the United Kingdom, the United States, Portugal and Switzerland. This completes the picture of how the indicators of inequality are located on the component planes of the Kohonen map.

Figure 14. The component planes of the latent variable Inequality.



In the third figure of component planes we have the indicators of social capital and how they contribute on the map. The outstanding feature in the first panel is the very low trust in institutions that is found in Greece. Another feature is the high trust of the citizens of Iceland. An explanation for this may be that the survey was performed before the well known events in the banking sector. The Nordic cluster performs well. Generalized trust, shown in the second panel, is at a high level in Sweden and Denmark, but also in other countries of the Nordic cluster. Trust in other citizens is lowest in Portugal, Greece and France. A look at participation, presented in the following two panels, shows that passive participation is at a low level in the countries of the fourth world cluster and in Germany and France. Passive participation is at a high level in the Netherlands. In the last panel the indicator of active participation is high in Sweden and Norway, while the lowest values are found in Germany and Portugal.

In figure 16 we finally see the component planes of the indicators of wellbeing used in this study. In the two first panels we can see some interesting features. The expected life time is low in Denmark but high in Japan. When we look at satisfaction with life it is the other way around. Now Danes show the greatest life satisfaction, whereas in Japan people seem dissatisfied with life. In general it is again the Nordic cluster where satisfaction with life is clearly above the average. In the last two panels, participation in public education is high in Australia and New Zealand. Low figures can be found in Switzerland, Japan, Germany and Luxembourg. It should be noticed, however, that the private school system is not included here. In the last panel, focusing on the gross domestic product, the leaders are Luxembourg, the United States and Norway. The Mediterranean countries represent here countries with a weak economic situation.

Figures 13–16 show the component planes resulting from the iterative algorithm of the self-organizing process. Each indicator participates in the final position of the nodes or models of the Kohonen map. The values of the component planes define the group profile of a country or a cluster.

Figure 15. The component planes of the latent variable Social capital.

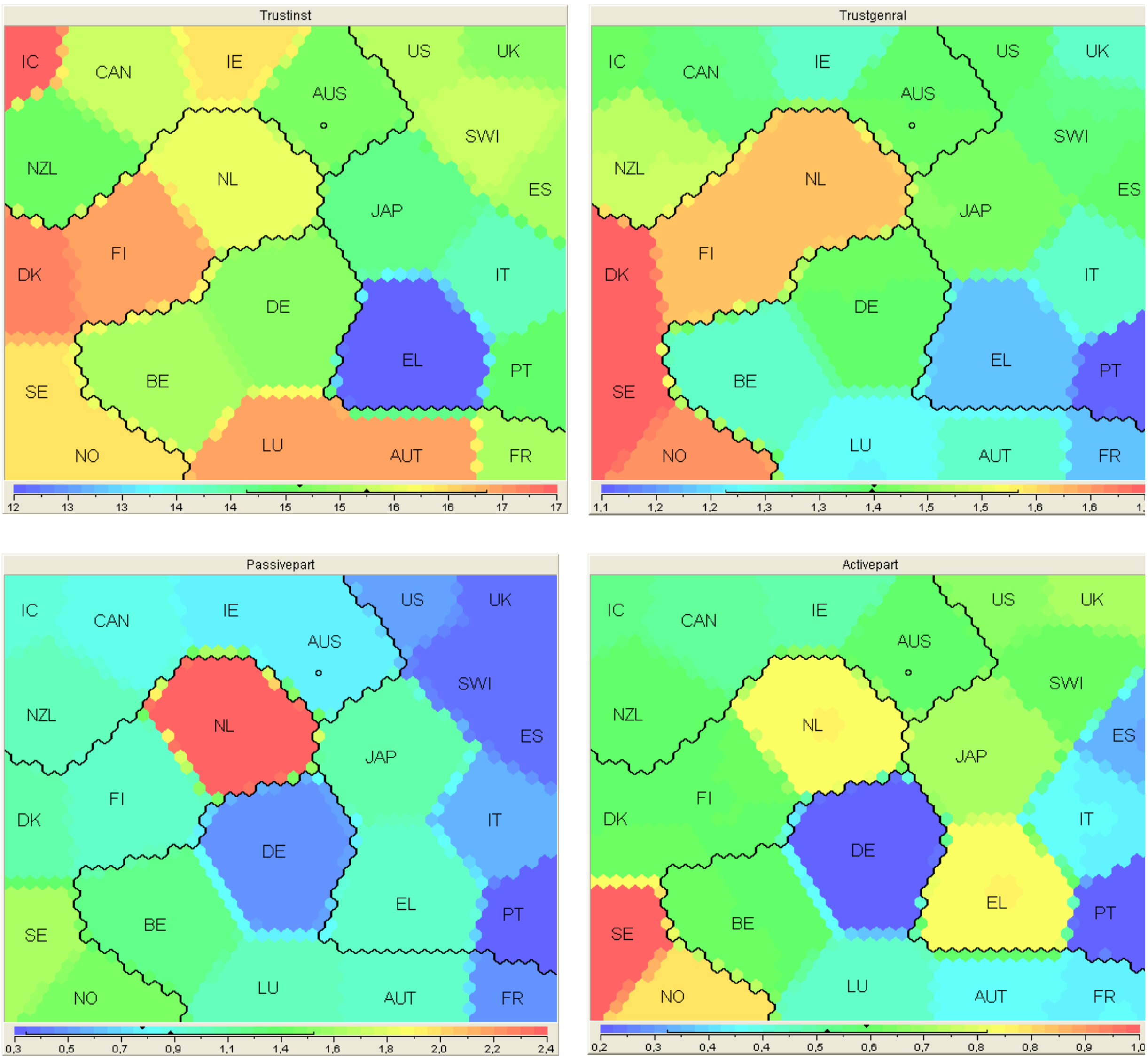
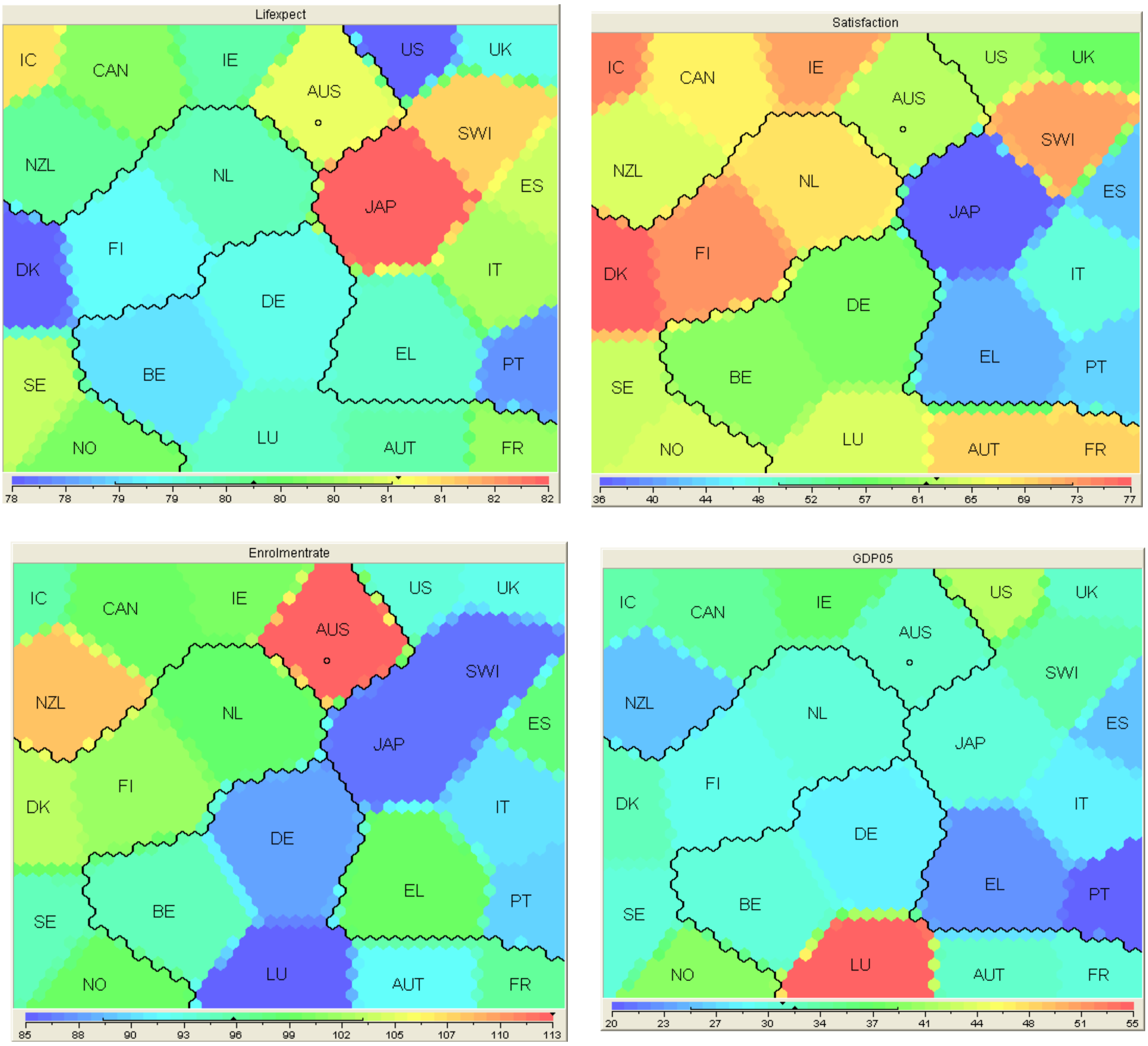


Figure 16. The component planes of the latent variable Wellbeing.



We shall next return to the question of clustering. According to Esping-Andersen "… there are three cases as the fourth regime: The Antipodes, The Mediterranean and Japan. Assuming the validity of all three claims, we will find ourselves with a total of six models for a total of 18-20 nations. The desired explanatory parsimony would be sacrificed and we might as well return to individual comparisons." (Esping-Andersen 1999, 88.) In spite of this rather strong opinion we will continue our classification by clustering the countries into six virtuous circle models. These are presented in the first panel of figure 17. Now we can see how the fourth cluster of the fourth world has split apart into three clusters. The Antipodean cluster, the Nordic cluster and the continental-conservative cluster remain unchanged. A cluster of three Mediterranean countries stands out. Spain clusters together with Japan and Switzerland. The third cluster is formed by the United Kingdom and the United States. If we have now lost some information, our relevant cluster number indicator should have fallen radically. This we do not observe. Thus we continue the experiment by increasing the number of clusters, first to 8, then to 10, 12, 14 and 16 clusters. The results are seen in the panels of figure17. As the number of clusters is increased from 6 to 8, in the Antipodean cluster New Zealand separates into a model of its own. Another major change occurs in the corporative-conservative cluster, where Germany and Belgium separate. As the number of clusters is increased from 8 to 10, Switzerland splits off into a separate model, as do Germany and Belgium. The Nordic cluster and the Mediterranean countries remain together, as they do when we increase the number of clusterings from 10 to 12. Now Iceland and Canada break off from Ireland and Australia. France becomes a distinct model type. As we now increase the number of clusters from 14 to 16 we see some interesting results. All countries form distinct models except the cluster of Nordic countries and the cluster containing the three Mediterranean countries. We might conclude that as we use the indicators of the virtuous circle as criteria for clustering these two represent a type of virtuous circle model of their own.

In order to see how tightly the countries are connected, we continue by increasing the number of clusters. In figure 18 the cluster number increases from 18 to 23. When the number of clusters is 18, the Nordic cluster divides into two groups, the first comprising Denmark and Finland and the second Sweden and Norway. The Mediterranean group remains together. In the next step Portugal splits off, and only one case remains before each country represents a model of its own. With 22 clusters, all countries are distinct except Norway and Sweden. These two countries remain together to the end, or looking at it the other way around, are the two countries that join together first. The Mediterranean countries follow the Nordic countries in the clustering process until the number of clusters is four, at which stage the Mediterranean group disbands.

Figure 17. Increasing the virtuous circle clusterings from 6 to 23.

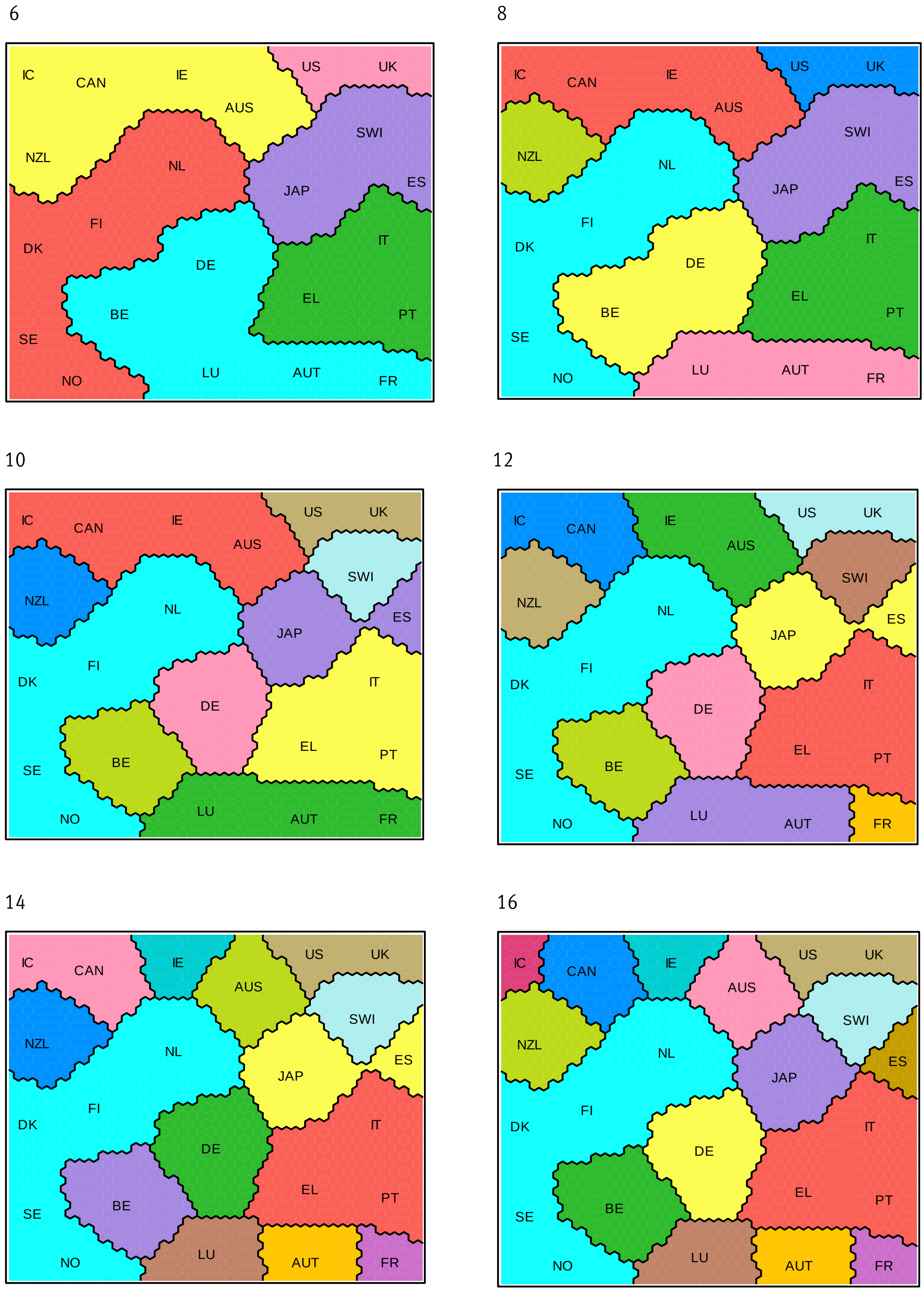
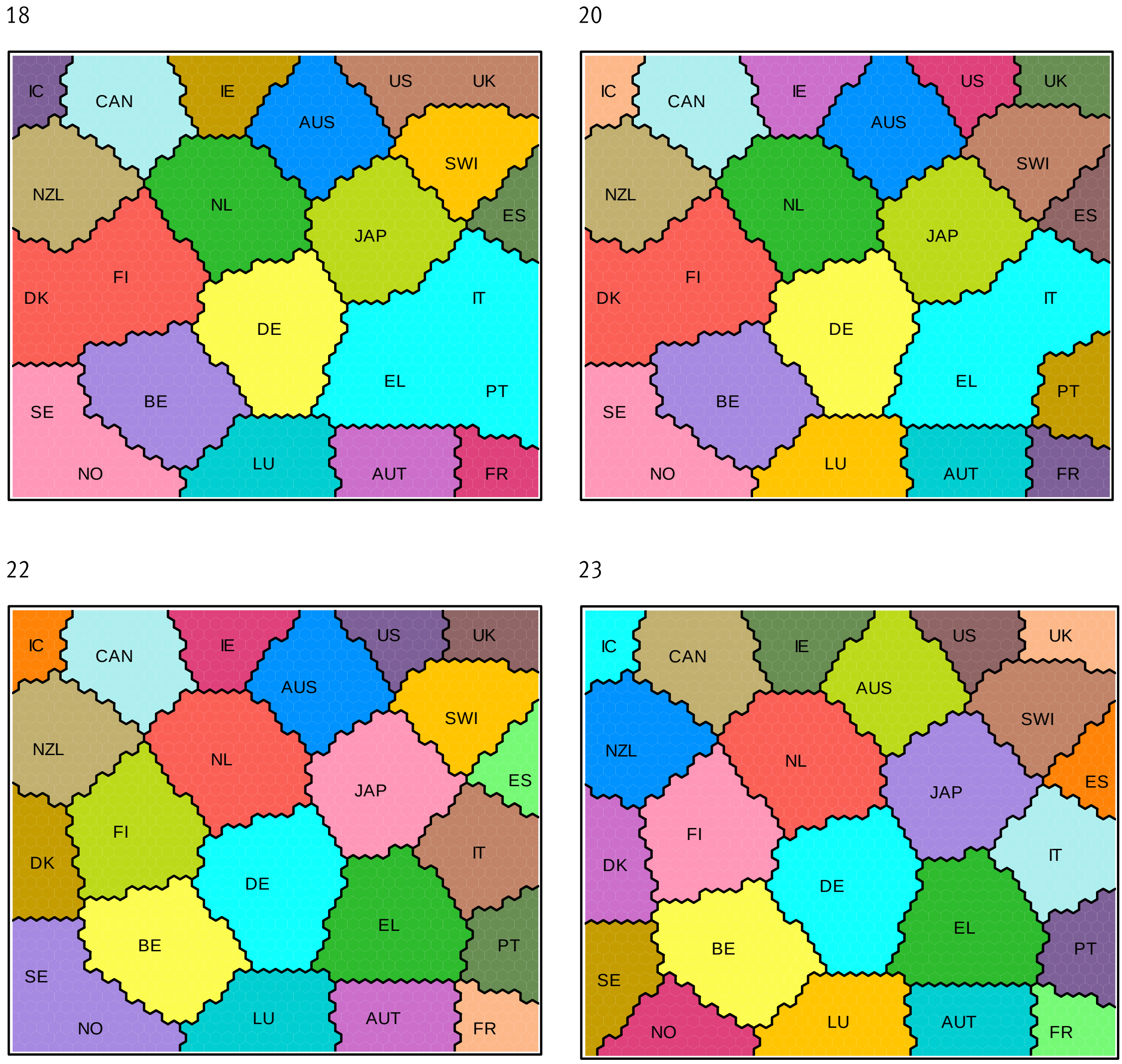


Figure 18. Increasing the virtuous circle clusterings from 18 to 23



Conclusions

We have proceeded in this study through two stages. To begin with, we defined and estimated the virtuous circle, in which welfare effort first decreases inequality. This has a positive effect on the existing social capital. Social capital has a positive effect on the wellbeing of citizens, and thus increases citizens' support towards the practised social policy. The key elements in this process are the indicators used to define the variables of the virtuous circle. Structural equation modelling was used for the estimation. It is clear that the causal relations of the virtuous circle cannot be verified by means of one cross section only. As of now the causal relations form a part of the structural equation modelling in the form of a hypothesis. Though we have carried out some causality studies that have produced evidence concerning for instance the relation between inequality and social capital, work on including data from several time points is still going on. As things stand, the results may be sufficient for the purposes of the second part of the study.

The second part of the study presented a method for clustering and classifying welfare states. The neural computation and the self organizing map were applied using the indicators of the virtuous circle. It was found that at a clustering using the virtuous circle as a criterion at first produced remarkably similar clusters to Esping-Andersen's three worlds of capitalism. The self organizing or Kohonen map allowed us to study the effect of changing the number of clusters. In that process some old clusters disbanded and new ones formed. It was found out that the Nordic countries were strongly clustered together as a well functioning virtuous circle group. The Mediterranean countries were another strongly integrated group. Looking at the group profile of these countries in appendix figure 1 we may say that this group, though coherent, is not a model group for the virtuous circle.

In the study by Bouget (ibid.) the purpose was to find countries which could act as models (attractors) for other countries. The conclusion was clear: the winner is Sweden. In our study the indicators are multidimensional and represent the virtuous circle character of the country. Two firm clusters appear, the Nordic cluster and the Mediterranean cluster. Only the countries in the former cluster can function in the manner of a virtuous circle. As we increase the number of clusters we may conclude that the winners are Sweden and Norway. The third place is shared by Finland and Denmark, with the Netherlands close behind. Social sustainability appeared to be firmly established in all of these countries.

References

Abrahamsson P. The welfare modelling business. Social Policy and Administration 1999; 33: 394–415.

Abu Sharkh M, Gough I. Global welfare regimes. A cluster analysis. Global Social Policy 2010; 10 (1): 27–58.

Adema W, Ladaique M. Net social expenditure, 2005 Edition. More comprehensive measures of social support. Paris: OECD, OECD Social, Employment and Migration Working Papers No. 29, 2005. Available at: <http//:www.oecd.org/els/workingpapers>.

Arbuckle J. Amos 7.0 User's Guide. Chicago, IL: SPSS, 2006.

Arts W, Gelissen J. Three worlds of welfare capitalism or more? A state-of-the-art report. Journal of European Social Policy 2002; 12 (2): 137–158.

Baldwin P. Can we define a European welfare state model? In: Greve B, ed. Comparative welfare systems. The Scandinavian model in a period of change. London: Macmillan, 1996: 29–44.

Boarini R, Johansson A, d'Ercole M. Alternative measures of well-being. Paris: OECD, OECD Social, Employment and Migration Working Papers, 2006. Available at: ‹http://www.oecd.org/els/workingpapers›.

Bouget D. Trends of social welfare systems. From convergence to attractiveness. An exploratory approach. REC-WP 13, 2009. Available at: ‹http://www.socialpolicy.ed.ac.uk/recwowepudiac/working papers›.

Byrne B. Structural equation modeling with AMOS. London: Lawrence Erlbaum Associates, 2001.

Castells M, Himanen P. The information society and the welfare state. Oxford: Oxford University Press, 2002.

Corak M, ed. Generational income mobility in North America and Europe. Cambridge: Cambridge University Press, 2004.

Dowling J Chin-Fang Y. Modern developments in behavioral economics. Social science perspectives on choice and decision making. London: World Scientific Publishing, 2007.

Esping-Andersen G. The three worlds of welfare capitalism. Oxford: Polity Press, 1990.

Esping-Andersen G. Social foundations of postindustrial economies. Oxford: Oxford University Press, 1999.

Garson D. Neural networks. An introductory guide for social scientists. London: Sage, 1998.

Hagfors R. Essays on the structure of social protection financing. Studies in social security and health 74. Helsinki: The Social Insurance Institution of Finland, 2004.

Hagfors R, Kangas O. (2004) Neural computation as a clustering method for comparative welfare state research. Paper presented at the second annual ESPANET conference, Comparative Methodology Stream, 9th–11th September 2004, St. Antony's College, Oxford.

Hagfors R, Kajanoja J. Hyvinvointivaltio ja sosiaalinen pääoma sosiaalisen kestävyyden perustana. Julkaisussa: Näkökulmia sosiaaliseen kestävyyteen. Helsinki: Valtioneuvoston kanslian raportteja 2, 2009a: 57–102.

Hagfors R, Kajanoja J. The Virtuous Circle. Unpublished manuscript, 2009b.

Hagfors R, Kajanoja J. Hyvän kehän hypoteesi. Teoreettista taustaa ja empiiristä arviointia. Helsinki: Kela, Nettityöpapereita (Online working papers) 11/2010. Available at: ‹http://www.kela.fi/research›.

Hausmann R, Tyson L, Zahidi S. The global gender gap report 2007. Geneva: World Economic Forum, 2007. Available at: ‹www.weforum.org›.

Hemani A, Postula A. Self-organising maps in computer aided design of electronic circuits. In: Oja E, Kaski S, eds. Kohonen maps. Amsterdam: Elsevier Science, 1999: 231–242.

Hill M, Bramley G. Analysing social policy. Oxford: Basil Blackwell, 1986.

Hofstede G. Culture's consequences. Thousand Oaks, CA: Sage, 2001.

Hudson J, Kühner S. Beyond the dependent variable problem. The methodologocal challenges of capturing productive and protective dimensions of social policy. Social Policy and Society 2010; 9 (2): 167–169.

Human Development Report 2007/2008. United Nations Development Programme. New York, NY: Palgrave Macmillan, 2007.

Inglehart R et al. World Values Surveys and European Values Surveys, 1999–2001 [electronic data]. ICPSR version. Ann Arbor, MI, Institute for Social Research [producer], 2002. Ann Arbor, MI: Inter University Consortium for Political and Social Research [deliver], 2004. Tampere: Yhteiskuntatieteellinen tietoarkisto [deliver].

Jäntti M, Bratsberg B, Røed K, Raaum O, Naylor R, Österbacka E, Björklund A, Eriksson T. American exeptionalism in a new light. A comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States. Bonn: IZA, Discussion Paper No. 1938, 2006.

Kajanoja J. Measuring social development. In: Mikko Perkiö, ed. Perspectives to global social development. Tampere: Tampere University Press, 2009: 81–95.

Kangas, O. and Palme J. Making social policy work for economic development. The Nordic experience. International Journal of Social Welfare 2009; 18: 62–72.

Kaski S, Kohonen T. Structures of welfare and poverty in the world discovered by the self-organizing map. Espoo: Helsinki University of Technology, Report A 24, 1995.

Kohonen T. Self-organized formation of topologically correct feature maps. Biological Cybernetics 1982; 43: 59–69.

Kohonen T. The self-organizing map. In: Proceedings of the IEEE 1990; 78 (9): 1464–1480.

Kohonen T. Self-organizing maps. Heidelberg: Springer-Verlag, Springer Series in Information Sciences 30, 2001.

Kvist J. Welfare reform in the Nordic Countries in the 1990s: Using fuzzy-set theory to assess conformity to ideal types. Journal of European Social Policy 1999; 9: 231–252.

Kvist J. Changing rights and obligations in unemployment compensation. Using fuzzy set theory to explore the diversity of Northern European policy developments in the 1990s. The Danish National Institute of Social Research, 2000. Available at: ‹http://cwis.kub.nl/~fsw_2/ home/worschot/ rc19/ papers/kvist.htm›. Retrieved on 2 August 2002.

Kvist J. Changing rights and obligations in unemployment insurance. In: Sigg R, Behrend C, eds. Social security in the global village. New Brunswick, NJ: Transaction Publishers, 2002: 227–245.

Luxembourg Income Study (LIS) Key Figures, {7.4.2009} Available at: ‹http://www.lisproject.org/keyfigures.htm›.

Maruyama G. Basics of structural equation modelling. Thousand Oaks, CA: Sage, 1998.

Mitchell D. Income transfers in ten welfare states. Revision of the author’s thesis (doctoral). Australian National University, SPRU University of York. Aldershot: Avebury, 1991.

Myrdal G. Rich lands and poor. The road to world prosperity. New York: Harper and Brothers, 1957.

Oja E, Kaski S, eds. Kohonen Maps. Amsterdam: Elsevier Science, 1999.

Oorschot W van, Arts W, Gelissen J. Social capital in Europe. Acta Sociologica 2006; 49 (2): 149–167.

Oorschot W van, Finsveen E. The welfare state and social capital inequality. Paper presented at the RC19 Annual Conference “The Future of Social Citizenship”, SOFI, Stockholm, 4–6 September 2008.

Putnam R. Making Democracy Work. Civic traditions in modern Italy. Princeton, NJ: Princeton University Press, 1993.

Putnam R. Bowling Alone. The collapse and revival of American community. New York, NY: Simon and Schuster, 2000.

Ragin C. The comparative method. Moving beyond qualitative and quantitative strategies. Berkeley, CA: University of California Press, 1987.

Ragin C. A qualitative comparative analysis of pension systems. In: Janoski T, Hicks A, eds. The comparative political economy of the welfare state. Cambridge: Cambridge University Press, 1994: 320–345.

Ragin C. Fuzzy-set social science. Chicago, IL: The University of Chicago Press, 2000.

Rothstein B. Is the universal welfare state a cause or an effect of social capital? The quality of government institute. Gothenburg: University of Gothenburg. QoG Working Paper Series 16, 2008.

Rothstein B, Uslaner E. All for all. Equality, corruption, and social trust. World Politics 2005; 58 (1): 41–72.

Saari J, Sailas R. Suomi – moniulotteinen menestyjä. Hyvinvointikatsaus 2006 (2).

Scruggs L, Allan J. Welfare state decommodification in eighteen OECD countries. A replication and revision. Available at: ‹http://www.sp.uconn.edu/~scruggs/wp.htm›.

Shalev M. Limits and alternatives to multiple regression in comparative political economy. Presented at the Seminar on the Methodological Problems in Comparative Research, University of Turku, Department of Social Policy, September 25–27, 2003, Turku.

Shapiro A. The merging of neural networks, fuzzy logic, and genetic algorithms. Insurance, Mathematics and Economics 2002; 31: 115–131.

Smith M. Neural networks for statistical modeling. Boston, MA: International Thomson Computer Press, 1996.

Society at a glance: OECD social indicators. 2006 edition. Paris: OECD, 2007.

Stiglitz J, Sen A, Fitoussi J-P. Report by the commission on the measurement of economic performance and social progress. 2009. Available at: ‹www.stiglitz-sen-fitoussi.fr›.

Svendsen G, Svendsen G, eds. Handbook of social capital. Cheltenham: Edward Elgar, 2009.

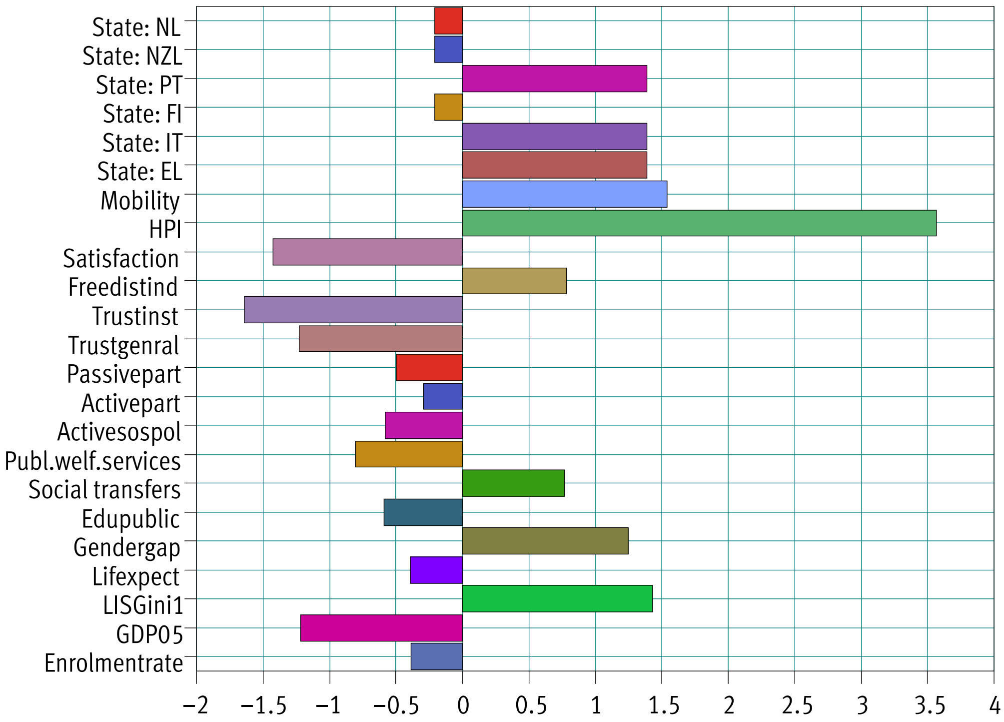
Viscovery SOMine 5.0. Available at: ‹http://www.viscovery.net›.

Vogel J, Wolf M. Index för internationella välfärdsjämförelser. Sverige i täten. Välfärd Nr 1, 2004.

Woolcock M. Social capital. The state of the notion. In: Kajanoja J, Simpura J, eds. Social capital. Global and local perspectives. Helsinki: Government Institute for Economic Research, 2000.

Appendices

Appendix figure. The group profile of the Mediterranean countries. (PT, IT, EL)



Appendix table. Indicator variables used in the study.

	N	Min	Max	Avg	Dispersion
Decommodification	18	17.90	37.30	27.4222	5.62636
Public education	17	3.50	8.30	5.5353	1.24396
Public welfare services	23	6.50	14.54	8.8852	2.06309
Income transfers	23	6.63	18.80	12.6017	3.43389
Active labour market policy	22	0.16	1.83	0.7355	0.43026
Human Poverty Index	19	6.50	29.90	12.0421	5.00648
Global Gender Gap	23	0.19	0.36	0.2731	0.04953
Gini Index	19	21.70	38.80	29.2789	5.28526
Income mobility	12	0.15	0.52	0.3100	0.13632
Restrictions on freedom of	16	20.39	46.53	34.7674	8.21853
Trust in institutions	23	12.10	16.99	15.2765	1.08420
Generalized trust	23	1.10	1.70	1.3983	0.16180
Passive participation	14	0.30	2.40	0.9143	0.56685
Active participation	14	0.20	1.00	0.5357	0.23732
Life expectancy	23	77.80	82.10	79.7348	1.11339
Satisfaction with life	23	36.00	77.00	61.1739	11.33639
Enrollment rate	23	84.70	113.00	95.8261	6.97008
GDP	23	19.53	54.69	31.8143	7.01214