



## **A network analytical approach to the study of labour market mobility**

Toubøl, Jonas; Larsen, Anton Grau; Jensen, Carsten Strøby

*Publication date:*  
2013

*Document version*  
Peer reviewed version

*Citation for published version (APA):*  
Toubøl, J., Larsen, A. G., & Jensen, C. S. (2013). *A network analytical approach to the study of labour market mobility*. Paper presented at XXXIII Sunbelt Social Networks Conference of the International Network for Social Network Analysis (INSNA), Hamborg, Germany.

# A NETWORK ANALYTICAL METHOD FOR THE STUDY OF SOCIAL MOBILITY WITH AN APPLICATION TO THE DANISH LABOUR MARKET

*By Jonas Toubøl, Anton Grau Larsen and Carsten Strøby Jensen*

**ABSTRACT:** The aim of this paper is to present a new network analytical method for analysis of social mobility between categories like occupations or industries. The method consists of two core components; the algorithm MONECA (Mobility Network Clustering Algorithm), and the intensity measure of Relative Risk (RR), which enable us to identify clusters of inter-mobile categories. We apply the method to data of the labour market mobility in Denmark 2000-2007 and demonstrate how this new method can overcome some long standing obstacles to the advance of labour market segmentation theory: Instead of the typical theory driven definition of the labour market segments, the use of social network analysis enable a data driven definition of the segments based on the direct observation of mobility between job-positions, which reveals a number of new findings.

**Keywords:** Social network analysis, cluster analysis, labour market segmentation theory, job mobility

## 1. INTRODUCTION: LABOUR MARKET MOBILITY AND SEGMENTATION

In this paper we approach the labour market as a network of industries and occupations. Individuals circulate within and between industries and occupations when they move from one job-position to another. Industries and occupations take the form of nodes in the network and the individuals' job mobility between industries and occupations are what generates the ties between the nodes of the network and thus constitute the labour market as a network. Such a conceptualization of the labour market serves mainly the purpose of mapping the mobility patterns of the labour market. Between what industries and occupations does the labour force flow freely and between which do barriers appear to disrupt the flow of labour force?

These questions have been central to the tradition of labour market segmentation theory. Labour market segmentation theory developed in the 1960's and 1970's as an alternative to human capital theory (Becker 1993; Hall 1970; Phelps Brown 1977): Not just the individual's human capital

determined the wage and opportunities at the labour market. Rather, selective barriers controlling the access to good and bad job in terms of wage and working conditions would determine the job allocation of the individual and hence the wage (Bluestone 1971; Doeringer & Piore 1971; Doeringer & Piore 1975; Gordon 1972; Reich et al. 1973). According to labour market segmentation theory, the selection mechanisms which determines who are allocated to what jobs can take a variety of forms. They can be said to be *functional*, as in the case of only people with the appropriate skills can enter certain job positions, *institutional*, as in the case of only persons with the right certificate being able to enter certain job position, and *normative*, as in the case of racial of sexist practices which exclude persons of a certain gender of race from certain job positions. According to this theory, job allocation in terms of whether an individual will end up in a good or bad job and accordingly receive a relatively high or low wage, depends more on factors like technological development, organization of the companies, interests and norms of management, and institutional regulations of the labour market, than a market driven distribution of individuals into job-position in accordance with the human capital they embody (Boje & Toft 1989; Leontaridi 1998; Loveridge & Mok 1979).

The quantitative based studies of labour market segmentation, suffer from a series of weaknesses. Of these the most crucial one is that the empirical definition of the segments: The classification of occupation or industries into various segments, did not take its starting point from the observation of the actual barriers structuring the job mobility at the labour market (Cain 1976; Leigh 1976; Wachter 1974). Instead, typically, the empirical segments are determined either in accordance with 1) a theory that designated certain job-position into certain segments (e.g. Osterman 1975; Fichtenbaum et al. 1994; Stier & Grusky 1990), or 2) measures of wage level, skill-requirements or working conditions are used to determine which segment a job-position belongs to (e.g. Boston 1990; Hudson 2007; Daw & Hardie 2012). Then, after this division of the labour market into segments, the characteristics of the workers in the segments are compared, and if for instance women are overrepresented in the group of bad jobs, it is deduced that a barrier discriminating against women exists between the good and bad jobs.

However, the segments' actual borders and properties in terms of job mobility and mobility-barriers, have not been observed and described. This is, as Cain (1976) points out, highly problematic because, we do not know whether the delimiting of the segments are correct and thus, we are not able to identify the exact location of the barriers, which we then are incapable of investigating. Hence, we are not able to draw any precise conclusions about the cause and effect of the segmentation of the labour markets.

For instance, if the defined segments actually consist of several segments with different relations to other segments, thus offering different opportunities in terms of moving between good and bad jobs, valuable information are lost, that might affect our conclusions about the labour market. Hence, traditional labour market segmentation studies have difficulties in identifying the actual segments and mapping out the borders of and relations between the segments in term of labour mobility.

In order to come up with a correct mapping of the mobility-patterns of the labour market and identify possible segments we have to observe the actual job mobility on the labour market. In section 2 we present a new network based clustering algorithm and a new distance measure we have developed in order to be able to identify the labour market segments and overcome some of the shortcomings of the traditional approach outlined above. The clustering method is data driven and explorative in nature and motivated by the need to analyze labour market mobility as well as other forms of social mobility.<sup>1</sup> Following the presentation of the method, in section 3 we apply it to micro data from the Danish labour market and go in some detail through the example of the working of the technique. In section 4 examples of the way network measures can be used to analyze the labour market segmentation are given and the findings are presented. Finally, in section 4, we discuss the prospects of these methodological developments.

## 2. A NETWORK ANALYTICAL METHOD FOR ANALYSIS OF SOCIAL MOBILITY

In dense networks where almost all nodes are connected, it is quite difficult to make sense of the structure. This will often be the case of networks describing social mobility: even though the mobility between some categories (eg. unskilled worker to senior manager) may extremely rare, they seldom will be null. In such a case, the challenge is to find the patterns of the network that is formed by the variation in the *intensity* of the connection between the nodes.

More specific, we are interested in finding groups of especially tightly connected nodes; to identify the *cohesive sub-groups* of the network. The search for such cohesive sub-groups are one of the fundamental tasks in social network analysis and some of the most basic concepts of social network analysis - *core*, *clique* and *clan* - all allude to this kind of phenomena (Frank 1995; Mokken 1979; Prell 2012:151pp; Scott 2000:108pp).

The problem with cluster analysis taking cliques or cores as its conceptual basis is that often the produced cluster solutions are not discrete but will overlap meaning that nodes will be assigned to more than one cluster (Balázs et al. 2006; Derényi et al. 2005; Farkas et al. 2007). This problem is relevant in our case, in which we want to identify clearly delimited, non-overlapping segments that can be said to make up separated categories of the labour market.

Compared to the concepts of core and clique, *clusters*, in the network analysis literature, generally refer to *non-overlapping* cohesive sub-groups (Scott 2000:126pp). Thus, cluster analysis offers a mean to analyze the graph in the sense of a set of non-overlapping cohesive sub-groups. Another

---

<sup>1</sup>The logic of the problems discussed with reference to labour market segmentation theory are not confined to the area of labour market studies but can also be said to be relevant within the broader field of social mobility including intergenerational social mobility where it has given rise to a vivid debate (Goldthorpe 2002; Grusky and Weeden 2002; Jonsson et al. 2009). Reframing the problem, it can be said to consist in identifying the categories that correctly describe the structure of the social mobility. For instance, in the case of analysis of intergenerational mobility conceptualized within class theory, should we assume 82 micro classes as Grusky et al. does (Grusky & Galescu 2005; Jonsson et al. 2009), or assume 9 classes as Goldthorpe and associates do (Goldthorpe 2000; Erikson et al. 1979), and what are the exact borders and properties of the various class structures?

advantage of cluster analysis over sociometric concepts as core and clique is that cluster analysis are much better suited for handling weighted networks (Snoke & Yang 2008:80-82). Thus, in the case of cluster analysis we can include the information represented by the variation of the weights of the ties in the analysis.

To sum up, when we set out to identify the cohesive sub-groups of a network, cluster analysis is better suited for the task if we want to 1) identify non-overlapping cohesive sub-groups and 2) if our network are weighted and we wish to take the information represented by the weights of the ties into consideration.

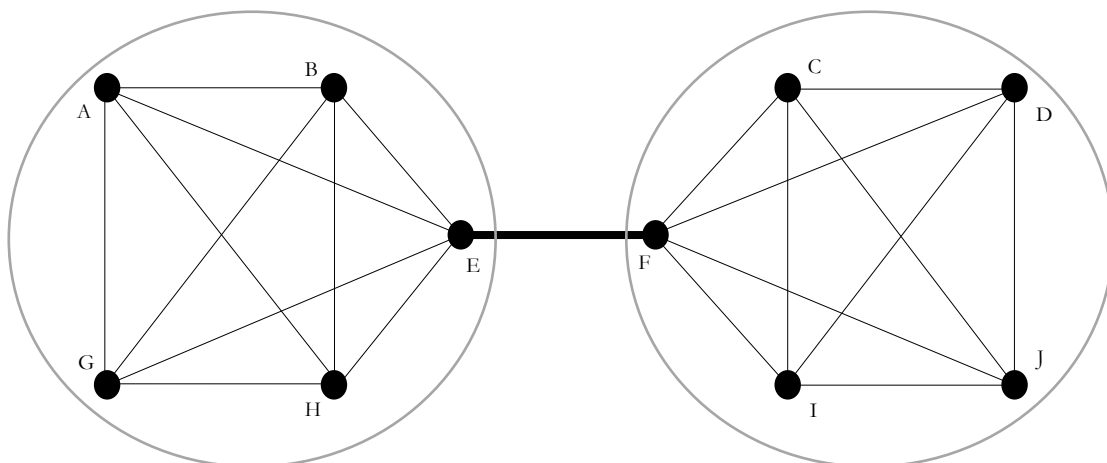
### 2.1 The shortcomings of traditional cluster analysis

Traditional cluster analysis suffers from one major weakness seen from the network analytical perspective: it ignores the sociometric and graph theoretical properties of social networks. This might seem a paradox, as all the advantages of cluster analysis listed above depends on breaking with a purely sociometric approach to the task of identifying cohesive sub-groups in the graph. Nonetheless, the cluster solution produced by most cluster analysis techniques might, when compared to the network they are supposed to represent, be quite unacceptable, because they separate two nodes like node E and F in figure 1, that are intensely connected.

In short, the problem can be said to consist in the fact that, traditional cluster analysis take the *similarity* (Everitt 2011:1pp; Jain et al. 1999; Sokal 1985), *likeness* (Williams 1971) or *proximity* (Zahn 1971; Ozawa 1986) of the cases as its principle of clustering, whereas the network analyst's principle is that of *connectedness* or *linkage* (Guha et al. 2000) of the cases, represented by the edges of a network graph. An example might help clarifying what this means in practice.

For instance, if our two intensely connected nodes completely differ with regard to their

FIGURE 1: SIMILARITY VERSUS CONNECTION



relations to all other nodes, like nodes E and F in figure 1, this might cause them to end up in two different clusters. Such a solution is indicated by the two grey circles in figure 1 demarcating the two clusters  $|ABEGH|$  and  $|CDFIJ|$  which represent the solution of a cluster analysis which focus on similarity. Due to nodes A, B, E, G and H all being connected with each other in the fashion of a clique, and their lack of connection with nodes C, D, F, I and J, (with the exception of node E's connection with node F), nodes A, B, E, G, and H are relatively similar and relatively dissimilar to nodes C, D, F, I and J, thus constituting a cluster. The remaining nodes, C, D, F, I, and J, however, are relatively similar and make up a cluster of their own. But in this case, following the principle of similarity/dissimilarity when forming the clusters, the two most intensely connected nodes, E and F are separated into two distinct clusters. Hence, this cluster analysis ignores their intense relation because they with regard to all other nodes are completely dissimilar. Seen from a network analytical perspective, a solution separating the most intensely connected nodes will often be unsatisfying.

To sum up, in alignment with most traditional cluster analysis we wish to classify the cases by partitioning them into non-overlapping (Bailey 1975) or “hard” (Jain et al. 1999) clusters, but we do not wish to classify by a principle of similarity of the cases' patterns of features or attributes (Omran et al. 2007: 584; Sandri 1969), but by a principle of *connection*.

It could be argued that the principle of *connectivity* is a variant of the well-known cluster analytical concepts of *distance* (Jain et al. 1999) or *proximity* (Zahn 1971), and the logic of the below proposed algorithm is indeed very similar to that of the nearest neighbor clustering principle (Lu & Fu 1978). Nonetheless, we think it would be wrong to conceptualize the relation between the nodes of a network as *distances* (distance is fundamentally just a measure of similarity (Omran et al. 2007:584)). The reason is that the conceptualization of the relations between nodes in a network as distances alludes to a perception of the network as points in a *space* (Cormack 1971) in which the relation between all points/nodes can be measured as a distance. This logic implies that potentially *all* points – no matter whether they are connected or not in the network – can be merged. In contrast to an approach based on a distance measure, the principle of connectivity implies that if two nodes are not connected, they cannot belong to the same cluster, no matter how similar they are in all other aspects. Therefore, by its adherence to the principle of connectivity the MONECA algorithm stay within the logic of network and graph theory. Hopefully, this will become clear, when we now turn to the specifics of the MONECA algorithm.

## 2.2 MONECA

MONECA is designed to identify discrete clusters of interconnected nodes in dense networks. The problem at hand can be perceived by inspecting the network in figure 2. Figure 2 depicts a rather dense weighted network's graph as well as its adjacency table. The maximal cliques of the graph, are  $|ABC|$  and  $|BCD|$ , which overlap each other in the case of 2 out of 3 nodes each. In this case, where identifying discrete clusters by a simple analysis of the cliques is a futile endeavor, MONECA

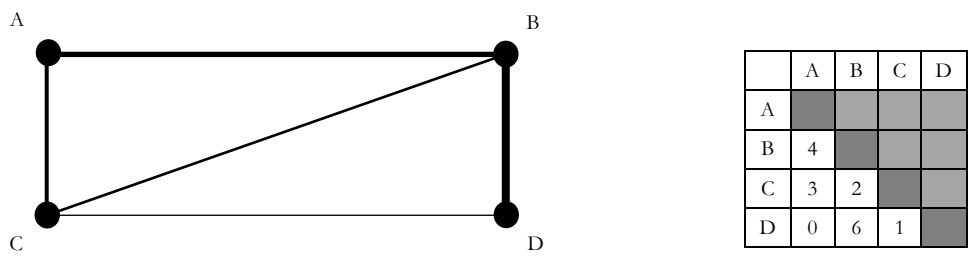
is a way to make sense of the discrete clustering structure of such a network. In the following, we start by explaining the logic of MONECA, then we give an example.

The logic of the algorithm is closely associated with the concept of the clique, and the task of the algorithm can be said to be to decide to which clique to allocate the nodes that are in the overlapping areas of the maximal clique solution in figure 2. Nodes B and C are members of both maximal cliques. However, the network is weighted, and the weights expressing the intensity of the relations of the nodes, is the information that enable the algorithm to decide which clique nodes B and C should belong to. The procedure to make this decision, however, takes a different starting point, than the maximal cliques of a network.

The algorithm is agglomerative, starting from the most disaggregated level, considering the connections of the single nodes. First step is to pair together the two most intensely connected nodes which then form a cluster. Subsequently, the connection between these two nodes is not considered. It proceeds in step two to pair together the two nodes which then are the most intensely connected in the same manner as in step one. If two nodes that already are members of clusters form the most intense connection, all the nodes in the respective clusters of the two nodes, are joined together forming one big cluster. However, a set of nodes can only be considered a cluster if they also form a clique. Therefore, in the case of pairing together two clusters, this is only possible if all the nodes of the two clusters form a clique. This criteria provides the stop rule for when no more single or sets of nodes should be paired together forming new clusters. Otherwise, the cluster solution would simply be the components of the network.

An example of how the algorithm works should clarify the procedure: In the case of figure 2, B and D are the most intensely connected nodes which can be seen from the width of the ties representing the intensity of the relation. Then, B and D are paired together, thus constituting a preliminary cluster,  $|BD|$ . The second most intense connection is that of A and B. However,  $|BD|$  are already a cluster so MONECA asks whether A can be paired with both B *and* D, forming the cluster of  $|ABD|$ . In order to settle this question, MONECA must determine whether  $|ABD|$  constitute a clique. In this case  $|ABD|$  is not a clique, because nodes A and D are not connected. MONECA then goes on and considers the third strongest connection, which is A and C. Neither A

FIGURE 2: EXEMPLIFYING THE MONECA ALGORITHM



or C are members of a preliminary cluster, thus, they can be paired without further ado. The fourth strongest connection is BC. However, B has already been paired with D, and C has been paired with A. The, MONECA asks whether  $|ABCD|$  constitutes a clique. The answer is *no*, because A and D are not connected. Hence,  $|AC|$  and  $|BD|$  cannot be paired. The same is the case with regard to the fifth connection,  $|CD|$ . As result, the cluster solution produced by the algorithm is  $|AC|$  and  $|BD|$ , and none of the maximal cliques,  $|ABC|$ ,  $|BCD|$ .

In this example the network was not directed, and MONECA will always consider networks as such. Therefore, in the case of a directed network, preliminary manipulation of the adjacency matrix is necessary.

In the following analysis of the segmentation of the labour market, we will demonstrate how weighted and directed networks can be handled in the case of a affiliation by affiliation network matrix. Before that, we introduce the second methodological novelty of this paper: Relative Risk as a network weight measure.

### 2.3 *Relative Risk as network weight measure*

The measure of Relative Risk (RR) is the ratio of proportions:  $RR = \pi_1/\pi_2$  (Agresti & Finlay 1997:271). As the name Relative Risk suggests, RR tells us something about the risk (or chance) of event A happening relative to event B. This measure is useful when we wish to assess the relative strength of the relation between two categories.

For instance, consider table 1 which is a section of job mobility table representing 10 out of 111 industries. The cells counts are the individuals moving from the industries of the row to the industries of the column. RR can be used to determine the relative likeliness of a given number of observed job shifts from one industry to another. If RR is 1, the number of job shifts is the likely one assuming independence between the row and column categories. If RR is more than 1 the number of job shifts from one industry to another is more than likely and if RR is less than 1 the number of job shifts is less than likely.

To obtain the RR of the  $i$ th row and the  $j$ th column we divide the observed cell count with the expected cell count:

$$RR_{ij} = \text{Observed}_{ij} / \text{Expected}_{ij}$$

The calculation of the expected cell value is done by the formula:

$$\text{Expected cell value} = \text{Row total} \times \text{Column total} / \text{Grand total}.$$

Then, the RR expresses the ratio of the proportion of individuals from industry  $i$  who is mobile to industry  $j$  ( $\pi_1$ ) to the proportion of *all* individuals including those of industry  $i$  who is mobile to industry  $j$  ( $\pi_2$ ).



Table 2 is the RR of table 1. If the value of a given cell is 1 or greater the proportion of individuals who are job mobile from industry  $i$  to industry  $j$  is equal to or greater than what we would expect assuming independence between the variables (randomness). If the value is less than 1 the proportion of mobile individual from industry  $i$  to  $j$  is less than expected.

The straight forward translation of table 2 into a network would be that if the RR, is 1 or greater, there is a directed tie between the nodes and if the RR is less than 1 the nodes are not

TABLE 1. COUNT OF JOB-TO-JOB MOBILITY BETWEEN 10 OF 110 INDUSTRIES, 2001-2007

	Landbrug	Gartnerier	Mask.stat	Skovbrug	Fiskeri	Olie/gas	Grus/ler	Slagterier	Mejerier	Bagerier	Total
Landbrug	25545	325	1622	252	41	9	46	217	162	100	44,060
Gartnerier	286	3670	163	37	3	1	5	53	27	74	9,916
Mask.stat	951	71	3403	83	7	2	34	37	21	7	9,662
Skovbrug	146	22	105	1082	4	1	1	14	7	5	2,560
Fiskeri	61	4	13	5	2833	17	8	39	5	3	4301
Olie/gas	3	0	0	0	2	622	0	2	0	1	1302
Grus/ler	36	1	16	3	2	1	486	7	2	3	1247
Slagterier	203	27	57	11	29	12	7	7762	152	46	16757
Mejerier	89	3	23	9	2	1	6	157	3289	15	7839
Bagerier	97	30	19	6	2	2	1	62	48	6227	18146
Total	37019	6439	9978	2469	3591	2131	1284	14916	7607	12071	2490258

TABLE 2. RELATIVE RISK OF JOB-TO-JOB MOBILITY BETWEEN 10 OF 110 INDUSTRIES, 2001-2007

	Landbrug	Gartnerier	Mask.stat	Skovbrug	Fiskeri	Olie/gas	Grus/ler	Slagterier	Mejerier	Bagerier	Total
Landbrug	39.001	2.853	9.188	5.769	0.645	0.239	2.025	0.822	1.204	0.468	44060
Gartnerier	1.940	143.138	4.103	3.763	0.210	0.118	0.978	0.892	0.891	1.540	9916
Mask.stat	6.621	2.842	87.901	8.664	0.502	0.242	6.825	0.639	0.712	0.149	9662
Skovbrug	3.836	3.324	10.236	426.295	1.084	0.456	0.758	0.913	0.895	0.403	2560
Fiskeri	0.954	0.360	0.754	1.173	456.779	4.619	3.607	1.514	0.381	0.144	4301
Olie/gas	0.155	0.000	0.000	0.000	1.065	558.265	0.000	0.256	0.000	0.158	1302
Grus/ler	1.942	0.310	3.202	2.426	1.112	0.937	755.874	0.937	0.525	0.496	1247
Slagterier	0.815	0.623	0.849	0.662	1.200	0.837	0.810	77.334	2.969	0.566	16757
Mejerier	0.764	0.148	0.732	1.158	0.177	0.149	1.484	2.918	137.352	0.395	7839
Bagerier	0.360	0.639	0.261	0.333	0.076	0.129	0.107	0.570	0.866	70.794	18146
Total	37019	6439	9978	2469	3591	2131	1284	14916	7607	12071	2490258

Source: Statistics Denmark

connected. However, depending on the subject matter, choosing 1 as the cut point may be more or less arbitrary and should be motivated by theoretical considerations. In the case of mobility at a labour *market* it makes theoretical sense to choose 1 as cut point, as a RR level of 1 is equivalent to perfect market conditions; under the assumption of randomness, a RR of 1 indicates that the expectation of random allocation has been met, and that no barriers influence the mobility of labour at the market. Then, if the RR of being job-mobile from industry A to industry B and vice versa is 1 or more, we can say that the relation of industries A and B meets the expectations of a market, and that they are part of the same labour market.

Thus, RR enables us to make sense of the relationship between rather abstract categories such as occupations, industries, social classes etc., in a way that can be understood in terms of social network analysis. A further advantage is that the calculation of the expected values can be subjected to statistical testing of whether the cell count of table A is random or not using standard  $\chi^2$ -test of the single cells' counts.

In the following, we apply the RR-weight measure as well as the MONECA algorithm to data describing the job mobility of the Danish labour market.

### 3. NETWORK ANALYSIS OF THE DANISH LABOUR MARKET'S MOBILITY STRUCTURE

The subsequent application of the combination of the RR measure and the MONECA algorithm to data of labour market mobility in Denmark 2001-2007, serves to demonstrate how a network analytical approach can help solve some longstanding problems within labour market studies, which were outlined in section 2 of this paper.

First, we will shortly summarize some of the central findings of the existing research concerning the Danish labour market with regard to mobility and segmentation. This serves to provide the reader a sense of the characteristics of the Danish labour market. Second, data is presented, and the transformation of microdata into relational network data is described. Third, the network of mobility between industries measured by the RR is presented, and the application of the MONECA algorithm in this specific case is explained. Fourth, the results of the cluster analysis is presented and various numerical network theoretical measures describing the properties of the network in its entirety, the segments in them self as well as the relationship between the segments are explained. Finally, we discuss the possible future applications of this network analytical method of social mobility.

#### *3.1 The Danish labour market: Flexicurity and powerful institutions*

The Danish labour market is one of the most flexible in the EU, famed for its flexicurity arrangements providing the employees with high level of social protection and unemployment benefits (security) at the cost of job-protection making it relatively easy and cheap for the employers to hire and fire and thus adjust the share of labour in the production to the demand of

the market (flexibility) (Jensen 2008, 2011, 2012; Jørgensen & Madsen 2007; Madsen 2006). That the Danish labour market is very flexible with a high degree of job-to-job mobility is confirmed by statistical records. In 2005 the share of employed persons who experienced a change of job in Denmark was the second highest in the EU, 11.5 % which should be compared to The EU average of 8.8 % (Danish Technological Institute 2008:21). In 2006 the average job tenure was the fourth lowest in the EU (DK: ca. 8.3 years; EU-average: ca. 10.5 years) (Danish Technological Institute 2008:22) and in 2005, the average job duration in Denmark was the shortest in the EU (DK: ca. 4.8 years; EU average: ca. 8.3 years) (Danish Technological Institute 2008:27).

These numbers all bear witness to the fact that the mobility is high. However, we cannot be sure that the high degree of job-to-job mobility also means that the degree of between industry and occupation mobility is high. The mobility might be within occupation or within industry, which would mean that the Danish labour market is very fragmented into small sub-markets. The aim of the subsequent analysis is to delve deeper into this kind of questions.

The flexicurity of the Danish labour market has been developed by strong labour market institutions. In 2007, 67 % of the labour force was organized in trade unions (own calculation), while around 55 % of the employees work in companies organized in an employers' association (Jensen 2012). The total collective bargaining coverage was around 80 % (Due et al. 2010:81). Together with the state, trade unions and employers associations form an IR-system which has been termed 'the Danish Model' (Due et al. 1994). This IR-system is special because rules regulating wages, working hours, working conditions etc. to a large degree are negotiated by the trade unions and employers and not the state. Also when it comes to labour market related areas regulated by the state, e.g. education, the unions and employers associations have a substantial influence.

All in all, the Danish labour market is highly regulated and institutionalized which most likely has consequences for the segmentation. According to Boje, who undertook a major analysis of the Danish labour market segmentation in the 1980's, the high level of institutionalization, flexibility, active labour market policies and relatively egalitarian and high level of education makes Denmark a special case in which we should not expect to find a few large segments as postulated by the classical segmentation theories discussed in section 2 (Boje 1985, 1986, 1990), but rather a number of sub-markets (Boje & Toft 1989). This conclusion is supported by his empirical findings (Boje 1987). Even though the mobility may be intra-categorical, still the most plausible expectation given the high level of job-to-job mobility is that the network will be rather dense and as such, the challenge is to identify the structures and patterns of the relative differences in the intensity of relations between the industries.

### *3.2 Data*

The data used comprises the entire Danish labour market from 2001-2007 and was collected by Statistics Denmark. We have information concerning the individual's labour market status with

regard to industry (NACE) and occupation (ISCO) and shifts of job-position. We are thus able to register whether they changed job or not and from and to which industry or occupation they moved year by year. For a person who was active on the labour market, we have registered 6 transitions<sup>2</sup>. For each transition, we know the industry and occupation they were in at the beginning and at the end of this transition. This job position might not have changed and still be the exact same job position both years, or it might have changed to another job position with the same or a different combination of occupation and industry.

Because the primary aim of this paper is to present a new method, and for the sake of simplicity, we are going to delimit the analytical scope to only consider mobility between industries in the following exemplification of the method. This is not to say that occupation is a less important variable, and if the aim was to make a thorough analysis of the social mobility at the Danish labour market occupation would certainly receive at least the same amount of attention as industry. However, the choice of industry is motivated by industry being the preferred categorization of the labour market of the labour market segmentation tradition.

Transforming the basic entity of observation from being individuals to transitions gives us a dataset with 16,227,672 industrial transitions of which 2,490,258 (15.3 %) was genuine job changes. The reason for the job-to-job mobility rate, 15.3 %, being so high as compared to the 11.5 % registered by the Danish Technological Institute in 2005 (2008:21) is due to transition in and out of the labour market, which do not count as job-to-job mobility, (e.g. from unemployment to employment, from employment to retirement, etc.) are not included in the dataset.

The final transformation of data is to make an adjacency matrices consisting of 111 by 111 industrial categories (total count = 2,490,258) with the cells counting the job shifts from the row categories to the column categories. Note that a job shift may be from and to the same category meaning that the diagonals of the matrices are not hollow (actually far from as can be seen from table T). This matrix form the basis of the networks of the relationship between the industrial categories, describing the mobility of the Danish labour market.

### *3.3 Application of Relative Risk measure and the MONECA clustering algorithm*

Variation in the number of job-mobile employees of the categories of the matrix is outspoken as can be seen from table 3, from which we can read some standard statistical descriptive measures regarding the distribution of the outgoing transition (the row totals). Varying from 392 to 163,051, the sheer variation in size makes it necessary to employ a measure of the relation between row and column categories that can take the differences in numeric size between the sending and receiving category into consideration: What might be a tiny fraction of a large category may be more than the total of a small category. Relative Risk enables us to take the relational nature of the exchange of labour force between occupations and industries into consideration.

---

<sup>2</sup> The six transition being 2001-2002, 2002-2003, 2003-2004, 2004-2005, 2005-2006, 2006-2007.

TABLE 3: DESCRIPTIVES FOR THE  
INDUSTRY MOBILITY MATRIX

<i>Mean</i>	22,434.76
<i>Median</i>	14,759.00
<i>SD</i>	24,590.29
<i>Minimum</i>	392.00
<i>Maximum</i>	163,051.00
<i>Diagonal's share</i>	51.10 %
<i>Cell-counts of zero</i>	4.24 %

We calculate the RR of the cells as outlined in section 2.3. 1787 number of cells in the matrix has a RR-value of 1 or more meaning that the network has 1787 of edges. If we ignore the loops, the diagonal, the number of edges are 1677. As indicated by the large share of the mobility being in the diagonal (see table 3), all categories of the two matrices are self-referential meaning that more than expected of the job-mobile shift to a job within the industry or occupation of their prior job-position. To students of the labour market mobility as well as common sense, this is no surprise.

An important restriction has to do with statistical nature of the RR-measure. RR is as mentioned in section 2.3 closely associated with the  $\chi^2$ -test values. Just as in the case of the  $\chi^2$ -test, the RR gets unreliable if the count is too small. The minimal industrial category had only a count of 392 transitions to describe its relation to 111 occupational categories. Each of the 15 transitions thus represent 0.283 % of the outgoing job-mobility of this category. This means if 1 or more of the transitions is to a category making up 0.283 % or less of the ingoing transition (which is quite likely as the smallest industry measured by its ingoing mobility makes up only 0.0001 % of the ingoing transitions), an edge should be drawn. Drawing an edge based on 1 or a few number of observation is not sound.

In order to counter drawing ‘false’ edges, we simply exclude all categories with less than 500 outgoing or ingoing transitions by replacing the values of their rows and columns in the matrix with the RR-values with zeros. We are now able to draw the initial network graph which can be seen in figure 3, depicting the industrial labour market mobility structure at the first and most disaggregated level with 111 nodes.

The basic layout is generated with the Fruchterman-Reingold layout algorithm but we have manipulated it in accordance with the clusters identified by MONECA which we shall turn to shortly. As indicated by the arrowheads, the network is directed, the arrowheads indicating the direction of the flow of labour pointing to the receiving node. The size of the nodes indicates the actual size of the category and the color indicates the percentage of job-transitions being internal to the category. The darker the color of the node, the higher the percentage of the job-transitions being internal.

In some cases the edges has arrowheads pointing towards both nodes indicating that the labour force flows in both direction, a reciprocal relation. In those cases the two categories can be said to be part of the same sub-labour market. If we can identify such relatively separated sub-labour markets we can talk about a segmentation of the labour market. The share of cell counts of zero is

4.24 % (see table 3). This number concerns the raw count of transition and indicates that truly discrete clusters with 0 exchange of labour will be extremely rare if not non-existing. Even when the cut point of a RR of 1 is used, the density of the network is quite high, 0.539 (See table 4).

In figure 4 the result of the MONECA algorithm is visible, the second level as compared to the first and most disaggregated level in figures P1. Before the cluster analysis has been performed, the directed network matrix has been transformed into an undirected network matrix in which only the reciprocal relations are maintained. MONECA has identified the discrete clusters meeting the criterias described in section 2.2, meaning they are cliques. We have marked the clusters identified by MONECA by drawing a line around them. Saying that each of the identified clusters (or sub-markets) forms a new category, we have reduced the number of categories from 111 to 41. The color of the line around the clusters indicated the total share of mobility being internal to the cluster in the same way the color of the individual nodes denote the internal mobility of the original categories. This information tells us about the relative discreteness of the clusters.

Now, we can form a new 41 x 41 matrix using the categories identified by MONECA and seek for clusters in this network with MONECA. The result is shown in figure 5. The new clusters are marked by lines encircling existing clusters of nodes and single nodes, forming new, bigger clusters. These level three clusters are more loose than the level two ones and does not fulfill the definition of a clique; otherwise, the clusters would have been formed the first time we analyzed the level one network for clusters. We have now reduced the number of categories to 28.

We reiterate the procedure once more, reducing the number of categories to 26, the new level four clusters being even more loose than the level three clusters. The results are depicted in figure 6. This procedure can be reiterated a number of time but it will eventually not be possible to merge more clusters and/or nodes. The reason for this is, that the matrix is not hollow, so for each level a larger share of the total mobility will be in the diagonal and eventually their will not be any reciprocal inter-categorical relations of mobility left, meaning that no more categories can be clustered. In our case, this situation occurs at the fourth level, and figure 6 depicts the segmentation of the Danish labour market by industry, identified by our method. 26 number of industry dependent segments have been identified.

#### 4. RESULTS: THE SEGMENTATION OF THE DANISH LABOUR MARKET

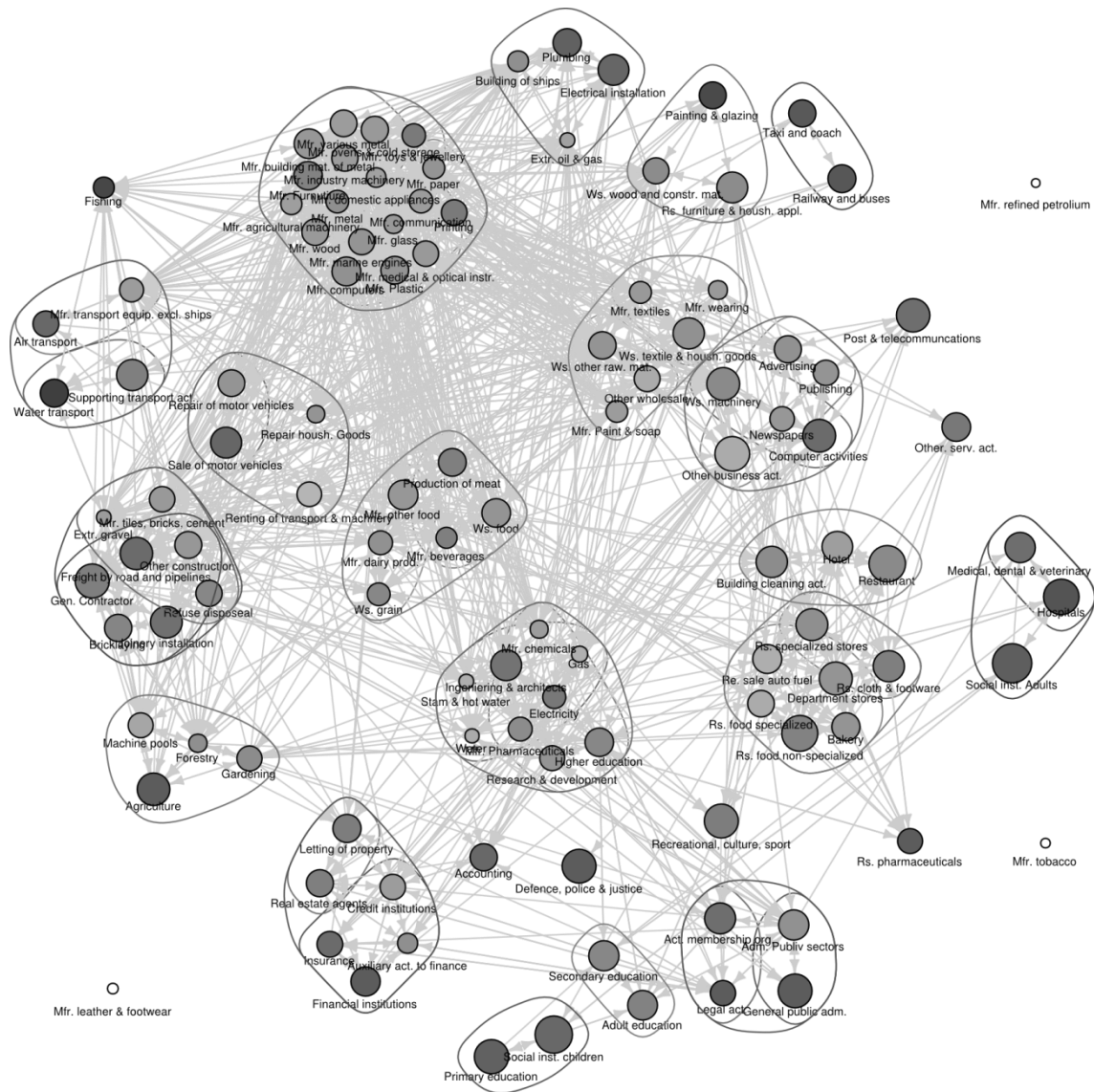
After this rather detailed examination of the technical steps of the method, we shall now turn to the more substantial description of the segmentation of the Danish labour market. We start out by presenting some measures, which express the overall structuration of the job-to-job mobility at the Danish labour market. These measures all in some sense express the deviation from total randomness in the job-to-job mobility. However, they all are good measures if we wanted to compare to different networks of labour market mobility. This could be comparison between different countries, but also between segments. This will be done in the second sub-section, in which we describe the identified segments and compare them.







FIGURE 5. LEVEL 3 NETWORK OF THE LABOUR MARKET MOBILITY BY INDUSTRIES





#### 4.1 The overall structuration of the mobility at the Danish labour market

Table 4 summarizes some general descriptive measures for the clustering procedure. The mobility explained by the industry clusters increases from 51.1 % at the first level with 111 clusters to 62.5 % at the fourth level with only 26 clusters. The increase in explained mobility is 11.4 %-point which is equal to 23.3 % percent of the between categories mobility at level 1 figures as within categories mobility at level 4. This tells us a lot about the Danish labour market: At level 1, 51.1 % of the job-to-job mobility happens within the 111 possible combinations of industry clusters that are from and to the same cluster. However, the total number of possible industry cluster combinations at level 1 are 12,321. Thus 51.1 % of the mobile labour moves within 111 out of 12,321 or 0.9 % of the theoretical possible paths. At the second level 57.2 % moves within 40 industry clusters out of 1600 combinations, or 2.5 % and at the third level 60.7 % moves within 3.5 % of the combinations and at the fourth level 62.5 % are mobile within 3.8 % of the possible combinations. Whether these numbers are high or low is a relative question, but what they bear witness to the, probably not very surprising fact, that labour market mobility is not random and that when we get a new job, the likelihood of staying within the same industry and maybe company as we came from, is very high.

Another measure that expresses the degree of the structuration of the labour market is how much of the total mobility which is within the edges of the networks, i.e. the cell counts that are greater than expected assuming randomness. 69.9 % of the mobility is nested in the edges. Under absolute randomness this number would have been 0 %. Now, to settle whether 69.9 % is a high or low level we will need a comparative study.

However, it is clear that the labour market is very structured in the sense that it is more likely to shift job to certain industries than others given the industry of origin. This is hardly a surprising fact, but nonetheless a fundamental assumption underlying segmentation theory. What is more important, it is a simple measure that enables comparison between different countries labour markets.

Of these 69.9 %, 51.1 %-points are within industry mobility at level 1. This implies that 18.8 % of the mobility is in the 1677 edges or the cells with an RR of more than 1 (minus the diagonal) which primarily guides the clustering algorithm because they express a more than likely pattern of mobility. The remaining 30.1 % of the mobility are located in the 10.533 cells with an RR of 1 or less, which as representing less than likely patterns of mobility. These also guide the clustering algorithm because they figure as disconnections in the network. Therefore, it would be wrong to

TABLE 4. GENERAL DESCRIPTIVES FOR THE CLUSTERING PROCEDURE

Level	1. Level	2. Level	3. Level	4. Level
Number of clusters	111	41	28	26
Change in number of clusters	-	-68	-15	-2
Within clusters mobility	51.1%	57.2%	60.7%	62.5%
Change in within clusters mobility	-	6.1%	3.5%	1.8%
Within edge - within clusters mobility	18.8%	12.7%	9.2%	7.4%
Density	0.272	0.157	0.048	0.019
Mean weight (RR)	2.5	1.3	1.1	1.2

say that 30.1 % of the mobility or information has been deleted and are not taken into consideration in the cluster analysis.

The concentration of mobility on the respectively more than likely and less than likely cells are however a telling measure in a comparative perspective. Just as the fact, that 51.1 % of the mobility are concentrated on the 111 cells in the diagonal bears witness to the degree of structuration, so does the concentration of mobility respectively on the more than likely and less than likely cells. This can be expressed as the mean share of mobility represented by each cell. For the 111 diagonal cells the number is 0.46036 (51.1 / 111). For the ‘more than likely’-cells the number is 0.01121 % and for the ‘less than likely’-cells the number is 0.00286 %. This means that the cells with the more than likely mobility patterns on average represent 3.920 (0.01121 / 0.00286) times more mobility than the ‘less than likely’-cells. This comparison can also be done to the overall average cell. The average cell represents 0.00812 % mobility (100 / 12321), which is less than the average ‘more than likely’-cell, which represents 1.381 (0.01121 / 0.00812) times more mobility.

This tells us that despite the heavy concentration of mobility on the diagonal, the remaining ‘more than likely’-cells, the edges of the networks, still represents relatively much mobility and that the invisible mobility in the ‘less than likely’ cells relatively are spread very thin and in general do not represent significant mobility patterns. In addition, even when we disregard the within-industry mobility in the diagonal, the between industry are concentrated on a Furthermore, these numbers can be used if we were comparing different labour markets and wanted to know the relative concentration of mobility on the more- and less than likely mobility patterns as well as the diagonal.

The measures of density and the mean edge-weight are also useful measures when we wish to assess the overall structuration of the labour market. The density of 0.272 tells us that on average, each node is connected to more than ¼ of the total number of nodes, a mean number of degrees of 30.2 (see table 5). The mean weight measure of 2.5 in table 5 tells us that the average edge of the network represent a pattern of mobility which is 2.5 times more likely to happen that expected under the assumption of randomness. The development in these measures as the cluster analysis aggregates the network to higher levels, gives us a clue about how successful the algorithm is in clustering together the most dense mobility patterns. In the level 4 network, the mean degree is only 0.5 meaning that a total of 13 edges among the 26 nodes are left. The mean weight is 1.2, which is quite close to one, meaning that the edges which are left does not represent particularly strong mobility patterns. This indicates that the final cluster solution and suggested segments are

TABLE 5. DEGREE AND WEIGHT STATISTICS BY LEVELS

	1. LEVEL		2. LEVEL		3. LEVEL		4. LEVEL	
	Degrees	Weight	Degrees	Weight	Degrees	Weight	Degrees	Weight
<i>Mean</i>	30.2	2.5	6.7	1.3	1.4	1.1	0.5	1.2
<i>Minimum</i>	0	1	0	1	0	1	0	1
<i>25. percentile</i>	17	1.2	3	1.1	0	1	0	1
<i>50. percentile (median)</i>	27	1.6	5	1.2	1	1.1	0	1.1
<i>75. percentile</i>	45	2.5	11.5	1.5	2	1.1	1	1.3
<i>Maximum</i>	67	78.4	20	6.3	7	1.5	2	1.5

relatively discrete. However, already at the second level, the mean weight has dropped from 2.5 to 1.3 and the mean number of degrees has dropped from 30.2 to 6.7.

In general we can conclude that the mobility between industries to a large extent is structured by other factors than chance. Also, the more than likely mobility patterns are relatively few but rather intense. Finally, they are to a great extent organized in a reciprocal way, which is indicated by the clustering algorithms success in minimizing mean degree and mean weight indicates, by clustering together the industries. The cluster solution of each level also seems to be relatively discrete. We now turn to the description of the segments identified by MONECA.

#### 4.2 *The segments of the Danish labour market*

The cluster analysis consists of four levels and choosing the highest level we get 26 clusters. The result of the MONECA algorithm at each level is represented in table 6 in which also the two key figures of size and within cluster mobility is are listed.

Table 6 allows for critical inspection of the cluster solution suggested by MONECA. To begin with, one should prefer higher level cluster solution over lower level solutions, but depending on the substantial research question guiding the analysis, it may make sense to choose a lower level solution. For instance, in this case, one might be interested in identifying market like clusters. Then, a too loose cluster may not meet the market criteria, because some of the level 1 industries does not exchange any level at all in the level 3 or 4 solution.

In table 7 a number of simple measures to assessment of the clusters are listed. These are proportion of mobility within the cluster, size as share of mobility, density and maximal path length. Table 7 only considers the level 3 and 4 clusters because we know that the level 2 clusters all have a density of 1 and a maximal path length of 1 due to the definition of the MONECA algorithm. Otherwise, within mobility and size for the level 1 and 2 cluster are listed in table 6.

Maximal path length tells us how integrated the cluster is. If it should be a true sub-market path length should be 1 and density 1 as well. This is true for all level 2 clusters. But when considering clusters at the higher levels we know that they have a maximal path length of more than one and a density of less than 1. Otherwise the cluster would have been formed at level 2. Therefore a level 3 or 4 cluster with a maximal path length of 2 is quite well integrated according to this measure, as the minimal path length for these levels are 2. Therefore we note that the level 3 cluster *Media, IT & business activity* and the two level 4 clusters have maximal path lengths of 3. This means that at least one of the industries are relatively decoupled from at least one of the other industries, and these should not be considered market like as labour does not flow freely within its boundaries. However, they can still be designated by the more loose concept of segments.

TABLE 6. CLUSTER STRUCTURE

Level 1			Level 2			Level 3			Level 3						
Title	Segment	W/Value add.	Seg	Title	Segment	W/Value add.	Seg	Title	Segment	W/Value add.	Seg	Title	Segment	W/Value add.	Seg
Bakery	10	0.412	0.006	Retail sale food	22	0.544	0.065	Retail sale	10	0.582	0.094	Sales, restaurants & cleaning	1	0.638	0.154
Re. sale auto fuel	57	0.317	0.007	Retail sale food	22	0.544	0.065	Retail sale	10	0.582	0.094	Sales, restaurants & cleaning	1	0.638	0.154
Rs. food non-specialized	65	0.482	0.03	Retail sale food	22	0.544	0.065	Retail sale	10	0.582	0.094	Sales, restaurants & cleaning	1	0.638	0.154
Rs. food specialized	66	0.322	0.005	Retail sale food	22	0.544	0.065	Retail sale	10	0.582	0.094	Sales, restaurants & cleaning	1	0.638	0.154
Department stores	67	0.42	0.017	Retail sale food	22	0.544	0.065	Retail sale	10	0.582	0.094	Sales, restaurants & cleaning	1	0.638	0.154
Rs. cloth & footwear	69	0.498	0.013	Retail sale non-food	29	0.498	0.029	Retail sale	10	0.582	0.094	Sales, restaurants & cleaning	1	0.638	0.154
Rs. specialized stores	71	0.437	0.015	Retail sale non-food	29	0.498	0.029	Retail sale	10	0.582	0.094	Sales, restaurants & cleaning	1	0.638	0.154
Hotel	73	0.373	0.011	Hotels, restaurants and cleaning	24	0.52	0.06	Hotels, restaurants and cleaning	13	NA	NA	Sales, restaurants & cleaning	1	0.638	0.154
Restaurant	74	0.453	0.035	Hotels, restaurants and cleaning	24	0.52	0.06	Hotels, restaurants and cleaning	13	NA	NA	Sales, restaurants & cleaning	1	0.638	0.154
Building cleaning act.	95	0.449	0.015	Hotels, restaurants and cleaning	24	0.52	0.06	Hotels, restaurants and cleaning	13	NA	NA	Sales, restaurants & cleaning	1	0.638	0.154
Mfr. textiles	14	0.399	0.002	Mfr. & ws. cloth and textiles	3	0.451	0.017	Mfr. & ws of textiles, ws. other	4	0.46	0.03	Ws, media, IT & business act.	2	0.561	0.114
Mfr. wearing	15	0.395	0.001	Mfr. & ws. cloth and textiles	3	0.451	0.017	Mfr. & ws of textiles, ws. other	4	0.46	0.03	Ws, media, IT & business act.	2	0.561	0.114
Ws. textile & housh. goods	60	0.423	0.014	Mfr. & ws. cloth and textiles	3	0.451	0.017	Mfr. & ws of textiles, ws. other	4	0.46	0.03	Ws, media, IT & business act.	2	0.561	0.114
Mfr. Paint & soap	24	0.39	0.002	Ws. other	26	0.403	0.013	Mfr. & ws of textiles, ws. other	4	0.46	0.03	Ws, media, IT & business act.	2	0.561	0.114
Ws. other raw. mat.	62	0.414	0.006	Ws. other	26	0.403	0.013	Mfr. & ws of textiles, ws. other	4	0.46	0.03	Ws, media, IT & business act.	2	0.561	0.114
Other wholesale	64	0.33	0.005	Ws. other	26	0.403	0.013	Mfr. & ws of textiles, ws. other	4	0.46	0.03	Ws, media, IT & business act.	2	0.561	0.114
Newspapers	19	0.407	0.003	Media & business act.	7	0.414	0.046	Media, IT & business act.	12	0.534	0.083	Ws, media, IT & business act.	2	0.561	0.114
Publishing	20	0.421	0.004	Media & business act.	7	0.414	0.046	Media, IT & business act.	12	0.534	0.083	Ws, media, IT & business act.	2	0.561	0.114
Advertising	94	0.429	0.006	Media & business act.	7	0.414	0.046	Media, IT & business act.	12	0.534	0.083	Ws, media, IT & business act.	2	0.561	0.114
Other business act.	96	0.331	0.033	Media & business act.	7	0.414	0.046	Media, IT & business act.	12	0.534	0.083	Ws, media, IT & business act.	2	0.561	0.114
Ws. machinery	63	0.456	0.018	IT-service & ws. of equipment	25	0.575	0.037	Media, IT & business act.	12	0.534	0.083	Ws, media, IT & business act.	2	0.561	0.114
Computer activities	89	0.572	0.019	IT-service & ws. of equipment	25	0.575	0.037	Media, IT & business act.	12	0.534	0.083	Ws, media, IT & business act.	2	0.561	0.114
Mfr. transport equip. excl.ships	41	0.379	0.003	Air transport and support	6	0.55	0.007	Air & water transport	1	0.639	0.029	-	3	n/a	n/a
Air transport	79	0.586	0.004	Air transport and support	6	0.55	0.007	Air & water transport	1	0.639	0.029	-	3	n/a	n/a
Water transport	78	0.751	0.007	Shipping	28	0.604	0.021	Air & water transport	1	0.639	0.029	-	3	n/a	n/a
Supporting transport act.	80	0.498	0.014	Shipping	28	0.604	0.021	Air & water transport	1	0.639	0.029	-	3	n/a	n/a
Mfr. dairy prod.	9	0.426	0.003	Mfr. & ws. food	13	0.466	0.015	Mfr.& ws. of eatables	2	0.518	0.03	-	4	n/a	n/a
Mfr. other food	11	0.408	0.009	Mfr. & ws. food	13	0.466	0.015	Mfr.& ws. of eatables	2	0.518	0.03	-	4	n/a	n/a
Ws. grain	58	0.468	0.002	Mfr. & ws. food	13	0.466	0.015	Mfr.& ws. of eatables	2	0.518	0.03	-	4	n/a	n/a
Production of meat	8	0.49	0.006	Meat, beverage	19	0.482	0.016	Mfr.& ws. of eatables	2	0.518	0.03	-	4	n/a	n/a
Mfr. beverages	12	0.49	0.002	Meat, beverage	19	0.482	0.016	Mfr.& ws. of eatables	2	0.518	0.03	-	4	n/a	n/a
Ws. food	59	0.411	0.008	Meat, beverage	19	0.482	0.016	Mfr.& ws. of eatables	2	0.518	0.03	-	4	n/a	n/a
Mfr. paper	18	0.421	0.002	Printing, paper & plastics	9	0.476	0.013	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Printing	21	0.519	0.004	Printing, paper & plastics	9	0.476	0.013	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. Plastic	26	0.412	0.007	Printing, paper & plastics	9	0.476	0.013	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. wood	17	0.392	0.005	Machinery & wooden products	10	0.551	0.044	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. metal	29	0.431	0.002	Machinery & wooden products	10	0.551	0.044	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. building mat. of metal	30	0.469	0.01	Machinery & wooden products	10	0.551	0.044	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. various metal	31	0.385	0.005	Machinery & wooden products	10	0.551	0.044	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. ovens & cold storage	33	0.394	0.006	Machinery & wooden products	10	0.551	0.044	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. agricultural machinery	34	0.374	0.002	Machinery & wooden products	10	0.551	0.044	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. industry machinery	35	0.382	0.005	Machinery & wooden products	10	0.551	0.044	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. domestic appliances	36	0.364	0.001	Machinery & wooden products	10	0.551	0.044	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. Furniture	42	0.481	0.007	Machinery & wooden products	10	0.551	0.044	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. marine engines	32	0.414	0.005	Advanced machinery	15	0.475	0.02	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. computers	37	0.428	0.009	Advanced machinery	15	0.475	0.02	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. communication	38	0.39	0.003	Advanced machinery	15	0.475	0.02	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. medical & optical instr.	39	0.402	0.004	Advanced machinery	15	0.475	0.02	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. Glass	27	0.393	0.001	Toys, glass & jewellery	21	0.49	0.004	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Mfr. toys & jewellery	43	0.518	0.003	Toys, glass & jewellery	21	0.49	0.004	Printing & mfr. of tangibles	3	0.602	0.081	-	5	n/a	n/a
Electricity	44	0.525	0.003	Utility & engineering	1	0.569	0.017	Utility, engineering, science &	5	0.574	0.034	-	6	n/a	n/a
Gas	45	0.309	0.001	Utility & engineering	1	0.569	0.017	Utility, engineering, science &	5	0.574	0.034	-	6	n/a	n/a
Stam & hot water	46	0.316	0.001	Utility & engineering	1	0.569	0.017	Utility, engineering, science &	5	0.574	0.034	-	6	n/a	n/a
Water	47	0.325	0	Utility & engineering	1	0.569	0.017	Utility, engineering, science &	5	0.574	0.034	-	6	n/a	n/a
Engineering & architects	93	0.548	0.013	Utility & engineering	1	0.569	0.017	Utility, engineering, science &	5	0.574	0.034	-	6	n/a	n/a
Mfr. chemicals	23	0.387	0.001	Science, pharmacy & chemistry	5	0.52	0.016	Utility, engineering, science &	5	0.574	0.034	-	6	n/a	n/a
Mfr. Pharmaceuticals	25	0.444	0.004	Science, pharmacy & chemistry	5	0.52	0.016	Utility, engineering, science &	5	0.574	0.034	-	6	n/a	n/a
Research & development	90	0.417	0.004	Science, pharmacy & chemistry	5	0.52	0.016	Utility, engineering, science &	5	0.574	0.034	-	6	n/a	n/a
Higher education	102	0.463	0.008	Science, pharmacy & chemistry	5	0.52	0.016	Utility, engineering, science &	5	0.574	0.034	-	6	n/a	n/a
Financial institutions	82	0.631	0.01	Finance & insurance	4	0.665	0.021	Finance & real estate	6	0.643	0.035	-	7	n/a	n/a
Credit institutions	83	0.383	0.004	Finance & insurance	4	0.665	0.021	Finance & real estate	6	0.643	0.035	-	7	n/a	n/a
Insurance	84	0.577	0.005	Finance & insurance	4	0.665	0.021	Finance & real estate	6	0.643	0.035	-	7	n/a	n/a
Auxiliary act. to finance	85	0.434	0.002	Finance & insurance	4	0.665	0.021	Finance & real estate	6	0.643	0.035	-	7	n/a	n/a
Letting of property	86	0.501	0.008	Real estate	17	0.546	0.014	Finance & real estate	6	0.643	0.035	-	7	n/a	n/a
Real estate agents	87	0.495	0.006	Real estate	17	0.546	0.014	Finance & real estate	6	0.643	0.035	-	7	n/a	n/a

TABLE CONTINUES ON NEXT PAGE ...

... TABLE 6 CONTINUED FROM PREVIOUS PAGE

Level 1				Level 2				Level 3				Level 3			
Title	Segment	Within mod.	Size	Title	Segment	Within mod.	Size	Title	Segment	Within mod.	Size	Title	Segment	Within mod.	Size
Hospitals	104	0.664	0.031	Health	27	0.687	0.044	Health & social inst. Adults	7	0.727	0.111	-	8	n/a	n/a
Medical, dental & veterinary	105	0.569	0.012	Health	27	0.687	0.044	Health & social inst. Adults	7	0.727	0.111	-	8	n/a	n/a
Social inst. Adults	107	0.622	0.067	Social inst. Adults	43	N/A	N/A	Health & social inst. Adults	7	0.727	0.111	-	8	n/a	n/a
Sale of motor vehicles	55	0.598	0.013	Auto sales & repair	8	0.631	0.018	Sale and repair of motor vehicles	8	0.579	0.023	-	9	n/a	n/a
Repair of motor vehicles	56	0.421	0.005	Auto sales & repair	8	0.631	0.018	Sale and repair of motor vehicles	8	0.579	0.023	-	9	n/a	n/a
Repair house. Goods	72	0.414	0.001	Repair & leasing of household appliances	31	0.313	0.005	Sale and repair of motor vehicles	8	0.579	0.023	-	9	n/a	n/a
Leasing of transport & machinery	88	0.284	0.004	Repair & leasing of household appliances	31	0.313	0.005	Sale and repair of motor vehicles	8	0.579	0.023	-	9	n/a	n/a
General public adm.	97	0.631	0.022	Public administration	18	0.662	0.033	Public adm., organizations & law	9	0.668	0.051	-	10	n/a	n/a
Adm. Public sectors	98	0.427	0.012	Public administration	18	0.662	0.033	Public adm., organizations & law	9	0.668	0.051	-	10	n/a	n/a
Legal act.	91	0.634	0.004	Law & organizations	32	0.39	0.017	Public adm., organizations & law	9	0.668	0.051	-	10	n/a	n/a
Act. membership org.	109	0.566	0.013	Law & organizations	32	0.39	0.017	Public adm., organizations & law	9	0.668	0.051	-	10	n/a	n/a
Gen. Contractor	48	0.495	0.025	Construction	14	0.678	0.052	Construction & trucks (Bob the Builder)	11	0.682	0.08	-	11	n/a	n/a
Bricklaying	49	0.476	0.006	Construction	14	0.678	0.052	Construction & trucks (Bob the Builder)	11	0.682	0.08	-	11	n/a	n/a
Joinery installation	52	0.553	0.015	Construction	14	0.678	0.052	Construction & trucks (Bob the Builder)	11	0.682	0.08	-	11	n/a	n/a
Other construction	54	0.408	0.006	Construction	14	0.678	0.052	Construction & trucks (Bob the Builder)	11	0.682	0.08	-	11	n/a	n/a
Esctr. gravel	7	0.384	0.001	Trucks	16	0.564	0.027	Construction & trucks (Bob the Builder)	11	0.682	0.08	-	11	n/a	n/a
Mfr. tiles, bricks, cement	28	0.401	0.004	Trucks	16	0.564	0.027	Construction & trucks (Bob the Builder)	11	0.682	0.08	-	11	n/a	n/a
Freight by road and pipelines	77	0.575	0.017	Trucks	16	0.564	0.027	Construction & trucks (Bob the Builder)	11	0.682	0.08	-	11	n/a	n/a
Refuse disposal	108	0.462	0.005	Trucks	16	0.564	0.027	Construction & trucks (Bob the Builder)	11	0.682	0.08	-	11	n/a	n/a
Esctr. oil & gas	6	0.363	0.001	Mfr. ships, electricity & plumbing	2	0.633	0.023	-	14	n/a	n/a	-	12	n/a	n/a
Building of ships	40	0.438	0.002	Mfr. ships, electricity & plumbing	2	0.633	0.023	-	14	n/a	n/a	-	12	n/a	n/a
Electrical installation	50	0.592	0.012	Mfr. ships, electricity & plumbing	2	0.633	0.023	-	14	n/a	n/a	-	12	n/a	n/a
Plumbing	51	0.614	0.008	Mfr. ships, electricity & plumbing	2	0.633	0.023	-	14	n/a	n/a	-	12	n/a	n/a
Agriculture	1	0.63	0.016	Forestry & agriculture	11	0.619	0.025	-	15	n/a	n/a	-	13	n/a	n/a
Gardening	2	0.449	0.003	Forestry & agriculture	11	0.619	0.025	-	15	n/a	n/a	-	13	n/a	n/a
Machine pools	3	0.347	0.004	Forestry & agriculture	11	0.619	0.025	-	15	n/a	n/a	-	13	n/a	n/a
Forestry	4	0.431	0.001	Forestry & agriculture	11	0.619	0.025	-	15	n/a	n/a	-	13	n/a	n/a
Railway and buses	75	0.651	0.007	Transp. of people by land	12	0.707	0.013	-	16	n/a	n/a	-	14	n/a	n/a
Taxis and coach	76	0.647	0.006	Transp. of people by land	12	0.707	0.013	-	16	n/a	n/a	-	14	n/a	n/a
Painting & glazing	53	0.706	0.006	Paint & household goods	20	0.548	0.023	-	17	n/a	n/a	-	15	n/a	n/a
W. wood and const. mat.	61	0.448	0.005	Paint & household goods	20	0.548	0.023	-	17	n/a	n/a	-	15	n/a	n/a
Rt. furniture & house. Appl.	70	0.444	0.012	Paint & household goods	20	0.548	0.023	-	17	n/a	n/a	-	15	n/a	n/a
Primary education	100	0.621	0.025	Primary education	30	0.653	0.071	-	17	n/a	n/a	-	15	n/a	n/a
Social inst. children	106	0.585	0.046	Primary education	30	0.653	0.071	-	17	n/a	n/a	-	15	n/a	n/a
Secondary education	101	0.465	0.011	Secondary education	23	0.519	0.02	-	18	n/a	n/a	-	16	n/a	n/a
Adult education	103	0.488	0.01	Secondary education	23	0.519	0.02	-	18	n/a	n/a	-	16	n/a	n/a
Fishing	5	0.718	0.002	-	35	n/a	n/a	-	19	n/a	n/a	-	17	n/a	n/a
Mfr. tobacco	13	-	-	-	38	n/a	n/a	-	20	n/a	n/a	-	18	n/a	n/a
Mfr. leather & footwear	16	-	-	-	36	n/a	n/a	-	21	n/a	n/a	-	19	n/a	n/a
Mfr. refined petroleum	22	-	-	-	37	n/a	n/a	-	22	n/a	n/a	-	20	n/a	n/a
Rt. pharmaceuticals	68	0.635	0.004	-	42	n/a	n/a	-	23	n/a	n/a	-	21	n/a	n/a
Post & telecommunications	81	0.56	0.02	-	40	n/a	n/a	-	24	n/a	n/a	-	22	n/a	n/a
Accounting	92	0.586	0.007	-	33	n/a	n/a	-	25	n/a	n/a	-	23	n/a	n/a
Defense, police & justice	99	0.632	0.022	-	34	n/a	n/a	-	26	n/a	n/a	-	24	n/a	n/a
Recreational, culture, sport	110	0.502	0.022	-	41	n/a	n/a	-	27	n/a	n/a	-	25	n/a	n/a
Other. serv. act.	111	0.53	0.008	-	39	n/a	n/a	-	28	n/a	n/a	-	26	n/a	n/a

We will only consider the level 4 solutions with a maximal path length of 3, which we consider problematically high, in order to exemplify how these numbers can be utilized when inspecting and assessing the validity of the cluster solutions produced by MONECA. When we do not consider *Media, IT & business activity*, it is because it is included in the second level 4 cluster.

Next we consider the integration measures of density and within mobility. We will use the two level 4 clusters as examples. The higher the density the more market like can the cluster be said to be and if density is 1 we can speak of a true sub labour market. The higher the within mobility mean that the cluster is more discretely separated from its surroundings. A within mobility of 100 % is equal to absolute separation. Both level 4 clusters have a maximal path length of 3, meaning at least one industry is relatively decoupled from one or more of the other industries.

*Wholesale, media & business activities* has a density of 0.455 density. This is relatively low. Especially, if we consider that this cluster consists of level 3 clusters number 4 and 12 (See table 6) with quite high densities, respectively 0.800 and 0.667, the cluster appears somewhat composed. Also within mobility is low, 56.1 % which is well below the level 4 mean of 62.5 %.

These numbers indicates that *Wholesale, media & business activities* is quite loosely integrated and given the research question, it could be argued that the level 3 solution is better with regard to these industries.

*Sales, restaurants & cleaning* has a high density of 0.644 indicating that it is quite well integrated. Nonetheless, this is substantial decline compared to the two lower level cluster which forms this cluster. The level 3 cluster *Retail sale* has a density of 0.810 and the level two cluster *Hotels, restaurants & cleaning* has a density of 1. However, within mobility, tells a different story. *Retail sale* has a within mobility of 58.2 %, which is below the level 3 average of 60.7 % (see table 4) and *Hotels, restaurants & cleaning* has a within mobility of 52 % which also is below the level 2 average of 57.2. When joined together at level four they form a cluster with a within mobility at 63.8 % which is above the level three mean of 62.5 %. Seen from this perspective, a lot speaks in favor of the level four

TABLE 7. INTEGRATION MEASURES FOR THE LEVEL 3 AND LEVEL 4 CLUSTERS

<i>Level 3</i>					
TITLE	NR.	MAXIMAL PATH	DENSITY	WITHIN MOBILITY	SIZE
Air & water transport	1	2	0.500	63.9 %	2.9%
Mfr.& ws. of eatables	2	2	0.867	51.8 %	3.0%
Printing & mfr. of tangibles	3	2	0.752	60.2 %	8.1%
Mfr. & ws of textiles, ws. other	4	2	0.800	46.0 %	3.0%
Utility, engineering, science & pharmacy	5	2	0.556	57.4 %	3.4%
Finance & real estate	6	2	0.733	64.3 %	3.5%
Health & social inst. Adults	7	2	0.667	72.7 %	11.1%
Sale and repair of motor vehichels	8	2	0.833	57.9 %	2.3%
Public adm., organizations & law	9	2	0.667	66.8 %	5.1%
Retail sale	10	2	0.810	58.2 %	9.4%
Construction & trucks (Bob the Builder)	11	2	0.750	68.2 %	8.0%
Media, IT & business act.	12	3	0.667	53.4 %	8.3%
<i>Level 4</i>					
TITLE	NR.	MAXIMAL PATH	DENSITY	WITHIN MOBILITY	SIZE
Sales, restaurants & cleaning	1	3	0.644	63.8 %	0.154
Ws., media, IT & business act.	2	3	0.455	56.1 %	0.114



solution as one or more proportionally large mobility patters joins together the two clusters.

The need to make a holistic assessment, and not just relying on one measure, is underlined if one considers how no co-variation between density and within mobility can be detected in table 7. In addition to these simple measures, inspection of the network graphs might prove helpful in determining the correct solution.

In the end, it is the meaningfulness of the solution that counts. In table 6 we have designated some descriptive names to the cluster solutions. We find that, the clusters suggested by MONECA, overall are quite meaningful, with the possible exception of the level 4 cluster *Wholesale, media & business activities*.

Table 8 lists the 10 largest segments. These ten clusters covers 76.1 % of the labour market. We shall briefly point to some interesting findings. *Retail sale, restaurant, hotels and cleaning* are clustered together. This makes sense as they all can be designated labour heavy service industries mainly occupied by semi-skilled and unskilled (female) workers. However, using standard NACE grouping this group would have been separated. With the NACE-9 standard grouping these are separated into 2 different groups and using NACE-27 standard grouping they are separated into 4 different groups.<sup>3</sup>

Second, we have *Wholesale, media, IT & business activities*. This is a very composed segment and it is very difficult to say something general about it. As discussed above one could argue for preferring a lower level solution with regard to these industries. At level three it makes up the two segments of *Manufacture & wholesale of textiles & wholesale of other* and *Media, IT & business activities*. The first one can be said to be structured by a kind of products, textiles, and a economical process, wholesale. The second is characterized by the use of and service of information technology.

Third is *Health & social institutions for adults* which is dominated by professions like doctors, nurses and other health and caretaking staff like specialized pedagogues. It is knowledge and labour heavy processes mainly situated within the public sector.

Fourth is *Printing & manufacture of tangibles*. This is a segment in which skilled blue collar (male) workers are dominating the scene of production. Operation of advanced machine and specialized

TABLE 8. TOP 10 SEGMENTS BY SIZE

NAME	WITHIN MOBILITY	SIZE	MAXIMAL PATH LENGTH	DENSITY
Sales, restaurants & cleaning	63.8%	15.4%	3	0.644
Ws., media, IT & business act.	56.1%	11.4%	3	0.455
Health & social inst. Adults	72.7%	11.1%	2	0.667
Printing & mfr. of tangibles	60.2%	8.1%	2	0.752
Construction & trucks (Bob the Builder)	68.2%	8.0%	2	0.750
Primary education	65.3%	7.1%	1	1.000
Public adm., organizations & law	66.8%	5.1%	2	0.667
Finance & real estate	64.3%	3.5%	2	0.733
Utility, engineering, science & pharmacy	57.4%	3.4%	2	0.556
Mfr. & ws. of eatables	51.8%	3.0%	2	0.867
Total (average)	63.6%	76.1%	2	0.709

<sup>3</sup> This is with reference to the Danish edition of the international NACE nomenclature, DB-03, which is available from Statistics Denmark's homepage, [www.dst.dk](http://www.dst.dk).

knowledge with regard to the production process is essential skills.

Fifth is *Construction & trucks* in which the production processes are characterized by the key element of large mobile machines like trucks, excavators and the like. Not only construction, but also transportation by trucks and refuse disposal is part of this segment. The common denominator is that you either operate trucks or other large machines or this sort of machines takes a central place in the production processes.

Sixth is *Primary education*. This is primary school and kindergartens and the segment is characterized by its educational aspect and the fact that it is children which are the focal point of the processes involved.

Seventh and eight can be taken together as this is white collar workers of respectively the public and the private sector. On the one hand we have the *public administration, legal activities and membership organizations* which in most cases are interest organizations which try to influence government and public administration. As many public officials have an law degree it makes sense that legal activity is included here. On the other we have *finance & real estate* which also includes insurance. This is where the money are and jobs involve a lot of number work.

Ninth, we have *Utility, engineering, science & pharmacy*. The for the companies in this segment denominator is all different kinds of knowledge. The dominating occupational layer is surely the professionals and technicians.

The final segment, *manufacture & wholesale of eatables*, is not defined by its workforce's occupational homogeneity, but by the products which are various kinds of food. Manufacture and wholesale are separated by the NACE nomenclature no matter at which level. Our analysis suggests that in this case the formative principle of labour market segments can be the product, rather than the position in the economical process assumed by NACE which sharply distinguish between manufacture and wholesale, as two very different things. From the perspective of labour market mobility, this distinction seems in this case to be wrong. This logic was also present in the case of *Manufacture & wholesale of textiles & wholesale of other*.

This, final analysis of the findings of MONECA, was made in order to show how already at the initial descriptive states have, this exploratory method opens up for a more nuanced view of what factors that may segment the labour market and how their does not seem to be one common logic to the barriers separating the segments. Traditional demarcation lines like professions, place in the economical process do not always take primacy: in some cases it seems that the product, a certain instrument of production or even gender may be the true organizing principle.

## 5. FUTURE APPLICATIONS

This approach to the study of social mobility can be applied to a number of themes and in the following we list those we think are the most promising.

- 1) Occupational mobility. This could be carried out in two ways answering two different, however, related question. First, it could be done looking at the movements at the labour market in a timespan, as we have just done just using the industry categories instead of occupation categories. This would tell us about labour market segmentation by occupation, but it would also tell us about society's structure of stratification. Differing from industries, occupations can be ordered as a hierarchy by mean wage, mean lifetime, mean social status and prestige of its members and so on. Therefore, occupational clusters may be ordered into hierarchy that resembles a class structures. However, a occupational cluster is not the same as a class, as for instance two clusters may share position and relation to the other clusters in the hierarchy and thus can be said to be the same class. The second approach, would be to look at intergenerational mobility. A central issue in the stratification theory is the percistence of stratification structures over time, and the much debated thesis about social reproduction and inheritance of social position. These questions could be addressed by the approach suggested in this paper.
- 2) Marriage mobility. Apart from looking at social mobility as something done by the individual or as the movement of child relative to parent, mobility can also be marriage across or within occupational-, educational-, ethnic-, geographic categories and probably many more categories. Marriage strategies has for a long time been known as a way to 'climb' the social hierarchy. However, a mapping of the typical marriage patterns would allow for investigation of to what extend the social reproduction happens throught marriage and choice of partner, and who marry 'up' or 'down' in the hierarchy.
- 3) Mobility between geographical entities. This could be migration between countries, travellers and tourist patterns, minorities migration between cities, the movement patternes between regions within a country, and so on.

One of the strength of MONECA in the example above, is that the use of the clique concept and the cut point of 1 RR give theoretically sense, as it resembles market conditions. However, in other instances given another research question, the clique may not be the ideal network theoretical concept to use. It may be too restrictive as more loose affiliations make theoretical sense to cluster together. If that is the case, one can simply change this is in the algorithm and replace the clique criteria by a core, K-plex, clan or N-clique criteria. Also, in our case, only considering reciprocal relationships make theoretical sense. In other instances, this may however not be the case and this can be adjusted as well.

These possibilities for careful adjustment of the methodological tools and thereby the theoretical assumptions made in the analysis are extremely important. MONECA is first and foremost an exploratory and descriptive method. When the task is to explore data it is crucial to how exactly the outcome was made in order to hypothesize about its meaning and to be able to construct fitting tests of these hypothesis. What MONECA calls for is careful descriptive analysis prior to the explanation of the phenomena. MONECA can in an exploratory way help us define the

categories we try to explain and guide us in our search for the explanation, that very well may involve a multiplicity of factors. The above exmaplifying analysis, hopefully, in a perfomative way, have demonstrated how the utilization of network analysis in an descriptive manner, can add a higher level of precision and empirical sensitivity to the discipline of social mobility studies.

## REFERENCES

- Agresti, Alan & Barbara Finlay (1997): *Statistical Methods for the Social Sciences*. New Jersey: Prentice Hall.
- Bailey, Kenneth D. (1975): "Cluster Analysis" in, *Sociological Methodology*, vol. 6: 59-128
- Balázs, Adamcsek, Gergely Palla, Illés J. Farkas, Imre Derényi and Tamás Vicsek (2006): "CFinder: Locating cliques and overlapping modules in biological networks" in, *Bioinformatics*, vol. 20(8): 1021-1023.
- Becker, Gary S. (1993): *Human Capital. A Theoretical and Empirical Analysis with Special Reference to Education*. 3rd Ed. Chicago: The University of Chicago Press.
- Bluestone, B. (1970): "The Tripartite Economy: Labor Market and the Working Poor" in, *Poverty and Human Resource Abstracts*, vol. 5:15-35.
- Boje, Thomas P. (1985): *Segmentation and Social Processes on the Danish Labour Market*. Working Paper no. 6/85. Institute of Economics and Planning, Roskilde University Center.
- Boje, Thomas P. (1986): "Segmentation and Mobility: An Analysis of Labour Market Flows on the Danish Labour Market" in, *Acta Sociologica*, vol. 29(2):171-178.
- Boje, Thomas P. (1987): *Mobilitets- og Beskæftigelsesmønstre på det Danske Arbejdsmarked 1980-1981: Bidrag til en teori om arbejdsmarkedets segmentering*. Roskilde: Institut for Samfundsøkonomi og Planlægning, Roskilde Universitetscenter.
- Boje, Thomas P. (1990): *Flexibility and Fragmentation in the Labour market. Recent trends in the structuring of employment and industrial realtions*. Paper presented at the EAEPE 1990 conference, Florence. Research Report no. 6/90. Institute of Economics and Planning, Roskilde University Center.
- Boje, Thomas P. & Christian Toft (1989): *Arbejdsmarkeder og segmenteringsteorier – en introduction til nyere segmenteringsteori*. Haderslev: Institut for Grænseregionsforskning.
- Boston, Thomas D. (1990): "Segmented Labor Markets: New Evidence From a Study of Four Race-Gender Groups" in, *Industrial & Labor Relations Review*, vol. 44(1):99-115.
- Cain, Glen G. (1976): "The Challenge of Segmented Labor Market Theories to Orthodox Theory: A Survey" in, *Journal of Economic Literature*, vol. 14(4):1215-1257.
- Cormack, R.M. (1971): "A Review of Classification" in, *Journal of the Royal Statistical Society, Series A (General)*, vol. 134(3):321-353.
- Danish Technological Institute (2008): *Job Mobility in the European Union: Optimising its Social and Economic Benefits*. Available at <http://www.dti.dk/job-mobility-in-europe/25567>.
- Daw, Jonathan & Jessica Halliday Hardie (2012): "Compensating differentials, labor market segmentation, and wage inequality" in, *Social Science Research*, vol. 41(5):1179-1197.
- Derényi, Imre, Gergely Palla and Tamás Vicsek (2005): "Clique Percolation in Random Networks" in, *Physical Review Letters*, vol. 94.

- Doeringer, Peter B. & Michael J. Piore (1971): "Unemployment and the "dual labor market" in, *Public Interest*, vol. 38
- Doeringer, Peter B. & Michael J. Piore (1975): *Internal Labour Markets and Manpower Analysis*. Lexington: Heath Lexington Books.
- Due, Jesper, Carsten Strøby Jensen, Jørgen Steen Madsen & Lars Kjerulf Petersen (1994): *The Survival of the Danish Model*. Copenhagen: DJØF Publishing.
- Due, Jesper, Jørgen Steen Madsen & Mie Dahlskov Pihl (2010): *LO-dokumentationen 1/2010. Udviklingen i den faglige organisering: årsager og konsekvenser for den danske model*. Copenhagen: Landsorganisationen i Danmark.
- Erikson, Robert, John H. Goldthorpe & Lucienne Portocarero (1979): "Intergenerational Class Mobility in Three Western European Countries: England, France and Sweden" in, *The British Journal of Sociology*, vol. 30(4):415-441.
- Everit, Brian S. (2011): *Cluster Analysis*. Hoboken: Wiley
- Fichtenbaum, Rudy, Kwabena Gyimah-Brempong & Paulette Olson (1994): "New Evidence on the Labour Segmentation Hypothesis" in, *Review of Social Economy*, vol. 52(1):20-39.
- Frank, Kenneth A. (1995): "Identifying Cohesive Subgroups" in, *Social Networks*, vol. 17:27-56.
- Farkas, Illés J., Dániel Ábel, Gergely Palla and Tamás Vicsek (2007): "Weighted network modules" in, *New Journal of Physics*, vol. 9(180).
- Goldthorpe, John H. (2000): "Social Class and the Differentiation of Employment Contracts" in, *On Sociology*. Oxford: Oxford University Press.
- Goldthorpe, John H. (2002): "Occupational Sociology, Yes: Class Analysis, No: Comment on Grusky and Weeden's Research Agenda" in, *Acta Sociologica*, vol. 45(3):211-217.
- Gordon, D.M. (1972): *Theories of Poverty and Unemployment: orthodox, radical, and dual labor market perspectives*. Lexington: Lexington Books.
- Grusky, David B. & Kim A. Weeden (2002): "Class Analysis and the Heavy Weight of Convention" in, *Acta Sociologica*, vol. 45(3):229-236.
- Grusky, David B. & G. Galescu (2005): "Foundations of a Neo-Durkheimian class analysis" in, Erik Olin Wright (ed): *Approaches to Class Analysis*: Cambridge: Cambridge University Press.
- Guha, Sudipto, Rajeev Rastogi and Kyuseok Shim (2000): "ROCK: A Robust Clustering Algorithm for Categorical Attributes" in, *Information Systems*, vol. 25(5): 345-366.
- Hall, R.E. (1970): "Why is the Unemployment Rate so High at Full Employment?" in, *Brookings Papers on Economic Activity*, no. 3:396-410.
- Hudson, Kenneth (2007): "The new labor market segmentation: Labor market dualism in the new economy" in, *Social Science Research*, vol. 36:286-312.

- Jain, A.K., M.N. Murty & P.J. Flynn (1999): "Data Clustering: A Review" in, *ACM Computing Surveys*, vol. 31(3): 264-323.
- Jensen, Carsten Strøby (2008): "Employment Relations in Denmark: Explaining flexicurity", paper presented at the *American Sociological Association, Annual Meeting, Boston*, 2005.
- Jensen, Carsten Strøby (2011): "The Flexibility of Flexicurity: the Danish model reconsidered" in, *Economic and Industrial Democracy*, vol. 32(4):721-737.
- Jensen, Carsten Strøby (2012): *Industrial Relations in Denmark – from conflict-based consensus to consensus-based conflict*. Copenhagen: DJØF Publishing.
- Jonsson, Jan O., David B. Grusky, Matthew Di Carlo, Reinhard Pollak & Mary C. Brinton (2009): "Microclass Mobility: Social Reproduction in Four Countries" in, *American Journal of Sociology*, vol. 114(4):977-1036.
- Jørgensen, Henning & Per Kongshøj Madsen (ed.) (2007): *Flexicurity and Beyond. Finding a new agenda for the European Social Model*. Copenhagen: DJØF Publishing.
- Leigh, Duane E. (1976): "The Occupational Mobility of Young Men, 1965-1970" in, *Industrial and Labor Relations Review*, vol. 30(1):68-78.
- Leontaridi, Marianthi Rannia (1998): "Segmented Labour Markets: Theory and Evidence" in, *Journal of Economic Surveys*, vol. 12(1):63-101.
- Loveridge, R & A.L. Mok (1979): *Theories of labour market segmentation. A critique*. London: Martinus Nijhoff Social Science Division.
- Lu, S.Y. & K.S. Fu (1978): "A sentence-to-sentence clustering procedure for pattern analysis". *IEEE Trans. Syst. Man Cybern.* Vol. 8: 381-389.
- Madsen, Per Kongshøj (2006): "Labour Market Flexibility and Social Protection In European Welfare States – Contrasts and Similarities" in, *Australian Bulletin of Labour*, vol. 32(2):139-162.
- Mokken, Robert J. (1979): "Cliques, Clubs and Clans" in, *Quality and Quantity*, vol. 13:161-173.
- Omran, Mahamed G.H., Andries P. Engelbrecht and Ayed Salman (2007): "An overview of clustering methods" in, *Intelligent Data Analysis*, vol. 11(6): 583-605.
- Ozawa, K (1985): "A stratificational overlapping overlapping cluster scheme" in, *Pattern Recognition*, vol. 18(3/4): 279-286.
- Phelps Brown, E.H. (1977): *The Inequality of Pay*. Oxford: Oxford University Press.
- Prell, Christina (2012): *Social Network Analysis - history, theory & Methodology*. London: Sage.
- Reich, Michael, David M. Gordon & Richard C. Edwards (1973): "A Theory of Labour Market Segmentation" in, *The American Economic Review*, vol. 63(2):359-365.
- Sandri, Giorgio (1969): "On the Logic of Classification" in, *Quality & Quantity*, vol. 3(1-2): 80-124.
- Snoke, David and Song Yang (2008): *Social Network Analysis*. London: Sage.
- Scott, John (2000): *Social Network Analysis. A Handbook*. London: Sage.

- Sokal, Robert R. (1985): "The Continuing Search for Order" in, *The American Naturalist*, vol. 126(6): 729-749.
- Stier, Haya & David B. Grusky (1990): "An Overlapping Persistence Model of Career Mobility" in, *American Sociological Review*, vol. 55:736-756.
- Wachter, Michael L. (1974): "Primary and Secondary Labor Markets: A Critique of the Dual Approach" in, *Brookings Papers on Economic Activity*, vol. 1974(3):637-693
- Williams, W.T. (1971): "Principles of Clustering" in, *Annual Review of Ecology and Systematics*, vol. 2(1): 303-326.
- Zahn, C.T. (1971): "Graph-theoretical methods for detecting and describing gestalt-clusters" in, *IEEE Transactions on Computers C-20*, (Apr.), 68-86.