The structure of European development — Community detection in inter-industry and external trade networks*

Nadia Garbellini[†]
and
Ariel Luis Wirkierman[‡]

Abstract Current debates on the future of the European Monetary Union (EMU, hereinafter) — as well as that of the European Union (EU27, hereinafter) — are mostly based on financial and aggregate parameters describing its constituent members, no reference being made to the disaggregated and interdependent productive features of the EMU/EU27 economies. The aim of the present paper is to analyse the clustering of industries within the EMU/EU27 and the characterisation of trade flows among its members, in order to unfold some structural features of European development.

Thanks to the recent Supply, Use and Input-Output tables database for the EMU/EU27 areas provided by EUROSTAT (2011), it is possible to analyse their respective flow structures as integrated networks of inter-industry relations, as opposed to the pattern of bilateral trade flows among EMU/EU27 countries (see EUROSTAT 2006), whether for total trade as well as for commodity-specific trade matrices.

In particular, we considered the application of a spectral bisection algorithm for community detection, and analysed the resulting clustering of industries (by means of EMU/EU27 intermediate consumption networks) and countries (by means of intra-EU external trade matrices). By exploiting the interplay between dependence and independence, hierarchy and circularity (as emphasised by Leontief 1963 [1986], p. 166) an analysis of the internal structure of the EMU/EU27 can be advanced, in order to assess the potentialities of European economic integration in productive structures, effective demand and income generation patterns.

Keywords Input-Output analysis, Network theory, Spectral Bisection algorithm, Industrial Clusters, International Trade Matrices

JEL classification C67, R15, O41

*Paper prepared for the *Final WIOD Conference: Causes and Consequences of Globalization*, Groningen (The Netherlands), April 24-26, 2012. Draft version, please do not quote without permission of the authors.

†Università degli studi di Pavia, Italy (nadia.garbellini@unipv.it)

‡Università Cattolica di Milano, Italy (ariwirkierman@gmail.com)

1 Introduction

When in 2006 the Council of the European Union defined the 'strategic priorities for innovation action at EU level',[1] it established that 'helping innovation in regions' was one of them, and that trans-national cluster-based analysis and policy could be the 'effective means to strengthen regional innovation'.

Since then, different initiatives have been promoted. One of these has been to establish a European Cluster Observatory,[2] which has undertaken the crucial methodological task of arriving at an operational definition, identification and measurement of 'clusters' in European regions (see EC 2008).[3]

At a definitional level it established that:

> clusters can be defined as a group of firms, related economic actors, and institutions that are located near each other and have reached a sufficient scale to develop specialised expertise, services, resources, suppliers and skills.
>
> (EC 2008, pp. 9)

which can be linked to the definition provided by Porter (2003):

> We define a cluster as a geographically proximate group of interconnected companies, suppliers, service providers and associated institutions in a particular field, linked by externalities of various types. [...] Clusters are important because of the externalities that connect the constituent industries, such as common technologies, skills, knowledge and purchased inputs. Note that a given industry can be part of more than one cluster based on different patterns of externalities.
>
> (Porter 2003, p. 562)

In both definitions the geographical dimension is privileged over the type of 'externality' (i.e. linkage relation between members of a cluster). Moreover, while Porter (2003) starts from the micro-economic agent ('group of interconnected companies'), it immediately switches to the 'industry' when detailing the 'externalities that connect the constituent industries' of the cluster, i.e. to a macro-economic unit of analysis. This change reflects the need to have an operational concept, as in the empirical implementation of these definitions both — the EU Commission and Porter — take the 'industry' level as the minimum unit being clustered.

In fact, the European Cluster Observatory adapted the methodology of Porter's US Cluster Mapping Project (see Porter 2003), labelled 'statistical cluster mapping', to identify and measure clusters in European regions:

[1] Conclusions of the Council of the EU meeting, December 4, 2006, available at: `http://www.consilium.europa.eu/uedocs/cms_Data/docs/pressdata/en/intm/91989.pdf`

[2] `http://www.clusterobservatory.eu`

[3] By 'region' we intend those geographical units as defined by the Nomenclature of Territorial Units for Statistics (NUTS) of EUROSTAT.

> The approach followed by the European Cluster Observatory is based on measuring indirectly the revealed effects of co-locations of businesses that are assumed to be observable when a cluster is present, such as concentrated employment rates or higher productivity.
>
> (EC 2008, pp. 15-6)

Concretely, the crux of this method consists in computing 'locational employment correlation coefficients' between couple of industries across regions and then grouping together those activities nearly always geographically associated.[4] For example, if the employment levels of 'Manufacture of musical instruments' and 'Sound recording and music publishing activities' are highly correlated across different regions, it is plausible to group these two industries in an 'Entertainment' cluster. It is clear that statistical cluster mapping "builds on an implicit definition of clusters based on the concept of co-location of industries, as well as on conventions for the categorisation of data, such as for the thresholds used for the definition of the relative strengths of clusters" (EC 2008, pp. 18).

Moreover, at its initial stage, the European initiative proceeded by translating the 'Porter Cluster model' (i.e. the clustering pattern of industries according to US data) directly into Europe's NACE Rev. 1 codes and adopting it to detect clusters in regions of selected European states (see the discussion in EC 2007, pp. 19-22). But this implied assuming a common template for the industry composition of clusters between the US and the EU. This seems to be a strong assumption.[5]

Even though regions are a most adequate unit of analysis given that policy measures are often implemented at this level (e.g. EU Structural Funds Programmes), the clustering of industries should disentangle country-specific features from common EU27 (or EMU) clustering patterns, in order to frame cross-regional policies with national awareness.

Hence, instead of imposing a clustering 'model' on regional data, it could be interesting to perform data mining on national and EU27/EMU Input-Output networks to discover country-specific industry composition of clusters, and measure (dis-)similarities among countries' clustering structures. In this way, while international comparisons should be more carefully performed, the percentage of the economy covered by clustered activities might increase, as many industries not included in the 'imposed' clustering tem-

[4]As detailed by Porter: "we identified pairs and then groups of tightly linked industries based on statistically significant locational correlations. Standard clustering algorithms proved inadequate to revealing the multiple patterns of linkages across industries. To build up clusters, then, we proceeded pragmatically, beginning with small groups of obviously related industries and then tracing correlation patterns to others" (Porter 2003, p. 563).

[5]See, in fact, the discussion in Porter (2003, p. 562).

plate might be part of the corresponding country-specific cluster. It might be even possible to detect which industries are always clustered together in *all* countries, to obtain a minimal 'standard' cluster template, augmented by national specificities.

Moreover, high industry-employment correlation across multiple geographical units is an indirect way to assess and identify clusters (see EC 2008, p. 16). On the contrary, Input-Output linkages provide a *direct* way of grouping activities, according to the strength of their monetary exchanges in the (re-)production process. Clusters so identified reflect *structural* relations, and not behavioural ones, i.e. they reflect how the single parts are mutually conditioned by the reproductive requirements of the whole inter-industry system. This view can be of great use to obtain an *objective* assessment of agglomeration patterns.

It would be surely interesting to perform a comparative analysis of both methodologies. Unfortunately, though, there is a data dimension problem. While the cluster template obtained by the European Cluster Observatory is formulated at a 4-digit level NACE Rev. 2 classification (corresponding to around 300 industries), European Supply, Use and Input-Output Tables (SUOIT, hereinafter) published by EUROSTAT are available at a 60-industry level of the NACE Rev. 1 classification.[6] This limitation suggests that to bridge EU cluster mapping results and Input-Output clusters aggregation schemes should be taken to a common level.

Nevertheless, the importance of devising alternative methods for the identification and measurement of cluster composition and relative weights within the EU economy remains, and advancing in this direction is the aim of the present paper.

We have identified industry clusters in EU27, EMU and the constituent national economies by means of a community detection strategy developed in the field of network theory: the spectral bisection algorithm (see Leicht & Newman 2008). This is the main methodological tool used throughout the paper, and is the point of departure of the analysis in Section 2. Afterwards, section 3 reports the results of community detection in EU27 and EMU Input-Output networks, while Section 4 frames the results in quantitative terms and assesses the productivity performance of individual European sectors.

In Section 5 we extend the analysis of Section 3 to the individual countries composing EU27 and EMU, with an emphasis on devising a similarity measure that results in a hierarchical ordering of countries' clustering structures, in order to grasp the underlying country-specific communities of activities.

[6]However, note that: "Up from reference year 2008, the revised NACE rev.2 will be applied" (EUROSTAT 2011, p. 4).

But clusters in each EU country are widely linked through external trade flows. In the search for connections between national economies within the EU, we have applied the spectral bisection algorithm to commodity-specific intra-EU trade matrices, in order to see whether persistent country-clustering patterns emerge. Results are reported in Section 6. A summary of findings and some concluding remarks follow in Section 7.

2 Methodology: Spectral Bisection algorithm for community detection

Throughout the paper, we make an extensive use of Spectral Bisection (SB, hereinafter) in order to identify industry and country clusters and sub-groups within clusters. Since all the empirical results derive from the choice of employing SB, it is worth devoting to it a few lines, in order to make clear the reasons for such a choice.

Leicht & Newman's (2008) SB algorithm for unweighted, *directed* graphs[7] can be straightforwardly generalised to take *weighted* flows into account. The logic at the basis of the algorithm runs as follows.

The starting point is the optimal partition of a network, which is defined as a division into indivisible subgraphs. In other words, the 'true' partition of a network into communities is found when all distinct communities have been detected, and thus none of them can be further divided into sub-communities. We are therefore indirectly provided with a definition of *community* as an *indivisible subgraph*. Any possible partition of the network has its associated modularity — a quantitative indicator of how good the partition is — whose maximum value is attained in correspondence to the optimal partition.

Consider a network with n nodes, represented by means of a *flow matrix* $\mathbf{F} = [f_{ij}]$, f_{ij} being the flows going from node i to j.[8] As a first step towards community detection, we define the kind of communities we are looking for. In what follows, we basically defined two kinds: (i) *flow-based* communities: groups of industries (when analysing inter-industry flows within a single country) or countries (when analysing trade flows between countries) connected by flows with above-than-average intensity; and (ii) *similarity-based* communities: groups of countries with a similar industry-clustering structure — or groups of commodities with a similar country-clustering structure of external trade.

With this distinction in mind, choosing the matrix to which SB has to

[7] A generalisation of Newman's (2006) algorithm for unweighted, undirected graphs.

[8] See Table 16 in Appendix A for a glossary of mathematical notation used throughout the paper.

be applied is straightforward: it will simply be the flow matrix $\mathbf{F}$ in the case of flow-based communities, and a *similarity matrix* $\mathbf{S} = [s_{ij}]$ in the case of similarity-based communities. In the latter case, there is a further preliminary step consisting in defining a similarity measure suitable to capture the characteristic to be compared among countries (or industries) and then computing it for each pair of nodes, i.e. computing the elements s_{ij}. In both cases, we get an $n \times n$ square matrix whose cells represent pair-wise connections between nodes, be them flows or similarities. Call this generic matrix $\mathbf{W}$.

Now, imagine we have some limited information about matrix $\mathbf{W}$: the totals by row $\mathbf{s}_{(out)} = \mathbf{We}$; the totals by column $\mathbf{s}^T_{(in)} = \mathbf{e}^T\mathbf{W}$; and the scalar grand total $m = \mathbf{e}^T\mathbf{We} = \sum_i s_{i(out)} = \sum_j s_{j(in)}$.[9]

On the basis of such limited information, we can exploit a bi-proportional average method to split row and column totals to get an averaged flow/similarity matrix, which we may call $\mathbf{W}^e$:[10]

$$w^e_{ij} = m\frac{s_{i(out)}}{m}\frac{s_{j(in)}}{m} = \frac{s_{i(out)}s_{j(in)}}{m}$$

or, in matrix terms:

$$\mathbf{W}^e = \frac{\mathbf{s}_{(out)}\mathbf{s}^T_{(in)}}{m}$$

The idea at the basis of modularity maximisation is that flows, or similarities, are in general greater than average within nodes in the same cluster, and below average otherwise.[11] The *modularity matrix* $\mathbf{B} = [b_{ij}]$ is given by the difference between actual and expected flow/similarity matrices:

$$\mathbf{B} = \mathbf{W} - \mathbf{W}^e$$

[9] In the case of flow-based communities we can interpret these magnitudes in terms of traditional network theory terminology: $\mathbf{s}_{(out)}$ are the total intermediate deliveries (exports) by industry (country) of origin, or industries' (countries') *out-strengths*; $\mathbf{s}^T_{(in)}$ are the total intermediate purchases (imports) by industry (country) of destination, or industries' (countries') *in-strengths*.

[10] Vector $\mathbf{s}^T_{(in)}/m$ gives us the proportions to split a row into columns, i.e. into single cells; *vice-versa*, vector $\mathbf{s}_{(out)}/m$ gives us an average rule to split a column into rows.

[11] There might be exceptions, especially in the case of flow-based communities for industries (countries) with exceptionally low or exceptionally high purchases/deliveries (exports/imports). The industries providing business services, for example, might deliver higher-then-average flows of inputs even to industries not belonging to their communities. On the contrary, industries with very few inter-industry connections can display lower-than-average exchanges even with industries in their same cluster.

Elements b_{ij} will in general be (the most) positive if nodes i and j actually belong to same cluster and (the most) negative otherwise. The original method for undirected graphs could not be applied to non-symmetric flow/similarity matrices. Leicht & Newman (2008) showed that it is possible to overcome this limitation by computing the *generalised modularity matrix*:

$$\tilde{\mathbf{B}} = \mathbf{B} + \mathbf{B}^T \tag{2.1}$$

Note that using this generalised matrix with undirected graphs as well does not change the results obtained by working instead with $\mathbf{B}$.

Take any initial (arbitrary) subdivision of the network in two communities α and β, and the *membership vector* $\mathbf{m} = [m_i]$, with $m_i = 1$ for $i \in \alpha$ and $m_i = -1$ for $i \in \beta$. *Modularity* can then be computed as a weighted sum of the $\tilde{b}_{ij}$'s:

$$Q = \mathbf{m}^T \tilde{\mathbf{B}} \mathbf{m} \tag{2.2}$$

the weights being $m_i m_j$, where $m_i m_j = +1$ if nodes i and j are assigned to the same community, $m_i m_j = -1$ otherwise. Therefore, correctly assigning two nodes (i) to the same group when they are actually in the same cluster; and (ii) to different groups when they actually belong to different clusters; improves modularity. On the contrary, incorrectly (i) separating nodes belonging to the same cluster; and (ii) grouping nodes belonging to different clusters; reduces modularity.

It can be shown (see Leicht & Newman 2008) that each industry can be assigned to community α or β according to the sign of the corresponding element of the leading eigenvector of matrix $\tilde{\mathbf{B}}$ (i.e. the eigenvector corresponding to the largest positive eigenvalue of the symmetric matrix).

After the first subdivision, the algorithm proceeds by further bisecting each resulting community, as long as such bisections lead to positive contributions to modularity. However, to be more precise, after the first iteration the procedure needs a slight modification. If we treated each subgraph as an autonomous network, we would disregard inter-modular connections, thus not being able to compute the modularity of the whole network, which depends on $\tilde{\mathbf{B}}$ and hence on the flows/similarities connecting *all* nodes. Rather, as stressed by Newman (2006), we compute the *additional contribution* to modularity (ΔQ) of each further bisection as:[12]

$$\Delta Q = \mathbf{m}_\alpha^T \tilde{\mathbf{B}}_\alpha \mathbf{m}_\alpha - \mathbf{m}_\alpha^T \widehat{(\tilde{\mathbf{B}}_\alpha \mathbf{e})} \mathbf{m}_\alpha = \mathbf{m}_\alpha^T (\tilde{\mathbf{B}}_\alpha - \widehat{(\tilde{\mathbf{B}}_\alpha \mathbf{e})}) \mathbf{m}_\alpha = \mathbf{m}_\alpha^T \tilde{\mathbf{B}}^{(\alpha)} \mathbf{m}_\alpha \tag{2.3}$$

Furthermore, after each consecutive bisection, a fine tuning of the result is performed; it consists in taking each couple of groups and moving each

[12]The presence of subscript α indicates those vectors ($\mathbf{m}_\alpha^T$, $\mathbf{m}_\alpha$) and matrices ($\tilde{\mathbf{B}}_\alpha$) composed only by the elements assigned to community α.

Table 1: Spectral Bisection (SB) Algorithm

1)	Compute the generalised modularity matrix $\tilde{\mathbf{B}}$ as in (2.1)
2)	Compute the Perron eigenvector of matrix $\tilde{\mathbf{B}}$: $\mathbf{z}^{\star}(\tilde{\mathbf{B}})$
3)	Bisect the network according to the sign of the elements of $\mathbf{z}^{\star}(\tilde{\mathbf{B}})$. Compute modularity Q as in (2.2)
4)	Fine Tuning: take the first node and move it to the other group. Compute modularity. Store the new bisection and the corresponding Q
5)	Repeat 4) for each single node. Stop when all nodes, one at a time and only once, have been moved from one group to the other
6)	Pick the maximum Q, and save the corresponding partition
7)	For each group g in the partition:
	a) Compute the modified modularity submatrix $\tilde{\mathbf{B}}^{(g)}$
	b) Compute the Perron eigenvector of matrix $\tilde{\mathbf{B}}^{(g)}$: $\mathbf{z}^{\star}(\tilde{\mathbf{B}}^{(g)})$
	c) Bisect the group according to the sign of the elements of $\mathbf{z}^{\star}(\tilde{\mathbf{B}}^{(g)})$. Compute contribution to modularity ΔQ as in (2.3)
	d) Fine Tuning. Choose the bisection associated to $\max(\Delta Q)$ and save the corresponding partition
8)	Repeat 7) while $\Delta Q > 0$.

node, one at a time and only once, to the other group, computing the associated modularity; then, within the set of intermediate states occupied by network, the one associated to the maximum value of modularity is chosen. The procedure of iteratively moving nodes from one group to the other is then repeated until no increase in modularity can be reached, ending the SB procedure. For a summary of the steps of the algorithm, see Table 1.

Some criticisms to modularity maximisation through SB as a community detection strategy might be summarised as follows: (i) modularity is not a good indicator of the goodness of a partition; (ii) the SB approach provides a partition of the network in the mathematical sense of the word: overlapping communities are not allowed for (see McNerney 2009, pp. 9-10); and (iii) this kind of algorithm always provides a partition, even if a well-defined clustering of the network does not exist.

As to point (i), it is worth stressing that many community detection methods known and used in recent Input-Output literature[13] are based on the identification of above-average direct flows between nodes. However, the very concept of 'above-average' is somewhat arbitrary, since it is usually associated to the choice of an exogenous level of significance to be used as cut-off point for above-average intermediate deliveries/purchases. The criteria of choice for such significance levels are not clear nor well defined, and thus the results are subject to a high degree of arbitrariness. The identification of

[13]For a review see, e.g. Santos, Almeida & Teixeira (2008) and Hoen (2002).

clusters through SB is also based on pinpointing above-average connections: $\mathbf{W}^e$ is a matrix of benchmark levels, one for each pair-wise flow (or similarity level) between nodes, though these are computed according to an *objective* criterion. In this sense, SB is superior to other clustering methods.

As to point (ii), it is our contention that ruling out overlapping communities, far from being a drawback, is advantageous in many respects. First, as a result of SB, it is possible to get isolated nodes. 'Clusters' of this kind have to be interpreted as nodes which do not belong to a specific group, but share strong links with more than one; their role is to connect different communities, which would be the same role played by nodes at the intersection of overlapping communities. The presence of the same node in more than one group, on the contrary, would be in contradiction with the definition of community itself: in the limit, we could have a node shared by all communities, implying that it has 'special' connections with *none* of them. Second, considering overlapping clusters would rule out the possibility of getting advantage of the additivity of some linear operators, which is essential, e.g., for mapping communities into vertical (hyper-)integrated sectors or subsystems (see Section 4 below).

As to the last point (iii), there might be cases in which no significant division in subgraphs is found, and thus maximum modularity is associated with the network as a whole.[14]

As a way of conclusion, let us stress some additional advantages of SB.

The first argument is of a general character. SB allows to identify both flow-based and similarity-based communities, while other algorithms working fine for flow matrices cannot be applied to similarity matrices, and *vice-versa*. The standard procedure for getting communities out of similarity (or distance) matrices is agglomerative hierarchical clustering, which consists in picking the cell with the highest value, agglomerating the two corresponding nodes, updating the matrix by computing the similarity between the new formed node and all the others, and iterating this process up to the point where the last agglomeration leads to the coverage of the complete network.[15]

The second argument specifically refers to the peculiarities of the flow matrix associated to an Input-Output (IO, hereinafter) network, as opposed to standard biological, social, physical, chemical and other networks usually being the object of network theory.

A most important peculiar feature of IO networks, besides the presence of self-loops and the fact that flows are both directed and weighted, is the

[14] See Newman (2006) for a formal proof.

[15] An exhaustive review of agglomeration methods has been recently provided by Murtagh & Contreras (2011).

presence of boundary flows: final demand and value added (see McNerney 2009, p. 1). The two boundaries respectively represent the 'entrance to' and 'exit from' the network.

Despite the fact that each flow f_{ij} can be alternatively seen either as a material flow (corresponding to the value at current prices of deliveries) from industry i to j, or as a monetary flow (corresponding to the payment received as a counterpart to such deliveries) from industry j to i, flow direction becomes relevant when considering the boundaries, as it defines the direction of causation. In our view, final demand is the 'entrance' — since the whole process of production is demand-driven: production aims at satisfying final demand, or the demand for intermediate products stimulated by it — and value added the 'exit' — a fraction of the monetary flows generated by each (final or intermediate) delivery immediately exits the inter-industry network becoming national (or foreign, in the case of imports) income.

We are, thus, in front of a 'dissipative' network: indirect flows become smaller and smaller, and they finally converge to zero — in so doing guaranteeing the convergence of the 'Leontief inverse' infinite power series. Some community detection techniques [16] specifically refer to non-dissipative networks, whose flow matrices can be reduced to row- or column-stochastic matrices, the elements of which can be given a probabilistic interpretation and whose Perron eigenvalue is always equal to 1. In our case, once the boundaries are excluded, such a possibility is excluded too; the accounting identity between row and column sums disappears, and the probabilistic interpretation vanishes too. These methodologies, therefore, cannot be applied, while SB still preserves its validity, since it simply compares a given matrix with the artificial one obtained by knowing only row and column sums.

There are, of course, other methodologies which might in principle be applied to a generic IO matrix without boundaries;[17] almost all of them have been developed without reference to a specific kind of network, and are based on engineering algorithms whose deterministic logic is remorseless, but are way too roundabout to be given a straightforward economic interpretation. All these methodologies lead to different results, i.e. different partitions of the network, and a choice among them can hardly be made according to economic arguments. SB has the invaluable advantage of working on the basis of an extremely simple logic and of a relatively smooth implementation. This makes its functioning clear and apparent, and allows a clear-cut identification of eventual biases in the results.

[16] See as an illustrative example Piccardi's (2011) methodology based on Lumped Markov Chains.

[17] See Rosvall, Axelsson & Bergstrom's (2009) article for an illustrative example, and Radicchi, Castellano, Cecconi, Loreto & Parisi's (2004) for an exhaustive review.

Let us now turn to the way in which SB has been applied in the present work. As mentioned above, we have been looking both for flow-based communities, and for similarity-based ones.

The empirical roadmap

First, we looked for industry clusters in the aggregate EU27 and EA17[18] inter-industry networks for years 2000 and 2007, using the intermediate consumption part corresponding to each country-group Input-Output table for domestic output at basic prices as flow matrices.[19] By applying the SB algorithm we obtained, for each aggregate and for each year, a partition into communities.

Then, we inspected relations between communities, on the one side, and relations between industries within the same community, on the other. In the first case, we aggregated the IO matrix according to the clustering structure obtained, and iteratively applied SB aggregating the resulting communities up to the point where no further aggregation was possible. In the second case, we took in isolation each of the community sub-matrices and iteratively applied SB up to the point where no further aggregation was possible in each of them.[20] Results are shown in Section 3.

As a second step, we applied SB to individual countries' square IO tables for domestic output at basic prices provided by EUROSTAT for 22 out of the EU27 member states for year 2005,[21] obtaining their corresponding communities of industries. For each country, we stored the results in a membership vector, and assembled all vectors into a membership matrix. On the basis of such a matrix, we advanced and computed a similarity measure between individual countries' clustering structures and built the associated similarity matrix. We applied SB to *this* matrix, and obtained similarity-based communities of national economies, *according to* how their industry clusters differ. Also in this case, we looked for relations between and within communities (by updating, rather than aggregating, the initial similarity matrix/sub-matrices). The results are shown in Section 5.

[18]A list of EU27 and EA17 (or EMU) country memberships is reported in Table 17 of Appendix B.

[19]For a description of the construction of EU27 and EMU Input-Output tables, see EUROSTAT (2011).

[20]It should be borne in mind that this procedure is not aimed at finding sub-clusters (a cluster being defined as an indivisible subgraph), but simply at qualitatively looking for particularly strong linkages within and between communities.

[21]A list of data availability by country for these computations is reported in Table 18 of Appendix B.

As a third step, we looked for a clustering structure in intra-European external trade by applying SB to total inter-country trade matrices, based on data obtained from EUROSTAT, for the years 2007-2010. We found a partition into communities of national economies and then looked for particularly intense trade relations between communities and among countries within communities.

Finally, we applied the same procedure, but starting instead from individual trade matrices for a subset of CPA commodities. We applied SB to each of these flow matrices looking for communities of *countries* characterised by particularly strong trade relations in each of the *commodities*. We stored the resulting clustering structure of countries in a membership matrix, and on the basis of it we applied our similarity measure between couples of commodities, and looked for similarity-based clusters, *according to* how commodity-specific trade networks among countries differ. Results are reported in Section 6.

3 Community detection in European IO networks

As a general feature of the application of the SB algorithm to EMU (or EA17) and EU27 IO networks — disaggregated into 59 industries[22] — for years 2000 and 2007, we found 12 industry groups whose activities are always grouped together.[23] We called these groups 'mini-clusters', and report them in Table 2.

Consider now Table 3, which reports the clustering of industries resulting from the application of SB to each of the four IO networks analysed (EMU 2000, EMU 2007, EU27 2000 and EU27 2007). This information can be used together with Table 4, which reports — for each network — what is the effect upon its own cluster of removing a given node (industry) and applying the SB algorithm again. In this way, removing one industry at a time might break, leave unaltered or enlarge (B, U and L in Table 4, respectively) the original cluster to which the node belongs, shedding light on how central each node is to its cluster within each network.

Furthermore, hierarchical relations within communities ('downstream' relations) have been obtained by the iterative application of SB to cluster-specific sub-matrices, aggregating nodes grouped together until no further aggregation was possible. Instead, to study hierarchical relations among

[22] See Table 19 of Appendix B for a glossary and description of the disaggregation scheme adopted.

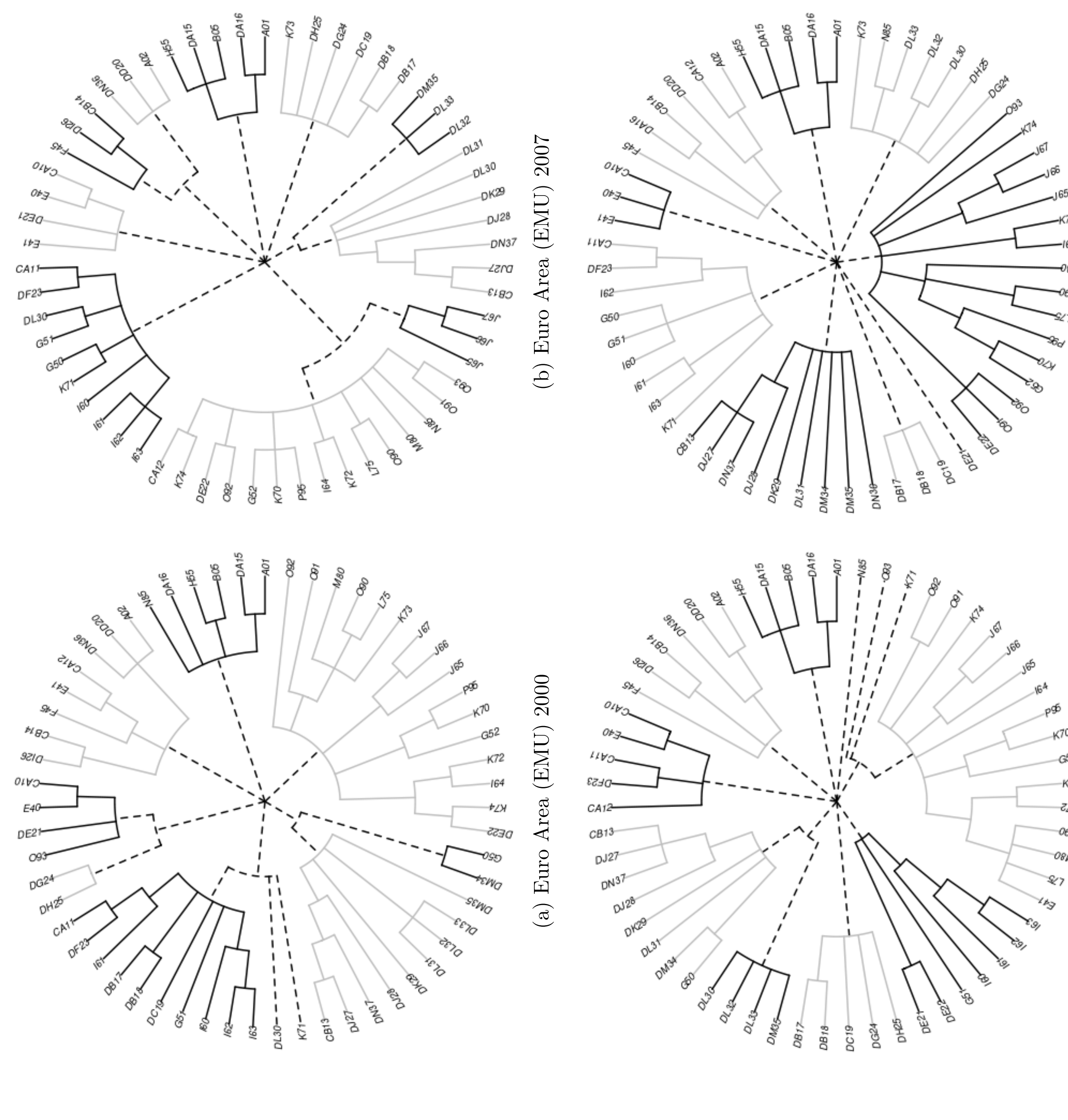
[23] This does not mean that these clusters are identically repeated in each network, but that each of these groups is the minimum subset of industries which are always grouped together in all networks, probably in clusters of differing global composition.

clusters (‘upstream’ relations), community sub-matrices have been aggregated to a single cell each, and *then* SB has been applied to detect nodes (now representing whole clusters) grouped together, iterating until no further aggregation was possible. Figures 1a,1b, 1c and 1d display the resulting dendrograms for each network, obtaining a quick visualisation of the whole clustering structure in each case.

Table 2: EMU/EU27 Persistent Mini-Clusters results (years: 2000, 2007)
(Groups of industries assigned to the same cluster for both country-groups and in all years)

Agri-Food	**Wood**	**Energy**
A01 Agriculture	A02 Forestry	CA10 Coal Mining
B05 Fishing	DD20 Wood	E40 Electricity-gas
DA15 Food-beverages		
DA16 Tobacco	**Construction**	**Petroleum-Gas**
H55 Hotel-restaurant	CB14 Stone-sand-clay-minerals	CA11 Petroleum-gas-extraction
	DI26 Glass-clay-cement-ceramic	DF23 Petroleum-refinery
Metal-Machinery	F45 Construction	
CB13 Metal-mining		**Chemicals**
DJ27 Iron-steel-aluminium-tub.	**Dressing**	DG24 Chemicals-pharma
DJ28 Structural-metal-products	DB17 Textiles	DH25 Rubber-plastics
DK29 Mechanical-machinery	DB18 Clothing	
DL31 Electrical-machinery	DC19 Leather	**Special-Machinery**
DN37 Recycling		DL32 ICT-equipment
	Services	DL33 Medical-precision-equip.
Transport-Trade	G52 Retail-trade	
G51 Wholesale-trade	I64 Post-telecomm.	**Finance**
I60 Transport-land	K70 Real-estate	J65 Finance
I61 Transport-water	K72 Computer-services	J66 Insurance
I62 Transport-air	K74 Business-services	J67 Brokerage-credit-cards
I63 Storage-travel-agencies	L75 Public-admin.	
	M80 Education	
	O90 Refuse-disposal	
	O91 Membership-organisations	
	O92 Arts-entertainment	
	P95 Household-services	

Source: Own computation based on EUROSTAT SUIOT Database.



(a) Euro Area (EMU) 2000

(b) Euro Area (EMU) 2007

(c) European Union (EU27) 2000

(d) European Union (EU27) 2000

Figure 1: Clustering structure of the European IO networks (2000, 2007)

Table 3: EMU/EU27 Clustering results (years 2000, 2007)

EMU, 2000	EMU, 2007	EU27, 2000	EU27, 2007
(C01) Agri-Food	**(C01) Agri-Food**	**(C01) Agri-Food**	**(C01) Agri-Food**
A01 Agriculture	A01 Agriculture	A01 Agriculture	A01 Agriculture
B05 Fishing	B05 Fishing	B05 Fishing	B05 Fishing
DA15 Food-beverages	DA15 Food-beverages	DA15 Food-beverages	DA15 Food-beverages
DA16 Tobacco	DA16 Tobacco	DA16 Tobacco	DA16 Tobacco
H55 Hotel-restaurant	H55 Hotel-restaurant	H55 Hotel-restaurant	H55 Hotel-restaurant
N85 Health			
(C02) Construction	**(C07) Construction**	**(C02) Construction**	**(C02) Construction**
F45 Construction	F45 Construction	F45 Construction	F45 Construction
CB14 Stone-sand-clay-minerals	CB14 Stone-sand-clay-minerals	CB14 Stone-sand-clay-minerals	CB14 Stone-sand-clay-minerals
DI26 Glass-clay-cement-ceramic	DI26 Glass-clay-cement-ceramic	DI26 Glass-clay-cement-ceramic	DI26 Glass-clay-cement-ceramic
E41 Water		A02 Forestry	A02 Forestry
CA12 Uranium	**(C02) Furniture**	DD20 Wood	DD20 Wood
A02 Forestry	A02 Forestry	DN36 Furniture-Sports-Toys	CA12 Uranium
DD20 Wood	DD20 Wood		
DN36 Furniture-Sports-Toys	DN36 Furniture-Sports-Toys	**(C03) Energy**	**(C03) Energy**
		CA10 Coal Mining	CA10 Coal Mining
(C03) Energy	**(C03) Energy**	E40 Electricity-gas	E40 Electricity-gas
CA10 Coal Mining	CA10 Coal Mining	CA11 Petroleum-gas-extraction	E41 Water
E40 Electricity-gas	E40 Electricity-gas	DF23 Petroleum-refinery	
O93 Personal-services	E41 Water	CA12 Uranium	**Isolated**
DE21 Paper	DE21 Paper		DE21 Paper (C07)
		(C05) Dressing-Chemicals	
(C07) Chemicals	**(C08) Dressing-Chemicals**	DG24 Chemicals-pharma	**(C06) Dressing**
DG24 Chemicals-pharma	DG24 Chemicals-pharma	DH25 Rubber-plastics	DB17 Textiles
DH25 Rubber-plastics	DH25 Rubber-plastics	DB17 Textiles	DB18 Clothing

(continued on next page)

(continued from previous page)

EMU, 2000	**EMU, 2007**	**EU27, 2000**	**EU27, 2007**
	K73 R&D	DB18 Clothing	DC19 Leather
(C04) Transport-Dressing	DB17 Textiles	DC19 Leather	
DB17 Textiles	DB18 Clothing		**(C04) Transport-Trade**
DB18 Clothing	DC19 Leather	**(C06) Transport-Media**	CA11 Petroleum-gas-extraction
DC19 Leather		I60 Transport-land	DF23 Petroleum-refinery
I60 Transport-land	**(C04) Transport-Trade**	I61 Transport-water	I60 Transport-land
I61 Transport-water	I60 Transport-land	I62 Transport-air	I61 Transport-water
I62 Transport-air	I61 Transport-water	I63 Storage-travel-agencies	I62 Transport-air
I63 Storage-travel-agencies	I62 Transport-air	G51 Wholesale-trade	I63 Storage-travel-agencies
CA11 Petroleum-gas-extraction	I63 Storage-travel-agencies	DE21 Paper	G50 Sale-repair-vehicles
DF23 Petroleum-refinery	CA11 Petroleum-gas-extraction	DE22 Publishing-printing	G51 Wholesale-trade
G51 Wholesale-trade	DF23 Petroleum-refinery		K71 Renting-equipment
	G51 Wholesale-trade	**(C07) Specialised Machin.**	
Isolated	G50 Sale-repair-vehicles	DM35 Ships-railway-aircrafts	**(C09) Pharma-Hi Tech**
K71 Renting-equipment (C10)	K71 Renting-equipment	DL32 ICT-equipment	DL32 ICT-equipment
DL30 Office-machinery-PC (C08)	DL30 Office-machinery-PC	DL30 Office-machinery-PC	DL30 Office-machinery-PC
		DL33 Medical-precision-equip.	DL33 Medical-precision-equip.
(C05) Heavy-HiTech Mach.	**(C09) Specialised Machin.**		DG24 Chemicals-pharma
DL32 ICT-equipment	DL32 ICT-equipment	**Isolated**	DH25 Rubber-plastics
DL33 Medical-precision-equip.	DL33 Medical-precision-equip.	N85 Health (C10)	N85 Health
DM35 Ships-railway-aircrafts	DM35 Ships-railway-aircrafts	O93 Personal-services (C11)	K73 R&D
CB13 Metal-mining		K71 Renting-equipment (C09)	
DJ27 Iron-steel-aluminium-tub.	**(C06) Heavy Machinery**		**(C05) Heavy Machinery**
DJ28 Structural-metal-products	CB13 Metal-mining	**(C04) Heavy Machinery**	DN36 Furniture-Sports-Toys
DK29 Mechanical-machinery	DJ27 Iron-steel-aluminium-tub.	CB13 Metal-mining	CB13 Metal-mining
DL31 Electrical-machinery	DJ28 Structural-metal-products	DJ27 Iron-steel-aluminium-tub.	DJ27 Iron-steel-aluminium-tub.
DN37 Recycling	DK29 Mechanical-machinery	DJ28 Structural-metal-products	DJ28 Structural-metal-products
	DL31 Electrical-machinery	DK29 Mechanical-machinery	DK29 Mechanical-machinery

(continued on next page)

(continued from previous page)

EMU, 2000	EMU, 2007	EU27, 2000	EU27, 2007
(C09) Motor-vehicles	DN37 Recycling	DL31 Electrical-machinery	DL31 Electrical-machinery
DM34 Motor-vehicles	DM34 Motor-vehicles	DM34 Motor-vehicles	DM34 Motor-vehicles
G50 Sale-repair-vehicles		DN37 Recycling	DN37 Recycling
	(C05) Services	G50 Sale-repair-vehicles	DM35 Ships-railway-aircrafts
	CA12 Uranium		
(C06) Services	N85 Health	**(C08) Services**	**(C08) Services**
K73 R&D	O93 Personal-services	G52 Retail-trade	G52 Retail-trade
G52 Retail-trade	G52 Retail-trade	I64 Post-telecomm.	I64 Post-telecomm.
DE22 Publishing-printing	DE22 Publishing-printing	J65 Finance	J65 Finance
I64 Post-telecomm.	I64 Post-telecomm.	J66 Insurance	J66 Insurance
K70 Real-estate	K70 Real-estate	J67 Brokerage-credit-cards	J67 Brokerage-credit-cards
K72 Computer-services	K72 Computer-services	K70 Real-estate	K70 Real-estate
K74 Business-services	K74 Business-services	K72 Computer-services	K72 Computer-services
L75 Public-admin.	L75 Public-admin.	K74 Business-services	K74 Business-services
M80 Education	M80 Education	L75 Public-admin.	L75 Public-admin.
O90 Refuse-disposal	O90 Refuse-disposal	M80 Education	M80 Education
O91 Membership-organisations	O91 Membership-organisations	O90 Refuse-disposal	O90 Refuse-disposal
O92 Arts-entertainment	O92 Arts-entertainment	O91 Membership-organisations	O91 Membership-organisations
P95 Household-services	P95 Household-services	O92 Arts-entertainment	O92 Arts-entertainment
J65 Finance		P95 Household-services	P95 Household-services
J66 Insurance	**(C10) Finance**	E41 Water	O93 Personal-services
J67 Brokerage-credit-cards	J65 Finance	K73 R&D	DE22 Publishing-printing
	J66 Insurance		
	J67 Brokerage-credit-cards		

Source: Own computation based on EUROSTAT SUIOT Database

Let us first compare the EMU clustering structure between 2000 and 2007.

Cluster C02 Construction breaks up in 2007 into C02 Furniture and C07 Construction. By looking at Table 4, we can see that there are three activities which, if missing, ‘break’ the Construction cluster in 2000: F45 Construction, DN36 Furniture-Sports-Toys and DD20 Wood. From Figure 2a, it is apparent that the link which weakened and lead to a separation is DD20 Wood $\rightarrow$ F45 Construction.

Another important difference is that cluster C04 Transport-Dressing broke up in 2007, partly into C08 Dressing-Chemicals (see Figure 2c) — including cluster C07 Chemicals of 2000 — and partly into C04 Transport-Trade — including the two isolated nodes K71 Renting-equipment and DL30 Office-machinery-PC of 2000.

It emerges from Figure 2b that — in 2000 — MC:Dressing[24] was connected to MC:Transport-Trade through nodes I62 Transport-air and, especially, G51 Wholesale-trade — both identified as breaking cluster C04 Transport-Dressing in 2000; while it was connected to MC:Chemicals in 2007, basically through edges DG24 Chemicals-pharma $\rightarrow$ DB17 Textiles $\rightarrow$ DH25 Rubber-plastics. Hence, in 2007, this latter connection became stronger relatively to the first one, and MC:Chemicals and MC:Dressing converged into cluster C08 Dressing-Chemicals.

A third difference is represented by the distribution of the metal-machinery group, broadly intended, into clusters C05 Heavy-HiTech Mach. and C09 Motor-vehicles in 2000, and C09 Specialised Machin. and C06 Heavy Machinery in 2007 (see Figure 2e).

It seems apparent, by looking at Table 4, that cluster C09 Motor-vehicles broke up due to the switch of G50 Sale-repair-vehicles, in 2007, to cluster C04 Transport-Trade. As a consequence industry DM34 Motor-vehicles converged into C05 Heavy-HiTech Mach.. As to cluster C09 Specialised Machin., it emerges from figure 2d and can be noticed from Table 4, that DJ27 Iron-steel-aluminium-tub. and DK29 Mechanical-machinery break cluster C05 Heavy-HiTech Mach. in 2000. In fact, the connections which became relatively weaker are the edges between the first (DJ27) and DM35 Ships-railway-aircrafts and between the second (DK29) and DL32 ICT-equipment/DL33 Medical-precision-equip., inducing a separation of these three industries into an own — and quite interdependent, as displayed in Figure 2f — cluster.

It can be further noticed that cluster C06 Services in 2000 also broke into C05 Services and C10 Finance in 2007. In this case, simple inspection of Table 4 and the corresponding graphs is not enough to pinpoint the missing

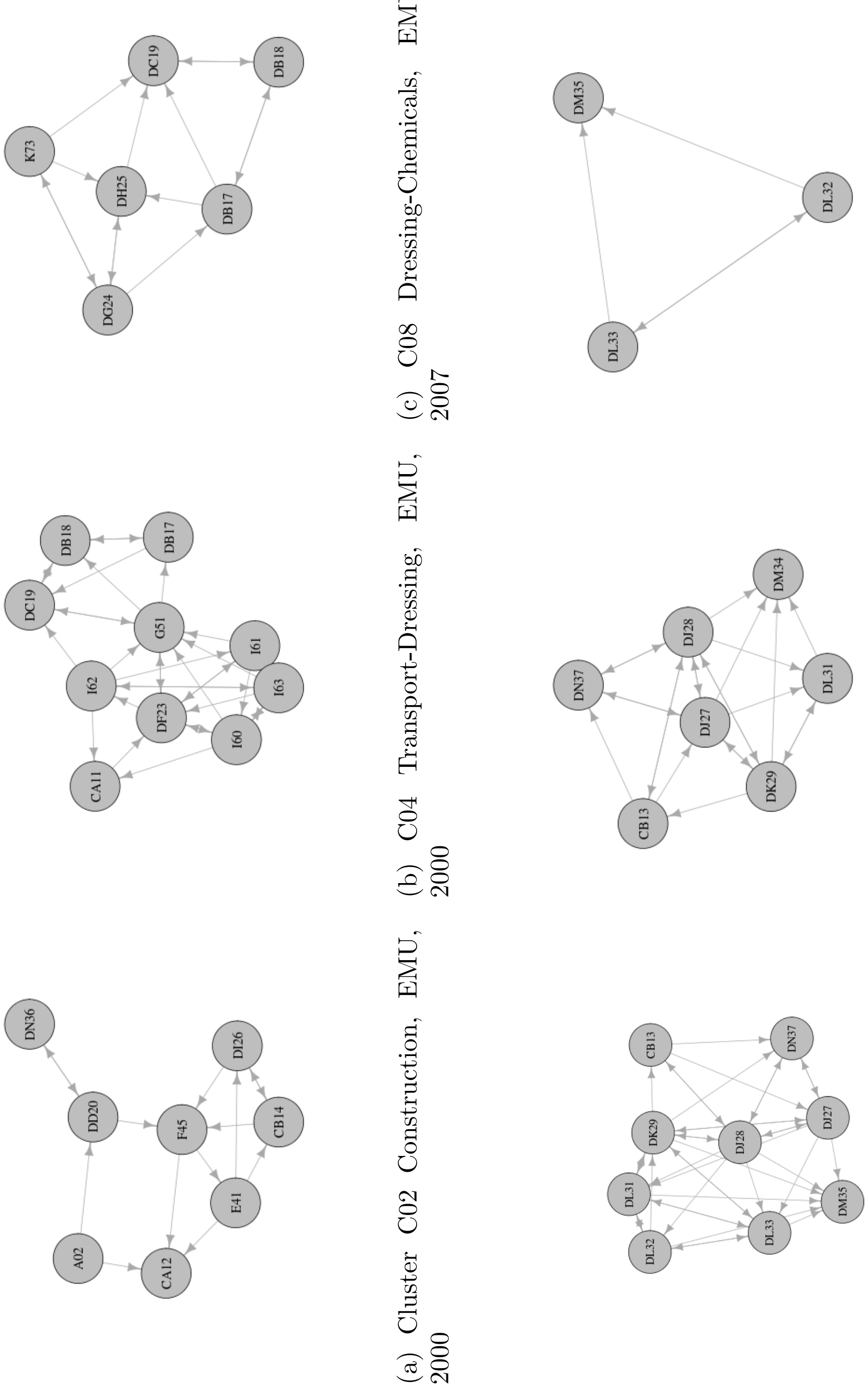
[24]MC stands for Mini-Cluster. Persistent mini-clusters are reported in Table 2.

Table 4: Cluster breakers in the EA17 and the EU27 (years 2000, 2007)

		EA17 2000				EA17 2007				EU27 2000				EU27 2007			
NACE	Industry	Cls.	B	U	L	Cls.	B	U	L	Cls.	B	U	L	Cls.	B	U	L
A01	Agriculture	C01	✓			C01	✓			C01	✓			C01	✓		
A02	Forestry	C02		✓		C02			✓	C02		✓		C02	✓		
B05	Fishing	C01		✓		C01		✓		C01		✓		C01		✓	
CA10	Coal Mining	C03		✓		C03		✓		C03		✓		C03		✓	
CA11	Petroleum-gas-extraction	C04		✓		C04	✓			C03	✓			C04	✓		
CA12	Uranium	C02		✓		C05		✓		C03		✓		C02		✓	
CB13	Metal-mining	C05		✓		C06		✓		C04		✓		C05			✓
CB14	Stone-sand-clay-minerals	C02		✓		C07		✓		C02		✓		C02	✓		
DA15	Food-beverages	C01	✓			C01	✓			C01	✓			C01	✓		
DA16	Tobacco	C01		✓		C01		✓		C01		✓		C01		✓	
DB17	Textiles	C04		✓		C08	✓			C05	✓			C06			✓
DB18	Clothing	C04	✓			C08	✓			C05	✓			C06			✓
DC19	Leather	C04		✓		C08		✓		C05		✓		C06			✓
DD20	Wood	C02	✓			C02	✓			C02	✓			C02	✓		
DE21	Paper	C03		✓		C03		✓		C06	✓			C07			
DE22	Publishing-printing	C06		✓		C05	✓			C06	✓			C08		✓	
DF23	Petroleum-refinery	C04	✓			C04	✓			C03		✓		C04	✓		
DG24	Chemicals-pharma	C07			✓	C08	✓			C05	✓			C09	✓		
DH25	Rubber-plastics	C07			✓	C08	✓			C05	✓			C09	✓		
DI26	Glass-clay-cement-ceramic	C02		✓		C07			✓	C02	✓			C02		✓	
DJ27	Iron-steel-aluminium-tub.	C05	✓			C06	✓			C04	✓			C05	✓		
DJ28	Structural-metal-products	C05			✓	C06		✓		C04	✓			C05	✓		
DK29	Mechanical-machinery	C05	✓			C06		✓		C04	✓			C05	✓		
DL30	Office-machinery-PC	C08				C04		✓		C07		✓		C09	✓		
DL31	Electrical-machinery	C05		✓		C06		✓		C04	✓			C05		✓	
DL32	ICT-equipment	C05		✓		C09			✓	C07	✓			C09	✓		
DL33	Medical-precision-equip.	C05		✓		C09			✓	C07			✓	C09	✓		
DM34	Motor-vehicles	C09			✓	C06			✓	C04	✓			C05	✓		
DM35	Ships-railway-aircrafts	C05		✓		C09			✓	C07			✓	C05		✓	
DN36	Furniture-Sports-Toys	C02	✓			C02			✓	C02	✓			C05		✓	
DN37	Recycling	C05		✓		C06		✓		C04		✓		C05		✓	
E40	Electricity-gas	C03	✓			C03	✓			C03	✓			C03	✓		
E41	Water	C02		✓		C03		✓		C08		✓		C03		✓	
F45	Construction	C02	✓			C07		✓		C02			✓	C02	✓		
G50	Sale-repair-vehicles	C09		✓		C04		✓		C04	✓			C04		✓	
G51	Wholesale-trade	C04	✓			C04	✓			C06	✓			C04		✓	
G52	Retail-trade	C06	✓			C05	✓			C08		✓		C08		✓	
H55	Hotel-restaurant	C01		✓		C01		✓		C01		✓		C01		✓	
I60	Transport-land	C04	✓			C04	✓			C06		✓		C04	✓		
I61	Transport-water	C04		✓		C04		✓		C06		✓		C04		✓	
I62	Transport-air	C04	✓			C04	✓			C06		✓		C04		✓	
I63	Storage-travel-agencies	C04	✓			C04	✓			C06	✓			C04		✓	
I64	Post-telecomm.	C06	✓			C05	✓			C08	✓			C08		✓	
J65	Finance	C06		✓		C10		✓		C08	✓			C08		✓	
J66	Insurance	C06		✓		C10			✓	C08		✓		C08		✓	
J67	Brokerage-credit-cards	C06		✓		C10			✓	C08		✓		C08		✓	
K70	Real-estate	C06			✓	C05	✓			C08	✓			C08	✓		
K71	Renting-equipment	C10				C04		✓		C09				C04		✓	
K72	Computer-services	C06	✓			C05	✓			C08		✓		C08		✓	
K73	R&D	C06		✓		C08		✓		C08		✓		C09	✓		
K74	Business-services	C06	✓			C05	✓			C08	✓			C08		✓	
L75	Public-admin.	C06	✓			C05	✓			C08		✓		C08			✓
M80	Education	C06	✓			C05		✓		C08		✓		C08		✓	
N85	Health	C01		✓		C05		✓		C10				C09	✓		
O90	Refuse-disposal	C06		✓		C05	✓			C08		✓		C08		✓	
O91	Membership-organisations	C06		✓		C05	✓			C08		✓		C08		✓	
O92	Arts-entertainment	C06		✓		C05	✓			C08		✓		C08		✓	
O93	Personal-services	C03		✓		C05		✓		C11				C08		✓	
P95	Household-services	C06		✓		C05		✓		C08		✓		C08		✓	

References: Removing the industry breaks (**B**), leaves unaltered (**U**), or enlarges (**L**) the cluster.

(Blanks in all three (B, U, L) correspond to isolated industries)



(a) Cluster C02 Construction, EMU, 2000

(b) C04 Transport-Dressing, EMU, 2000

(c) C08 Dressing-Chemicals, EMU, 2007

(d) C05 Heavy-HiTech Mach., EMU, 2000

(e) C06 Heavy Machinery, EMU, 2007

(f) C09 Specialised Machin., EMU, 2007

Figure 2: Selected Graphs from EMU Clusters (2000, 2007)

edge(s) that broke up the initial unified cluster, as it is a densely interconnected one. However, it is possible to see from Figure 1a that already in 2000 the MC:Finance was particularly connected within cluster C06 Services, thus it is most likely that such connections became relatively stronger causing the separation of the mini-cluster from the rest of services.

Finally, from Figures 1a and 1b it can be inferred that, in 2000, C04 Transport-Dressing groups with DL30 Office-machinery-PC and K71 Renting-equipment, C09 Motor-vehicles groups with C05 Heavy-HiTech Mach.; while in 2007, C05 Services groups with C10 Finance, C06 Heavy Machinery groups with C09 Specialised Machin., and C02 Furniture groups with C07 Construction.

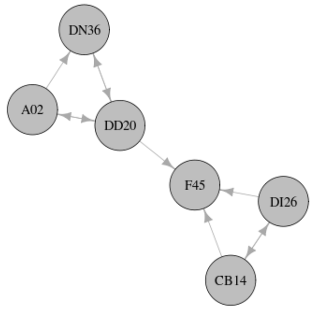
Let us now look at the differences in EU27 clustering structures between 2000 and 2007.

Consider first cluster C02 Construction, which coincides in both years except for the presence of industry CA12 Uranium (in 2007) instead of DN36 Furniture-Sports-Toys (which in 2007 moved to cluster C05 Heavy Machinery). According to Table 4, this latter industry breaks its cluster in 2000, and so do DD20 Wood and DI26 Glass-clay-cement-ceramic (as can be seen from Figure 3a).

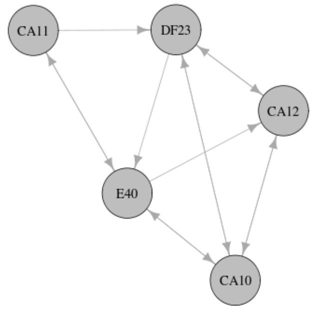
Second, even though a core of cluster C03 Energy consists of MC:Energy, as to the remaining industries the groups are quite different: in 2000 we also have CA11 Petroleum-gas-extraction, CA12 Uranium and DF23 Petroleum-refinery, while in 2007 we only have E41 Water. Table 4 identifies CA11 Petroleum-gas-extraction as a node that would break the cluster C03 Energy in 2000. By looking at Figures 3b and 1c it is apparent that DF23 Petroleum-refinery is particularly connected to CA11 Petroleum-gas-extraction. Altogether, these facts suggest that the edge whose weakening caused the break of cluster C03 Energy is precisely CA11 Petroleum-gas-extraction $\rightarrow$ DF23 Petroleum-refinery.

In each year, the Heavy Machinery cluster (C04 in 2000 and C05 in 2007) includes MC:Metal-Machinery plus industry DM34 Motor-vehicles. Beyond this core, we have industry G50 Sale-repair-vehicles in the 2000 cluster and industries DM35 Ships-railway-aircrafts and DN36 Furniture-Sports-Toys in the 2007 cluster (see Figures 3d and 3e).

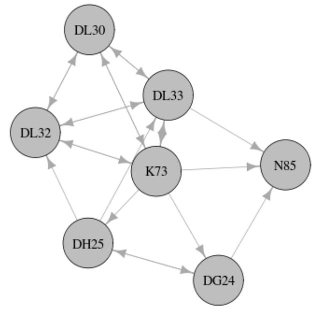
Finally, a most interesting configuration arising in the EU27 network for year 2007 corresponds to the breaking up of cluster C05 Dressing-Chemicals (actually consisting of two mini-clusters: MC:Dressing and MC:Chemicals), partly into a Dressing cluster (C06), and partly into a *new* type of cluster: by combining the MC:Chemicals mini-cluster with cluster C07 Specialised Machin. (of year 2000, excluding DM35 Ships-railway-aircrafts), and incorporating industries K73 R&D and N85 Health, cluster C09 Pharma-Hi Tech



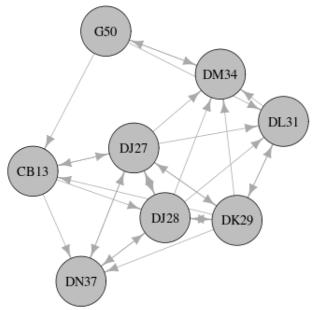
(a) Cluster C02 Construction, EU27, 2000



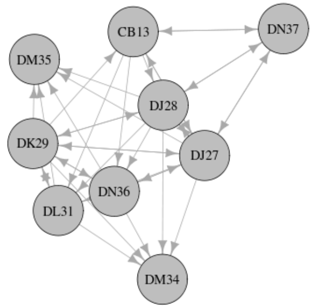
(b) Cluster C03 Energy, EU27, 2000.



(c) Cluster C09 Pharma-Hi Tech, EU27, 2007



(d) Cluster C04 Heavy Machinery, EU27, 2000



(e) Cluster C05 Heavy Machinery, EU27, 2007

Figure 3: Selected Graphs from EU27 Clusters (2000, 2007)

has been obtained as a result of the SB algorithm. As can be seen from Figure 3c and Table 4, this is a quite interconnected cluster: the removal of *any* node would break it.

To complete the picture for EU27, let us add that in 2000 (see Figure 1c) there are particularly strong connections between clusters C04 Heavy Machinery and C07 Specialised Machin. and between C08 Services and the isolated nodes K71 Renting-equipment and O93 Personal-services. On the contrary, in 2007 no particularly strong interconnection can be detected between clusters, converging to the whole network in an symmetric way (see Figure 1d).

As a way of conclusion, there are two peculiarities of EU27 with respect to EMU which are worth highlighting.

The first one is the emergence of a very well defined Pharma-Hi Tech cluster in the EU27 for year 2007, grouping the MC:Special-Machinery and MC:Chemicals mini-clusters into an industrial agglomeration characterised by hi-tech and a great R&D effort. On the contrary, the traditional metal-machinery group remains separated forming a specific Heavy Machinery cluster. This tendency is not present in the EMU, where C09 Specialised Machin. forms an independent cluster (though cluster C08 Dressing-Chemicals *does* include the MC:Chemicals *as well as* industry K73 R&D).

The second one concerns the Services cluster in 2007. By looking at Table 4, it can be immediately seen that while in the EMU a great majority of its member industries (10 out of 14), if removed, would break the cluster, in the EU27 only industry K70 Real-estate would eventually break it. This means that on average the economy of the EU27 is tertiarised up to the point that connections between Service industries are so strong and pervasive as to make none of them, besides K70, essential for the survival of the cluster.

These differences cannot but depend on the 10 countries which belong to EU27 but not to the Euro Area (EMU). Looking at data about individual countries' gross outputs, it emerges that UK's share in total gross output of EU27 was about 17% in 2005, suggesting that the above-mentioned peculiarities of the EU27 with respect to the EMU reflect, to some extent, some specificities in the evolution of UK's economic structure.

4 A quantitative characterisation of cluster features

So far the analysis has dealt with the identification of clusters of industries. The aim of this section is to quantify the implications of looking at the economy by taking as a unit of analysis each of these clusters, and its associated net output subsystem (in the sense of Sraffa 1960, p. 89).

Consider the following Supply-Use accounting framework:[25]

$$\mathbf{z} = \mathbf{V}\mathbf{e} = \mathbf{U}\mathbf{e} + \mathbf{f} \tag{4.1}$$

$$\mathbf{g}^T = \mathbf{e}^T\mathbf{V} = \mathbf{e}^T\mathbf{U} + \mathbf{m}^T + \mathbf{t}_z^T + \mathbf{t}_g^T + \mathbf{w}^T + \boldsymbol{\pi}^T \tag{4.2}$$

$$\mathbf{D} = \widehat{\mathbf{z}}^{-1}\mathbf{V} \tag{4.3}$$

$$\mathbf{g} = \mathbf{V}^T\mathbf{e} = \mathbf{D}^T\mathbf{z} \tag{4.4}$$

$$L = \mathbf{l}^T\mathbf{e} \tag{4.5}$$

with (4.1) and (4.2) being the expenditure side and value added side of the system, respectively, (4.3) defines the market shares matrix (the percentage of commodity gross output produced by every industry for each product), (4.4) relates gross output by industry ($\mathbf{g}$) and by commodity ($\mathbf{z}$), and (4.5) defines the employment requirements by industry.

The spectral bisection (SB) algorithm for the identification of EMU and EU27 industry clusters has been performed on square Input-Output tables at basic prices for domestic output of industry $\times$ industry type. To obtain these tables we have applied the fixed product sales structure technology assumption to the expenditure side (4.1), as follows:[26]

$$\mathbf{D}^T\mathbf{z} = \mathbf{D}^T\mathbf{U}\mathbf{e} + \mathbf{D}^T\mathbf{f}$$

$$\text{define: } \mathbf{X} := \mathbf{D}^T\mathbf{U}, \mathbf{d} := \mathbf{D}^T\mathbf{f}, \text{ to obtain:}$$

$$\mathbf{g} = \mathbf{X}\mathbf{e} + \mathbf{d} \tag{4.6}$$

where $\mathbf{X}$ is the matrix of interindustry transactions, and $\mathbf{d}$ the vector of final uses by industry. Tables 20 and 21 in Appendix B display the resulting square Input-Output schemes, aggregated by clusters for EMU and EU27, years 2000 and 2007.

Consider four related aspects that characterise the role of each cluster in the system: (i) value added components per unit of cluster final demand, (ii) intra/inter-cluster intermediate uses, (iii) cluster participation in total final demand, and (iv) labour requirements by cluster industries and their associated subsystems.

As regards (i), express some current period value-added-side components (wages, profits and imports, respectively), as well as the interindustry trans-

[25] See Appendix A for the specification of the notation used.

[26] See ten Raa & Rueda-Cantuche (2009, p. 364) for a description of industry coefficient Input-Output technology assumptions.

actions matrix, in 'intensive' units (i.e. per unit of industry gross output):

$$\mathbf{a}_w^T := \mathbf{w}^T \widehat{\mathbf{g}}^{-1} \tag{4.7}$$
$$\mathbf{a}_\pi^T := \boldsymbol{\pi}^T \widehat{\mathbf{g}}^{-1} \tag{4.8}$$
$$\mathbf{a}_m^T := \mathbf{m}^T \widehat{\mathbf{g}}^{-1} \tag{4.9}$$
$$\boldsymbol{\Lambda} := \mathbf{X} \widehat{\mathbf{g}}^{-1} \tag{4.10}$$

By introducing (4.10) into (4.6), the expenditure system in intensive units is:

$$\mathbf{g} = \boldsymbol{\Lambda}\mathbf{g} + \mathbf{d} \tag{4.11}$$
$$\mathbf{g} = (\mathbf{I} - \boldsymbol{\Lambda})^{-1}\mathbf{d} \tag{4.12}$$

where (4.12) expresses gross output by industry in terms of total (direct and indirect) requirements to reproduce final demand. Note that $\mathbf{g}$ can be partitioned into as many parts $\mathbf{g}^{(i)}$ as there are elements in $\mathbf{d}$, by computing:

$$\mathbf{g}^{(i)} = (\mathbf{I} - \boldsymbol{\Lambda})^{-1}\mathbf{d}^{(i)} \tag{4.13}$$
$$\text{with } \mathbf{d}^{(i)} = \mathbf{e}_i d_i, \text{such that: } \mathbf{d} = \sum_{i=1}^{N} \mathbf{d}^{(i)}, \text{obtaining: } \mathbf{g} = \sum_{i=1}^{N} \mathbf{g}^{(i)}$$

For each value-added component it is possible to compute its total requirements per unit of industry-specific final demand d_i. For example, if the wage-bill is given by $W = \mathbf{a}_w^T\mathbf{g}$, total wages per unit of final demand of industry i are:

$$w^{(i)} = \mathbf{a}_w^T\mathbf{g}^{(i)} = \mathbf{a}_w^T(\mathbf{I} - \boldsymbol{\Lambda})^{-1}\mathbf{d}^{(i)} = \mathbf{a}_w^T(\mathbf{I} - \boldsymbol{\Lambda})^{-1}\mathbf{e}_i d_i = v_{w_i} d_i, \tag{4.14}$$
$$\text{where: } v_{w_i} = \mathbf{a}_w^T(\mathbf{I} - \boldsymbol{\Lambda})^{-1}\mathbf{e}_i \tag{4.15}$$

But we can take advantage of the fact that the SB algorithm produces a partition of the network, so that each cluster is associated to a mutually exclusive subset of industries. By noting with C the set of industries that belong to a given cluster, for each C we can compute the total wages per unit of cluster-final demand:

$$v_w^{(C)} = \frac{\sum_{i\in C} w^{(i)}}{\sum_{i\in C} d_i} = \frac{\sum_{i\in C} v_{w_i} d_i}{\sum_{i\in C} d_i} \tag{4.16}$$

which is a weighted (by the final demand components d_i of cluster industries) average of industries' v_{w_i}. In this way, $v_w^{(C)}$ reflects wage requirements to reproduce the product mix of final demand of cluster C, going through the whole network of intermediate consumption by industries. It is a general interdependence measure — as each v_{w_i} depends on the whole network as follows from (4.15) — but at the same time is cluster-specific.

By proceeding as in (4.14) and (4.16) for profits ($\Pi = \mathbf{a}_\pi^T \mathbf{g}$) and intermediate imports ($M = \mathbf{a}_m^T \mathbf{g}$), we get:

$$v_\pi^{(C)} = \frac{\sum_{i \in C} \pi^{(i)}}{\sum_{i \in C} d_i} = \frac{\sum_{i \in C} v_{\pi_i} d_i}{\sum_{i \in C} d_i} \tag{4.17}$$

$$v_m^{(C)} = \frac{\sum_{i \in C} m^{(i)}}{\sum_{i \in C} d_i} = \frac{\sum_{i \in C} v_{m_i} d_i}{\sum_{i \in C} d_i} \tag{4.18}$$

A comparison of $v_w^{(C)}$, $v_\pi^{(C)}$ and $v_m^{(C)}$ across clusters allows to distinguish the intensity in their *total* use of each of these value-added components.

As regards (ii) intra/inter-cluster intermediate uses, we inquire about the relevance of the main diagonal of the intermediate consumption part of the Input-Output Tables aggregated in clusters (Tables 20 and 21). To do so, we compute — for each cluster C — the proportion of intra-cluster to total-cluster transactions, sales and purchases, respectively:

$$IU^{(C)} = \frac{\sum_{i \in C} \sum_{j \in C} x_{ij} + \sum_{i \in C} \sum_{j \in C} x_{ij}}{\sum_{i=1}^{N} \sum_{j \in C} x_{ij} + \sum_{i \in C} \sum_{j=1}^{N} x_{ij}} \tag{4.19}$$

$$IS^{(C)} = \frac{\sum_{i \in C} \sum_{j \in C} x_{ij}}{\sum_{i \in C} \sum_{j=1}^{N} x_{ij}} \tag{4.20}$$

$$IP^{(C)} = \frac{\sum_{i \in C} \sum_{j \in C} x_{ij}}{\sum_{i=1}^{N} \sum_{j \in C} x_{ij}} \tag{4.21}$$

As to (iii) cluster participation in total final demand, we look at the relative importance of each cluster-specific component of final uses within the economy-wide value for the respective component, noting that final demand domestically produced can be decomposed into: $\mathbf{d} = \mathbf{d}_c + \mathbf{d}_g + \mathbf{d}_k + \mathbf{d}_{vk} + \mathbf{d}_x$, i.e. final private consumption ($\mathbf{d}_c$), government consumption ($\mathbf{d}_g$), gross fixed capital formation ($\mathbf{d}_k$), changes in inventories and valuables ($\mathbf{d}_{vk}$) and exports ($\mathbf{d}_x$).

Finally, as regards (iv) labour requirements by cluster industries and their associated subsystems, we propose to intersect the cluster and subsystem approaches, to obtain a novel decomposition of the division of labour in the economic system.

Consider subsystem analysis first.[27] Labour requirements of industry i to produce gross output g_i (L_i), and those of its associated subsystem to

[27] See, for example, Pasinetti (1973, pp. 5-7).

produce final demand d_i ($L^{(i)}$), are given by:

$$L_i = \mathbf{l}^T \mathbf{e}_i = \mathbf{a}_l^T \widehat{\mathbf{g}}_i \mathbf{e}_i = a_{l_i} g_i, \tag{4.22}$$

$$\text{where: } \mathbf{a}_l^T := \mathbf{l}^T \widehat{\mathbf{g}}^{-1} \tag{4.23}$$

$$L^{(i)} = \mathbf{a}_l^T (\mathbf{I} - \mathbf{\Lambda})^{-1} \mathbf{d}^{(i)} = v_{l_i} d_i \tag{4.24}$$

Note that adding-up over industries or subsystems adds-up to total employment L, i.e. $\sum_{i=1}^N L_i = \sum_{i=1}^N L^{(i)} = L$.

Exploting the additivity of final demand subsystems and the mutually exclusive clustering of industries, we may also compute labour requirements for cluster industries to reproduce cluster-gross output, as well as the labour requirements to reproduce cluster-final demand, for a given cluster C:

$$L_C = \sum_{i \in C} L_i \tag{4.25}$$

$$L^{(C)} = \sum_{i \in C} v_{l_i} d_i \tag{4.26}$$

Also in this case we have that $\sum_C L_C = \sum_C L^{(C)} = L$.

For each cluster C, the labour redistribution between the aggregate of cluster-industries and its associated subsystem is $L^{(C)} - L_C$. However, in aggregating over all industries and subsystems of a cluster, redistribution across individual units might compensate for each other, showing a modest aggregate figure. To capture the degree of labour redistribution that takes place in the units composing a cluster, we compute the mean absolute deviation of individual labour redistributions between industries and subsystems of cluster C:

$$MAD_l^{(C)} = \sum_{i \in C} \left| (L^{(i)} - L_i) - \frac{\sum_{i \in C}(L^{(i)} - L_i)}{N_C} \right| \tag{4.27}$$

A high value for $MAD_l^{(C)}$ means that cluster-industries and/or their associated subsystems are relevant in the redistribution of employment through the interindustry network of intermediate consumption.

Note, however, that $L_i = a_{l_i} g_i$ and $L^{(i)} = v_{l_i} d_i$ refer to a different concept of output, and while L_i is a partial measure, in $L^{(i)}$ we have general interdependence (through v_{l_i}). Thus, while $L^{(i)} - L_i$ would give an idea of the labour redistribution that occurs when switching from industry to subsystem as a unit of analysis, its detailed interpretation is not straightforward.

In fact, take industry i and subsystem i: where does the labour 'exported' by the industry to other subsystems go? Where does the labour 'imported' by the subsystem from other industries come from? What are the additional complications of accounting for labour redistribution among intra-cluster and extra-cluster industries/subsystems?

In order to answer these questions, it is necessary to further decompose $L^{(i)} - L_i$. Accordingly, note that:

$$L^{(i)} - L_i = \underbrace{\mathbf{e}^T \widehat{\mathbf{a}}_{l(-i)} (\mathbf{I} - \mathbf{\Lambda})^{-1}_{(-i,i)} d_i}_{(I)} - \underbrace{a_{l_i} (\mathbf{I} - \mathbf{\Lambda})^{-1}_{(i,-i)} \widehat{\mathbf{d}}_{(-i)} \mathbf{e}}_{(II)} \tag{4.28}$$

where $\widehat{\mathbf{a}}_{l(-i)}$ is a diagonal matrix with all elements but the i-th. one, $(\mathbf{I} - \mathbf{\Lambda})^{-1}_{(-i,i)}$ is a column vector obtained by extracting the i-th. column and removing the i-th. row of the Leontief inverse, $(\mathbf{I} - \mathbf{\Lambda})^{-1}_{(i,-i)}$ is a row vector obtained by extracting the i-th. row and removing the i-th. column of the Leontief inverse, and $\widehat{\mathbf{d}}_{(-i)}$ is a diagonal matrix with all elements but the i-th. one.

Hence, at the level of the individual industry/subsystem i:

(I) represents the labour coming from other industries to subsystem i, and

(II) represents the labour going from industry i to other subsystems.

By further decomposing (I) into intra- and extra-cluster industries, and (II) into intra- and extra-cluster subsystems, and summing over all industries/subsystems for a given cluster C, it is possible to compute:

$$(I.A) = \sum_{i \in C} \sum_{j \in C} \mathbf{e}_j^T \widehat{\mathbf{a}}_{l(-i)} (\mathbf{I} - \mathbf{\Lambda})^{-1}_{(-i,i)} d_i \tag{4.29}$$

$$(I.B) = \sum_{i \in C} \sum_{j \notin C} \mathbf{e}_j^T \widehat{\mathbf{a}}_{l(-i)} (\mathbf{I} - \mathbf{\Lambda})^{-1}_{(-i,i)} d_i \tag{4.30}$$

$$(II.A) = \sum_{i \in C} \sum_{j \in C} a_{l_i} (\mathbf{I} - \mathbf{\Lambda})^{-1}_{(i,-i)} \widehat{\mathbf{d}}_{(-i)} \mathbf{e}_j \tag{4.31}$$

$$(II.B) = \sum_{i \in C} \sum_{j \notin C} a_{l_i} (\mathbf{I} - \mathbf{\Lambda})^{-1}_{(i,-i)} \widehat{\mathbf{d}}_{(-i)} \mathbf{e}_j \tag{4.32}$$

where, at the level of the individual cluster C:

$(I.A)$ Labour coming from intra-cluster industries to intra-cluster subsystems

$(I.B)$ Labour coming from extra-cluster industries to intra-cluster subsystems

$(II.A)$ Labour going from intra-cluster industries to intra-cluster subsystems

$(II.B)$ Labour going from intra-cluster industries to extra-cluster subsystems

Note that $(I.A) \equiv (II.A)$.

Moreover, with these magnitudes we compute:

$$LR^{(C,C)} := (I.A)/L \tag{4.33}$$

$$LR^{(-C,C)} := (I.B)/L \tag{4.34}$$

$$LR^{(C,-C)} := (II.B)/L \tag{4.35}$$

$$PLR^{(C)}_{(in)} := \frac{(I.A)}{(I.A) + (I.B)} \tag{4.36}$$

$$PLR^{(C)}_{(out)} := \frac{(II.A)}{(II.A) + (II.B)} \tag{4.37}$$

Measures (4.33)-(4.34) quantify the importance (in terms of total employment) of labour being redistributed within cluster C, labour coming from extra-cluster industries to cluster C subsystems, and labour going from cluster C industries to extra-cluster subsystems, respectively. In addition, measures (4.36)-(4.37) quantify the proportion of labour from intra-cluster industries remaining in intra-cluster subsystems and the proportion of labour going to intra-cluster subsystems coming from intra-cluster industries, respectively. In this way, the inner-persistence of labour redistribution within each cluster can be assessed.

All the indicators related to aspects (i)-(iv) detailed above have been computed for EMU/EU27-2000/2007 Input-Output Tables. Results are reported in Tables 5 and 6.

Table 5: Quantitative features of Clusters of the European Monetary Union (EMU/EA17), years 2000, 2007

Clusters (Year 2000)			Value added			Intermediate uses			Final demand				Labour of Cluster industries and subsystems							
Code	Description	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
C01	Agri-Food	6	49.2	39.9	8.7	57.8	75.9	46.6	23.6	28.2	0.7	6.8	19.7	22.7	7.9	2.8	1.6	4.6	63.6	37.9
C02	Construction	8	48.2	36.2	11.8	45.2	54.4	38.6	3.2	0.3	51.0	5.1	10.4	12.3	4.8	0.8	2.1	4.0	27.9	17.1
C03	Energy	4	35.7	46.9	13.8	30.6	26.7	35.8	3.9	0.2	0.6	2.5	2.4	2.0	0.3	0.0	1.0	0.7	4.5	6.8
C04	Transport-Dressing	10	42.3	36.0	16.9	42.8	39.4	46.9	13.4	1.7	4.5	18.4	10.8	9.1	4.3	1.0	4.5	2.9	18.4	26.1
C05	Heavy-HiTech Mach.	9	47.8	28.3	20.7	50.6	51.4	49.8	2.3	0.6	20.1	29.5	7.0	7.7	3.1	0.8	2.4	3.1	25.3	21.1
C06	Services	16	49.2	41.6	4.8	60.3	53.9	68.5	44.9	66.7	14.8	11.4	44.1	38.3	13.6	5.0	9.6	3.8	34.3	57.1
C07	Chemicals	2	40.2	32.7	23.3	35.6	34.8	36.6	1.8	2.1	0.4	13.1	1.8	2.2	1.0	0.0	0.9	1.3	4.8	3.5
C08	Office-machinery-PC	1	26.1	22.3	49.1	18.1	21.4	15.6	0.1	0.0	1.1	1.9	0.1	0.3	0.0	0.0	0.0	0.2	0.0	0.0
C09	Motor-vehicles	2	48.0	30.2	18.5	44.9	63.4	34.7	6.5	0.1	6.8	11.2	3.4	5.4	1.9	0.1	0.7	2.6	14.4	4.1
C10	Renting-equipment	1	19.9	72.9	3.8	14.5	9.8	28.2	0.3	0.0	0.0	0.1	0.2	0.1	0.0	0.0	0.2	0.0	0.0	0.0
	Economy-wide	59	47.3	37.9	11.0	50.2			100.0	100.0	100.0	100.0	100.0	100.0	0.0	10.7	23.2	23.2	31.6	31.6

Clusters (Year 2007)			Value added			Intermediate uses			Final demand				Labour of Cluster industries and subsystems							
Code	Description	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
C01	Agri-Food	5	40.2	44.8	12.4	57.5	72.0	47.9	18.2	0.2	0.5	6.4	10.7	12.5	5.8	2.1	1.5	3.3	57.4	38.4
C02	Furniture	3	43.1	36.2	19.0	38.4	40.5	36.5	1.3	0.1	1.8	2.3	1.6	1.5	0.6	0.1	0.6	0.6	14.3	14.9
C03	Energy	4	31.6	43.5	20.6	37.0	32.3	43.2	3.2	0.1	0.5	2.2	1.0	1.0	0.1	0.0	0.6	0.6	3.6	3.6
C04	Transport-Trade	10	40.4	38.5	17.5	43.4	39.6	48.1	14.7	2.0	6.4	18.1	11.5	9.7	3.7	1.1	4.9	3.1	18.2	26.0
C05	Services	15	49.2	39.7	5.0	53.3	49.6	57.5	46.8	93.5	16.1	9.8	52.7	47.7	14.8	4.7	9.6	4.6	33.0	50.8
C06	Heavy Machinery	7	43.3	31.7	23.9	59.5	64.0	55.6	4.0	0.2	18.2	32.6	6.4	8.4	4.6	1.0	1.8	3.8	34.7	20.1
C07	Construction	3	45.7	42.7	11.3	49.6	60.7	41.9	1.2	0.1	51.6	1.9	8.6	10.7	3.7	0.4	1.6	3.6	20.6	10.2
C08	Dressing-Chemicals	6	40.1	33.0	24.5	39.7	40.9	38.5	3.6	3.3	0.4	16.0	3.4	4.1	1.6	0.2	1.2	2.0	14.7	9.4
C09	Specialised Machin.	3	42.8	28.8	26.7	33.8	40.3	29.0	0.7	0.5	4.3	7.8	1.3	2.0	0.2	0.0	0.3	1.1	8.5	2.7
C10	Finance	3	45.9	36.0	7.1	49.4	43.1	57.9	6.3	0.0	0.1	3.1	2.8	2.3	1.5	0.4	1.4	0.9	20.4	28.8
	Economy-wide	59	45.1	38.7	12.1	49.9			100.0	100.0	100.0	100.0	100.0	100.0	0.0	10.0	23.5	23.5	29.7	29.7

(1) Number of industries per cluster
(2) Subsystem labour cost per monetary unit of net output (in %) (4.16)
(3) Subsystem surplus per monetary unit of net output (in %) (4.17)
(4) Subsystem imports per monetary unit of net output (in %) (4.18)
(5) Proportion of intra-cluster to total intermediate transactions (in %) (4.19)
(6) Proportion of intra-cluster to total intermediate sales (in %) (4.20)
(7) Proportion of intra-cluster to total intermediate purchases (in %) (4.21)
(8) Percentage of cluster to total final private consumption (in %)
(9) Percentage of cluster to total final government consumption (in %)
(10) Percentage of cluster to total gross fixed capital formation (in %)
(11) Percentage of cluster to total exports (in %)
(12) Percentage of cluster to total employment by industry (in %) (4.25)
(13) Percentage of cluster to total employment by subsystem (in %) (4.26)
(14) Mean absolute deviation of labour redistribution between subsystem and industry (in p.p.) (4.27)
(15) Labour going from intra-cluster industries to intra-cluster subsystems/total labour (in p.p.) (4.33)
(16) Labour going from intra-cluster industries to extra-cluster subsystems/total labour (in p.p.) (4.34)
(17) Labour coming from extra-cluster industries to intra-cluster subsystems/total labour (in p.p.) (4.35)
(18) Proportion of labour from intra-cluster industries remaining in intra-cluster subsystems (in %) (4.37)
(19) Proportion of labour to intra-cluster subsystems coming from intra-cluster industries (in %) (4.36)

Source: Own computation based on EUROSTAT SUIOT and National Accounts Databases

Table 6: Quantitative features of Clusters of the European Union (EU27), years 2000, 2007

Clusters (Year 2000)			Value added			Intermediate uses			Final demand				Labour of Cluster industries and subsystems							
Code	Description	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
C01	Agri-Food	5	53.6	36.7	8.2	58.9	72.7	49.6	18.8	0.2	0.6	5.7	13.5	15.0	9.4	3.6	2.0	3.5	64.2	50.7
C02	Construction	6	59.7	28.2	8.9	46.7	55.6	40.3	3.0	0.3	48.8	4.9	9.9	11.7	4.7	0.9	2.2	4.0	28.3	18.0
C03	Energy	5	34.7	35.1	26.1	43.9	36.9	54.2	3.8	0.1	0.4	3.0	1.2	1.1	0.4	0.1	0.7	0.6	13.7	15.9
C04	Heavy Machinery	8	59.8	24.0	13.3	54.4	58.6	50.8	8.2	0.2	22.0	29.8	8.6	10.1	4.8	1.2	2.6	4.1	31.7	22.8
C05	Dressing-Chemicals	5	54.0	27.2	15.6	41.8	42.6	41.0	4.7	1.9	0.6	15.4	3.8	4.4	1.4	0.3	1.3	1.9	16.5	11.9
C06	Transport-Media	7	57.0	30.2	8.2	39.2	35.6	43.7	11.4	1.6	4.7	15.4	10.2	7.5	3.1	0.8	5.2	2.6	13.8	24.6
C07	Specialised Machin.	4	52.4	21.5	23.7	34.2	39.4	30.3	1.2	0.4	7.6	11.7	1.5	2.4	0.3	0.1	0.4	1.4	10.5	3.5
C08	Services	16	57.6	34.5	3.7	58.5	52.9	65.4	42.9	65.5	15.0	13.6	41.2	36.6	12.8	4.8	9.2	4.5	34.5	51.6
C09	Renting-equipment	1	34.5	58.6	3.5	12.1	8.5	20.9	0.4	0.0	0.0	0.1	0.3	0.1	0.0	0.0	0.2	0.1	0.0	0.0
C10	Health	1	72.0	20.8	4.2	29.6	68.0	18.9	4.1	29.7	0.0	0.0	8.6	10.1	0.0	0.0	0.2	1.7	0.0	0.0
C11	Personal-services	1	43.1	49.8	4.0	12.3	12.4	12.2	1.3	0.1	0.2	0.3	1.3	1.1	0.0	0.0	0.4	0.2	0.0	0.0
	Economy-wide	59	57.7	30.7	8.1	49.8			100.0	100.0	100.0	100.0	100.0	100.0	0.0	11.8	24.6	24.6	32.5	32.5

Clusters (Year 2007)			Value added			Intermediate uses			Final demand				Labour of Cluster industries and subsystems							
Code	Description	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
C01	Agri-Food	5	43.7	44.9	8.7	56.6	71.2	47.0	17.9	0.2	0.5	5.1	12.3	13.9	7.5	2.8	1.8	3.5	60.4	44.5
C02	Construction	6	47.1	43.5	8.1	52.3	61.4	45.6	1.5	0.2	50.8	2.5	9.3	11.1	5.1	0.7	2.0	3.8	24.8	14.7
C03	Energy	3	30.0	47.8	16.7	36.7	32.2	42.7	3.2	0.1	0.4	0.5	0.9	0.9	0.2	0.0	0.5	0.5	7.1	7.5
C04	Transport-Trade	9	43.4	39.8	12.4	43.7	39.2	49.3	15.0	1.9	5.7	18.6	11.6	9.3	3.9	1.1	5.3	3.0	17.2	26.9
C05	Heavy Machinery	9	48.0	33.7	16.2	59.4	64.8	54.8	6.1	0.3	22.1	34.0	7.5	9.8	4.8	1.2	2.1	4.5	35.9	20.8
C06	Dressing	3	46.5	34.4	17.0	44.7	55.9	37.2	1.9	0.0	0.1	2.3	1.5	1.7	0.3	0.1	0.3	0.5	24.4	15.7
C07	Paper	1	43.6	37.3	12.1	27.8	25.5	30.7	0.3	0.0	0.0	1.4	0.3	0.3	0.0	0.0	0.2	0.2	0.0	0.0
C08	Services	16	48.6	41.1	3.9	62.4	56.5	69.7	47.0	61.9	16.5	16.8	44.5	38.5	13.4	5.8	10.1	4.2	36.3	57.9
C09	Pharma-Hi Tech	7	55.4	30.9	10.0	41.8	47.1	37.5	7.1	35.3	3.8	18.8	12.2	14.5	3.5	0.3	1.5	3.8	18.0	8.0
	Economy-wide	59	47.9	39.3	8.4	53.7			100.0	100.0	100.0	100.0	100.0	100.0	0.0	11.9	23.8	23.8	33.4	33.4

(1) Number of industries per cluster
(2) Subsystem labour cost per monetary unit of net output (in %) (4.16)
(3) Subsystem surplus per monetary unit of net output (in %) (4.17)
(4) Subsystem imports per monetary unit of net output (in %) (4.18)
(5) Proportion of intra-cluster to total intermediate transactions (in %) (4.19)
(6) Proportion of intra-cluster to total intermediate sales (in %) (4.20)
(7) Proportion of intra-cluster to total intermediate purchases (in %) (4.21)
(8) Percentage of cluster to total final private consumption (in %)
(9) Percentage of cluster to total final government consumption (in %)
(10) Percentage of cluster to total gross fixed capital formation (in %)
(11) Percentage of cluster to total exports (in %)
(12) Percentage of cluster to total employment by industry (in %) (4.25)
(13) Percentage of cluster to total employment by subsystem (in %) (4.26)
(14) Mean absolute deviation of labour redistribution between subsystem and industry (in p.p.) (4.27)
(15) Labour going from intra-cluster industries to intra-cluster subsystems/total labour (in p.p.) (4.33)
(16) Labour going from intra-cluster industries to extra-cluster subsystems/total labour (in p.p.) (4.34)
(17) Labour coming from extra-cluster industries to intra-cluster subsystems/total labour (in p.p.) (4.35)
(18) Proportion of labour from intra-cluster industries remaining in intra-cluster subsystems (in %) (4.37)
(19) Proportion of labour to intra-cluster subsystems coming from intra-cluster industries (in %) (4.36)

Source: Own computation based on EUROSTAT SUIOT and National Accounts Databases

Consider EMU for the year 2000 (Table 5). Half of the cluster-subsystems are close to the economy-wide value of wages per unit of final demand, which is around 47.3%. Moreover, note the total import intensity of Office-machinery-PC, Chemicals and Heavy-HiTech Mach. (49 %, 23.3% and 20.7%, respectively). Agri-Food, Heavy-HiTech Mach. and Services are above the average (which is around 50%) value of intra-cluster intermediate consumption, being particularly important in intra-cluster sales of Agri-Food (76%) and intra-cluster purchases of Services (68.5%).

As to final demand, note that Agri-Food, Transport-Dressing and Services add-up to 82% of private consumption, and note the essential role of Heavy-HiTech Mach. and Chemicals on exports (this last cluster being composed of only two industries but accounting for 13% of foreign demand). Finally, the high value in column (14) for Agri-Food and Services reflects the fact that these two clusters, to an important extent, reproduce themselves by means of themselves (as more than 63% of labour coming from intra-Agri-Food industries remains in the associated subsystems and around 57% of labour going to intra-Services subsystems comes from intra-cluster industries).

As to EMU for the year 2007 (Table 5), the changing cluster composition results in differences in individual clusters (with respect to the year 2000) in the use of value added components per unit of final demand that compensate each other (as economy-wide values are very similar in the two years). Note the interesting intra-cluster intermediate consumption of Finance, which is around 49%, the fact that more than 93% of government consumption is directed to Services and that Heavy Machinery has fewer industries than in 2000 but has gained relative importance in exports (more than 32% of total foreign demand). By inspecting columns (18) and (19) it is clear that Energy and Specialised Machin. are mainly suppliers in the interindustry network, as a very small proportion of their labour remains in intra-cluster subsystems or comes from intra-cluster industries.

Switching to the analysis of EU27 for the year 2000 (Table 6), it is noticeable the sharp difference in labour cost per unit of final demand with respect to the EMU value for the same year (57.7% against 47.3%). Thus, incorporating extra-EMU members in the analysis suggests their having a different functional income distribution in 2000 which, by looking at the 2007 value (47.9%), has been taken to EMU standards. Additionally, note that Agri-Food, Heavy Machinery and Services have above average values for intra-cluster intermediate consumption and Health represents almost 30% of government final demand.

Finally, as regards EU27 for the year 2007 (6), note that Heavy Machinery and Pharma-Hi Tech together add-up to 53% of total exports and 10 % of total employment goes from intra-Services industries to extra-Services cluster subsystems. In switching from EMU to EU27, note the decline in imports per unit of final demand (from 12% to 8%), which suggests the importance of EMU/extra-EMU (but intra-EU27) external trade.

All in all, making a rough approximation on all four period-country-group combinations, some highlight points are:

1. Labour costs (wages plus taxes on labour), Profits and Imports per unit of final demand have stabilised around 45-47%, 38-39% and 8-12%, respectively (leaving a minor role of around 5% of net output to taxes on products and production).

2. Intra-cluster productive consumption accounts for 49-53% of the value of intermediate transactions.

3. Agri-Food, Services and Transport-Trade clusters account for 73-82% of final private consumption.

4. Services clusters (including Health) account for more than 93% of government consumption.

5. Heavy Machinery, Specialised Machin. and Chemicals clusters account for 50-60% of total EU27 exports.

6. Around 2 $\times$ 10-12% = 20-24% of total employment is redistributed between intra-cluster industries and intra-cluster subsystems.

7. Around 23-24% of total employment is redistributed between intra-cluster and extra-cluster industries/subsystems.

8. Around 30-33% of the labour from (to) intra-cluster industries (subsystems) remains in (comes from) intra-cluster subsystems (industries).

The last three points imply that there is still an important part of labour redistribution that occurs between industries and subsystems of different clusters, showing the relevance of combining the cluster and subsystem approaches.

4.1 Underlying productivity performance of individual subsystems in the EU

The analysis of total labour productivity growth by looking at the economy as a set of self-replacing subsystems with a high degree of autonomy allows to pinpoint the most dynamic sectors in each country-group: EMU and EU27.

Note from (4.24) that subsystem labour is given by: $L^{(i)} = v_{l_i} d_i$ and that v_{l_i} involves the whole network of interindustry intermediate consumption. Hence, the reciprocal of labour intensity per unit of final demand for each subsystem provides a measure of total (direct and indirect) labour productivity ($\alpha_l^{(i)}$) and its associated rate of growth (ρ_i):

$$\alpha_l^{(i)} = \frac{1}{v_{l_i}} = \frac{d_i}{L^{(i)}} \tag{4.38}$$

$$\rho_i = d \ln \alpha_l^{(i)} \tag{4.39}$$

By computing a weighted (by the levels of subsystem labour) average of subsystem-specific growth rates we obtain an economy-wide rate of total labour productivity growth:

$$\rho^* = \frac{\sum_{i=1} \rho_i L^{(i)}}{\sum_{i=1} L^{(i)}} \tag{4.40}$$

Given that ρ_i reflects labour saving trends it is always important to consider the employment dimension associated to changes in productivity, as high values for ρ_i might be due to employment reduction, which would hinder effective demand. Therefore, ρ_i should always be accompanied by its associated subsystem labour growth rate: $d \ln L^{(i)}$.

Discrete approximations for ρ_i and $d \ln L^{(i)} (\approx \Delta\% L^{(i)})$ have been computed for both EMU and EU27 between 2000 and 2007, on an average yearly growth rate basis. Figures 4 and 5 report the results, respectively.[28]

Consider the Euro Area (EMU) first (Figure 4). Note that while almost all manufacturing subsystems are increasing productivity (brown circles above the dashed horizontal line), most of them are doing so by decreasing subsystem labour (being to the left of the dashed vertical line). On the contrary, many of the crucial (as regards their employment importance) subsystems are decreasing productivity while increasing employment (e.g. F45 Construction, H55 Hotel-restaurant, M80 Education, K74 Business-services).

[28] In order to improve visualisation some subsystems have been left out of the Figures (being outliers in the order of magnitude of their growth rates). However, the sectors displayed account for more than 97.5% of total employment.

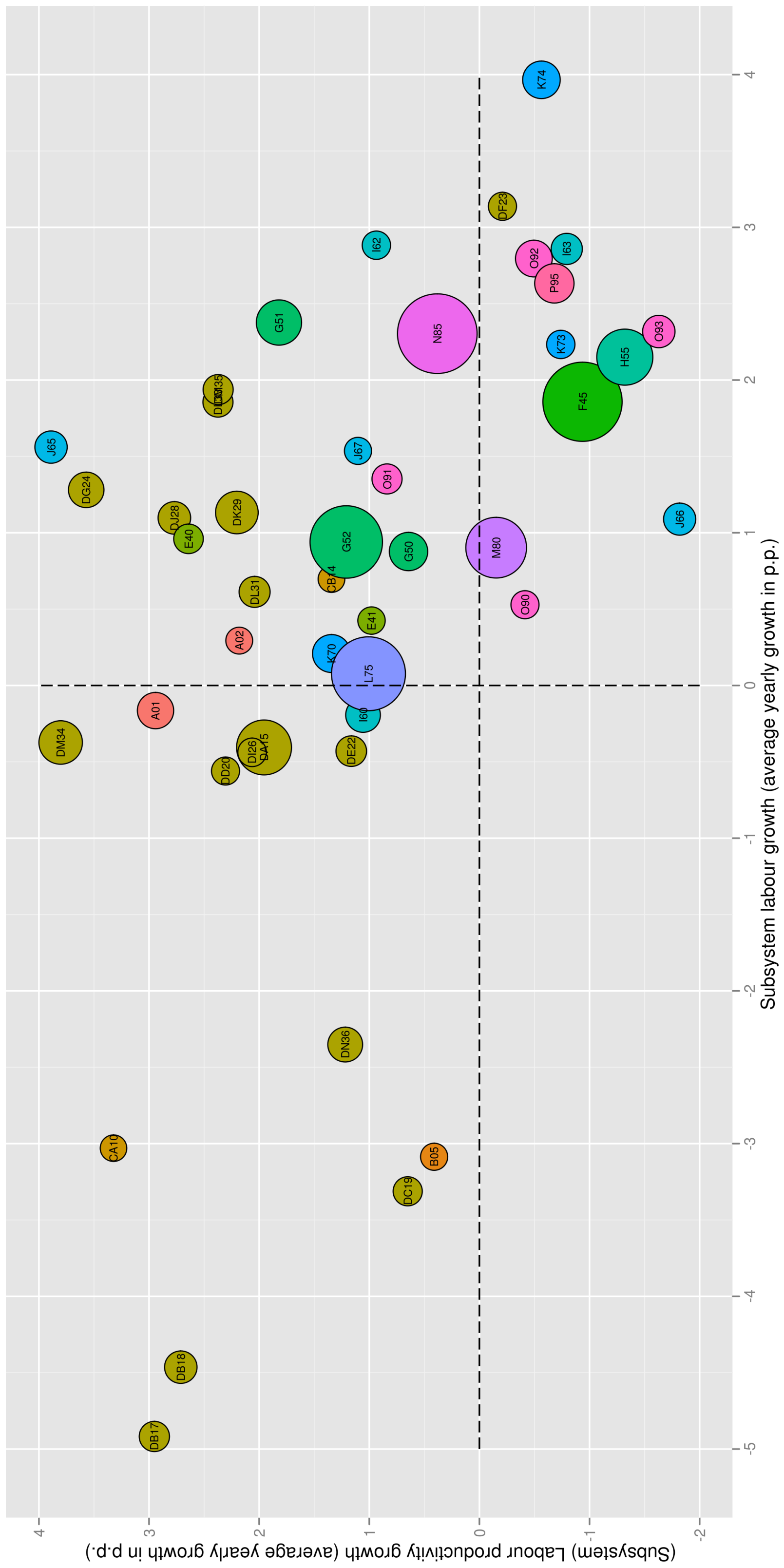


Figure 4: Subsystem Employment and Labour Productivity Dynamics for the Euro Area (2000-2007)
(Size of each point represents subsystem labour to total labour of the initial year)
Source: Own computation based on EUROSTAT SUIOT and National Accounts Databases

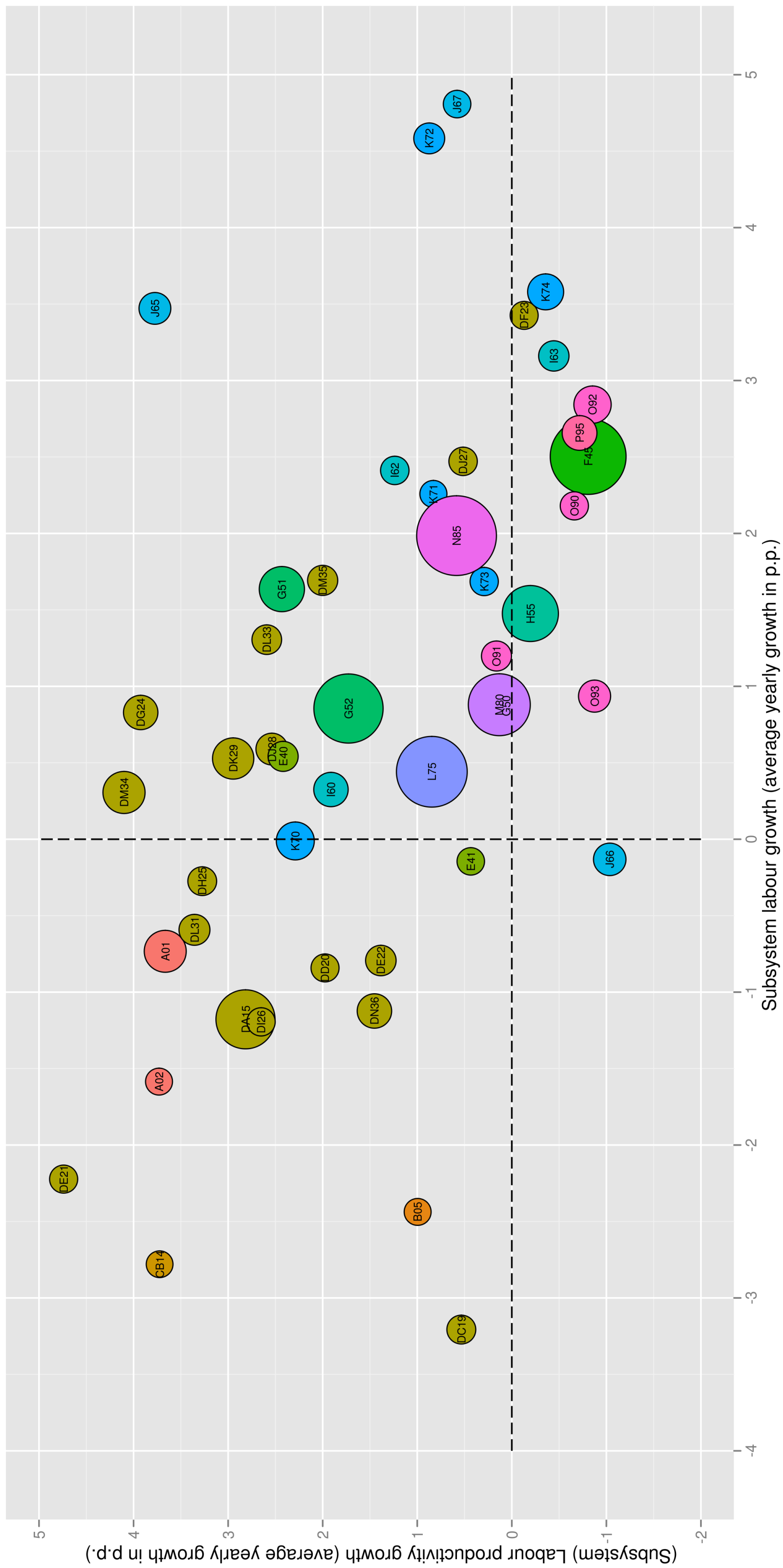


Figure 5: Subsystem Employment and Labour Productivity Dynamics for the European Union (2000-2007)
(Size of each point represents subsystem labour to total labour of the initial year)
Source: Own computation based on EUROSTAT SUIOT and National Accounts Databases

To have a quantitative account of the most dynamic subsystems, Table 7 reports those sectors with higher than average productivity increase (i.e. with $\rho_i \geq \rho^* = 1.0143$ p.p. on a yearly average growth basis) and positive subsystem labour growth.

Table 7: Dynamic Subsystems for EMU (EA17) between 2000 and 2007

Subsystems with $\rho_i \geq \rho^* = 1.0143$ and $\Delta\%L^{(i)} > 0$				
NACE	Subsystem	ρ_i (in p.p.)	$\Delta\%L^{(i)}$ (in p.p.)	$L^{(i)}/L$ (in %)
A02	Forestry	2.181	0.294	0.073
CB14	Stone-sand-clay-minerals	1.343	0.697	0.047
DG24	Chemicals-pharma	3.571	1.281	1.678
DH25	Rubber-plastics	4.202	0.832	0.510
DJ28	Structural-metal-products	2.770	1.097	1.137
DK29	Mechanical-machinery	2.203	1.133	2.993
DL31	Electrical-machinery	2.041	0.613	0.844
DL33	Medical-precision-equip.	2.374	1.856	0.670
DM35	Ships-railway-aircrafts	2.373	1.937	0.675
E40	Electricity-gas	2.642	0.960	0.580
G51	Wholesale-trade	1.820	2.377	3.485
G52	Retail-trade	1.209	0.940	8.517
I61	Transport-water	6.094	5.710	0.229
I64	Post-telecomm.	6.042	1.311	0.934
J65	Finance	3.890	1.561	1.089
J67	Brokerage-credit-cards	1.103	1.536	0.123
K70	Real-estate	1.343	0.209	2.137
K72	Computer-services	1.091	4.132	0.874

Source: Own computation based on EUROSTAT SUIOT and National Accounts Databases

It emerges that a manufacturing core that accounts for 8.5 % of total employment (subsystems DG24 Chemicals-pharma to DM35 Ships-railway-aircrafts in Table 7) has been particularly dynamic, as well as ICT-services subsystems (I64 Post-telecomm. and K72 Computer-services), together with Trade (G51 Wholesale-trade and G52 Retail-trade) and Financial services (J65 Finance and J67 Brokerage-credit-cards) subsystems.

Switching to EU27, Figure 5 shows that the situation worsens for many dynamic manufacturing subsystems in the EMU region, and silghtly improves for some sluggish service subsystems of the EMU. Note that J65 Finance has been the best performing subsystem during 2000-2007, which is interesting considering recent developments in Europe. However, as a subsystem (i.e. considering all the labour provided by its supporting industries through intermediate inputs) it accounts for only 1% of total employment. Also in this case, Table 8 reports those sectors with higher than average productivity growth (i.e. with $\rho_i \geq \rho^* = 1.346$) and with positive subsystem employment growth.

Table 8: Dynamic Subsystems for EU27 between 2000 and 2007

Subsystems with $\rho_i \geq \rho^* = 1.346$ and $\Delta\% L^{(i)} > 0$				
NACE	Subsystem	ρ_i (in p.p.)	$\Delta\% L^{(i)}$ (in p.p.)	$L^{(i)}/L$ (in %)
DG24	Chemicals-pharma	3.925	0.829	1.515
DJ28	Structural-metal-products	2.537	0.590	1.022
DK29	Mechanical-machinery	2.947	0.528	2.814
DL33	Medical-precision-equip.	2.591	1.305	0.609
DM34	Motor-vehicles	4.101	0.306	2.970
DM35	Ships-railway-aircrafts	2.000	1.693	0.707
E40	Electricity-gas	2.419	0.542	0.698
G51	Wholesale-trade	2.432	1.636	3.548
G52	Retail-trade	1.727	0.855	8.086
I60	Transport-land	1.913	0.325	1.506
I61	Transport-water	6.975	3.996	0.259
I64	Post-telecomm.	5.427	1.498	0.939
J65	Finance	3.774	3.472	1.022

Source: Own computation based on EUROSTAT SUIOT and National Accounts Databases

With respect to EMU area, the manufacturing core of dynamic subsystems in the EU27 (subsystems DG24 Chemicals-pharma to DM35 Ships-railway-aircrafts in Table 8) accounts for a greater share in total employment (more than 9.63% against 8.5%). It is interesting that the EU27 group includes DM34 Motor-vehicles (which was not in the EMU group of Table 7), highlighting the importance of extra EMU-countries in the European automobile value chain. As to other sectors, Transport subsystems (I60 Transport-land and I61 Transport-water) have improved their performance (note that this coincides with the geographical enlargement of the area considered), while Trade subsystems (which account for more than 11.5% of total employment) and Finance have also seen a positive productivity growth path accompanied by employment expansion.

This quantitative characterisation has had the scope to provide a birds' eye view on the structure of EMU and EU27 underlying the clustering results. However, it is also of importance to see whether this Euro Area or EU27 clustering of industries has come about from convergent national economies, or has been a mere statistical aggregation of completely dissimilar interindustry structures. The next section explores this issue.

5 Clustering structure of individual countries behind the aggregates

After having analysed, in Section 3, the clustering structure of the EU27 and the EMU in years 2000 and 2007, it is interesting to inspect it further by uncovering the underlying structures of the single countries. In order to do so, we define a pair-wise measure of similarity between countries' clustering structures and we use it to build a similarity matrix; based on it, we iteratively apply SB to look for above-average forward — between different clusters — and backward — between the industries of each cluster — connections and draw the resulting dendrogram.

Before showing the results, it is worth illustrating, by means of a simple example, the similarity measure we developed.

Consider a network of four countries ($\varphi = DE, FR, IT, NL$) with six industries ($i = A, B, C, D, E, F$; $N = 6$); in each country φ, we identify c^{φ} (indexed $\alpha^{\varphi} = 1, \ldots, c^{\varphi}$) clusters including m^{φ}_{α} industries each (see Figures 6a and 6b). The number of all the possible pair-wise combinations of industries within each cluster — i.e. the number of *significant* edges — is given by $\omega^{\varphi}_{\alpha} = m^{\varphi}_{\alpha}(m^{\varphi}_{\alpha} - 1)/2$.[29] Given two countries in the network, the measure of similarity we developed is based on the number of significant edges in the first one which are significant in second one too, and *vice-versa*.

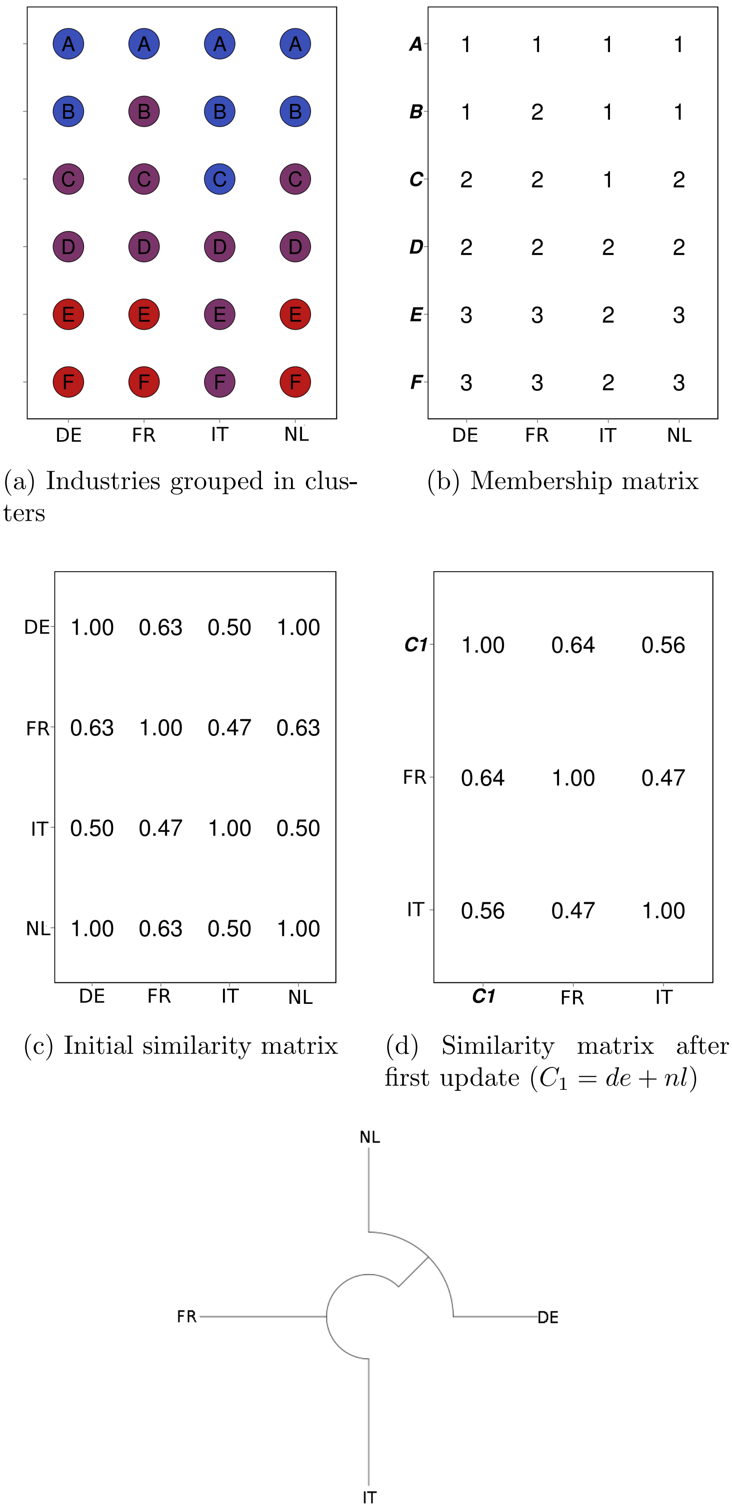
Take, as a way of example, Germany and Italy. In Germany, industries are grouped in $c^{DE} = 3$ clusters, including $m^{DE}_1 = m^{DE}_2 = m^{DE}_3 = 2$ industries — meaning $2(2-1)/2 = 1$ edges — each; in Italy, we identify $c^{IT} = 2$ clusters made up by $m^{IT}_1 = m^{IT}_2 = 3$ — and $3(3-1)/2 = 3$ edges — each:

$$\Omega^{IT} = \{\Omega^{IT}_1 = \{(A,B),(A,C),(B,C)\}, \quad \Omega^{IT}_2 = \{(D,E),(D,F),(E,F)\}\}$$
$$\Omega^{DE} = \{\Omega^{DE}_1 = \{(A,B)\}, \quad \Omega^{DE}_2 = \{(C,D)\}, \quad \Omega^{DE}_3 = \{(E,F)\}\}$$

The directed similarity between Italy and Germany is a weighted average of the ratio of the cardinality of $\Omega^{it}_{\alpha} \cap \Omega^{DE}$ to the cardinality of Ω^{IT}_{α}, ($\alpha^{IT} = 1, 2$), the weights being the importance, relative to the whole inter-industry network, of each cluster in terms of the number of its members:

$$\sigma^{it}_{DE} = \sum_{\alpha^{IT}=1}^{2} \frac{\left|\Omega^{IT}_{\alpha} \cap \Omega^{DE}\right|}{\left|\Omega^{IT}_{\alpha}\right|} \frac{m^{IT}_{\alpha}}{N^{IT}} = \frac{1}{3}\frac{3}{6} + \frac{1}{3}\frac{3}{6} = \frac{1}{3}$$

[29]The total number of edges in the network is computed this way because, in this example, we are dealing with an undirected network; in the case of industry clusters derived from intermediate production flows, on the contrary, the direction of the edges is relevant, and therefore their total number is given by $\omega^{\varphi}_{\alpha} = m^{\varphi}_{\alpha}(m^{\varphi}_{\alpha} - 1)$. The same will hold below, in the case of the total number of significant edges.



(a) Industries grouped in clusters

(b) Membership matrix

(c) Initial similarity matrix

(d) Similarity matrix after first update ($C_1 = de + nl$)

(e) Dendrogram

Figure 6: Similarity measure of clustering structure: Example

where N^{IT} is the number of industries, excluding those which are isolated nodes in the clustering structure of the country. In the same way, we can compute the directed similarity between Germany and Italy, $\sigma_{IT}^{DE} = 2/3$, the undirected similarity $\sigma_{DE,IT}$ being a simple arithmetic average between the two:

$$\sigma_{DE,IT} = \frac{1}{2}\left(\sigma_{DE}^{IT} + \sigma_{IT}^{DE}\right) = \left(\frac{1}{3} + \frac{2}{3}\right)\frac{1}{2} = 0.5$$

By repeating the same procedure, we get the similarity matrix shown in Figure 6c: consistently with the fact that Germany and Netherlands have exactly the same clustering structure, the corresponding similarity values are equal to 1.

SB is iteratively applied forward up to the point where no further agglomeration is possible; in the same way, starting from the complete clustering structure, SB is iteratively applied, backwards, to each corresponding similarity sub-matrix, in the search for particularly strong similarities between the countries of each group, again up to the point where no further agglomeration of sub-groups is possible. The results are summarised in a dendrogram. After each successive agglomeration, the similarity matrix (or sub-matrix) is updated.

In our toy-example, the first application of SB to the complete similarity matrix leads to the identification of three 'clusters': $\{C_1 = (DE, NL), C_2 = (FR), C_3 = (IT)\}$. We then proceed to the update of the similarity matrix; for instance, by taking $C_{13} = (DE, NL, IT)$:

$$\sigma_{IT,NL}^{DE} = \sum_{\alpha^{DE}=1}^{m_{\alpha}^{DE}} \frac{|\Omega_{\alpha}^{DE} \cap \left(\Omega^{IT} \cap \Omega^{NL}\right)|}{\omega_{\alpha}^{DE}} = 1\frac{2}{6} + 0\frac{2}{6} + 1\frac{2}{6} = \frac{2}{3}$$

$$\sigma_{DE,IT,NL} = \frac{1}{3}\left(\sigma_{IT,NL}^{DE} + \sigma_{DE,NL}^{IT} + \sigma_{DE,IT}^{NL}\right) = \frac{1}{3}\frac{5}{3} \simeq 0.56$$

By applying SB to the updated matrix (shown in Figure 6d) we conclude that no further forward agglomeration is possible. The same holds for backwards iterations: the iterative process is thus concluded, and the resulting dendrogram is shown in Figure 6e.

Coming back to the real world, we applied the SB algorithm on a set of square Input-Output Tables for domestically produced output at basic prices for 2005 of 22 EU countries,[30] built an associated *industry* × *country* membership matrix,[31] obtained a symmetric *country* × *country* similarity

[30]Of course it would have been preferable to perform this exercise on all EU27 countries, however the required data were not available. See Table 18 in Appendix B for details on data availability.

[31]See Table 22 in Appendix B.

matrix, and then applied SB again on this latter matrix to find clusters of countries based on the similarity of their industry-clustering structures. This procedure lead to the identification of six communities (see the dendrogram in Figure 7). We can now examine these groups with the aid of Figure 8 and Table 9.

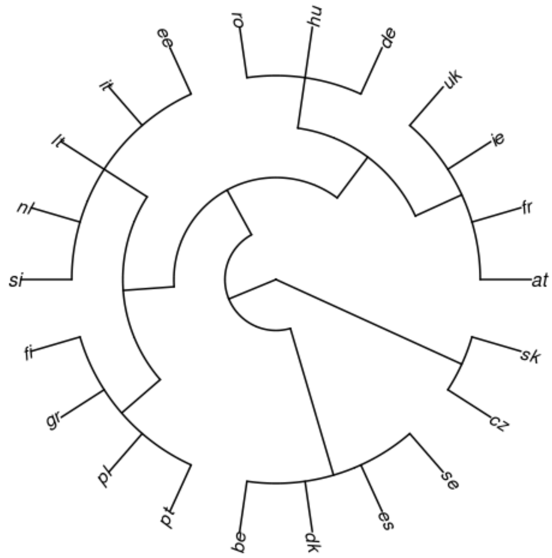
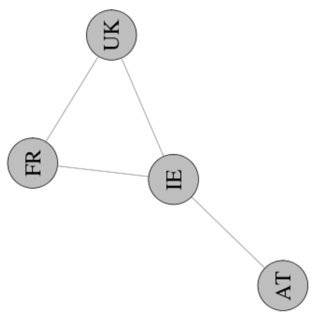


Figure 7: Hierarchical similarity in clustering of industries across countries
(based on square Input-Output tables for domestic output at basic prices, year 2005)
Source: Own computation based on EUROSTAT SUIOT Database

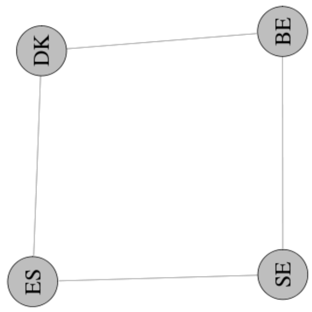
Table 9: Clusters breakers and enlargers (EU27 for the year 2005)
(Similarity between countries' interindustry-clustering structures)

Cluster	Breakers	Broken pieces (a)	(b)	(c)	Enlargers	Entrants
1	FR	AT, IE	UK			
	IE	AT	FR	UK		
	UK	AT, IE	FR			
2	BE	DK	ES, SE			
	ES	BE, DK	SE			
	SE	BE, DK	ES			
4					DE	PL, PT
5	IT	EE, LT	NL, SI			
6	GR	FI	PL, PT			

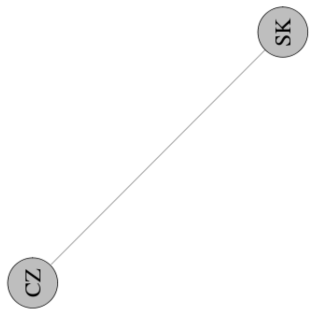
Based on Input-Output Tables for domestic output at basic prices for the year 2005
Not available from EU27:
Bulgaria (BG), Cyprus (CY), Luxembourg (LU), Latvia (LV), Malta (MT)
Source: Own computation based on EUROSTAT SUIOT Database



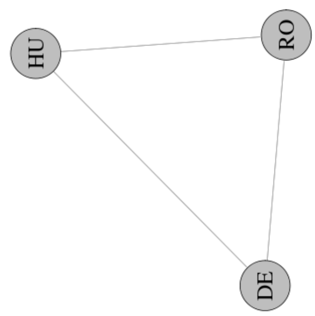
(a) Cluster 1



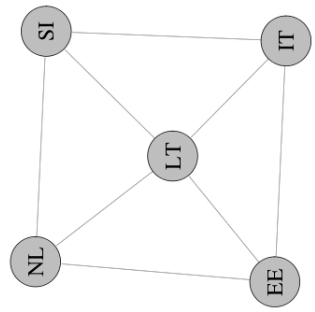
(b) Cluster 2



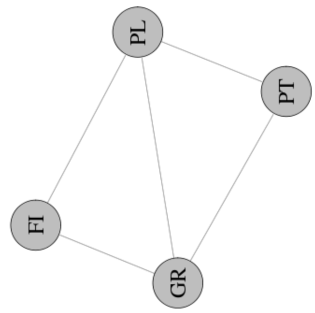
(c) Cluster 3



(d) Cluster 4



(e) Cluster 5



(f) Cluster 6

Figure 8: Graphs for similarity-based country clusters, 2005

Cluster 1 is composed by Austria (AT), France (FR), Ireland (IE) and United Kingdom (UK). By looking at Figure 8a we can see that all countries are quite similar to each other with the exception of Austria, which is marginally connected to the rest of the cluster via its similarity to Ireland. And in fact, Table 9 tells us that Austria is the only node that, if removed, would not break the cluster. On the contrary, removing Ireland would disintegrate it, with all the remainder countries going into different clusters; removing either France or UK would keep Ireland and Austria together while separating them from the remainder country.

Cluster 2 includes Belgium (BE), Denmark (DK), Spain (ES) and Sweden (SE). Looking at Figure 8b we see that the cluster has a pair-wise chain structure connecting Belgium with Denmark, Denmark with Spain, Spain with Sweden and Sweden back with Belgium. Denmark is the only node which, if removed, would not break the cluster. Removing Belgium would leave Denmark alone while keeping Spain and Sweden connected, while removing one of the latter two countries would leave the other one alone while keeping together Belgium and Denmark.

In Cluster 3 we find Czech Republic (CZ) and Slovakia (SK).

Cluster 4 is composed by Germany (DE), Hungary (HU) and Romania (RO). As can be seen from Figure 8d, the three countries are quite interconnected, and in fact removing none of them would break the cluster. However, Table 9 shows a peculiarity of this group: it has one node, namely Germany, which, if removed, would enlarge the cluster connecting Romania and Hungary with Poland and Portugal.

Cluster 5 includes Estonia (EE), Italy (IT), Lithuania (LT), Netherlands (NL) and Slovenia (SI). Figure 8e shows a structure quite similar to that of Cluster 2, with the difference that there we have a central node, Lithuania, connected to all the others. However, Italy is the only country that, if removed, would break the group, leaving together Estonia and Lithuania on the one side and Netherlands and Slovenia on the other.

Finally, Cluster 6 is composed by Finland (FI), Greece (GR), Poland (PL) and Portugal (PT). Also in this case, Figure 8f shows a structure similar to that of Cluster 2, but with an additional diagonal edge connecting Poland and Greece. The latter is the only node that, if removed, breaks the cluster, leaving Finland alone while keeping Poland and Portugal together.

Looking at these clusters from a geographical perspective, it is possible to note that four of them — clusters 2, 5, 6 and, to a smaller extent, 1 — connect Northern, Central or Eastern countries to the Mediterranean. The exceptions are cluster 3, whose origins are quite clear, and 4. In the latter case, however, we saw that removing Germany would make Portugal enter the cluster, though creating a connection with the periphery.

The same methodology used for computing similarities between countries on the basis of their industry-clustering structure was used to compute bilateral similarities between each of the 22 considered countries on the one side, and EMU and EU27 on the other. Results are shown in the first and second column, respectively, of Table 10.

Table 10: Similarity in the clustering structure of individual countries with respect to EMU and EU27 Input-Output Tables for 2005

Euro Area (EMU/EA17)			European Union (EU27)		
EU Code	Country	Similarity	EU Code	Country	Similarity
PT	Portugal	0.5536	UK	United Kingdom	0.5010
DE	Germany	0.5183	HU	Hungary	0.4980
BE	Belgium	0.4531	DK	Denmark	0.4806
ES	Spain	0.4296	PT	Portugal	0.4776
NL	Netherlands	0.4237	FR	France	0.4596
AT	Austria	0.4112	DE	Germany	0.4501
EE	Estonia	0.4103	ES	Spain	0.4310
IT	Italy	0.4078	IT	Italy	0.4304
FI	Finland	0.4076	PL	Poland	0.4249
FR	France	0.3979	BE	Belgium	0.4160
SI	Slovenia	0.3852	AT	Austria	0.4072
SK	Slovakia	0.3696	FI	Finland	0.4033
IE	Ireland	0.3458	EE	Estonia	0.3967
GR	Greece	0.2933	SI	Slovenia	0.3963
			NL	Netherlands	0.3956
			IE	Ireland	0.3913
			CZ	Czech Republic	0.3602
			SK	Slovakia	0.3574
			RO	Romania	0.3424
			LT	Lithuania	0.3102
			SE	Sweden	0.3089
			GR	Greece	0.3072

Not available from EMU/EU27: Bulgaria (BG), Cyprus (CY), Luxembourg (LU), Latvia (LV), Malta (MT)
Source: Own elaboration based on EUROSTAT SUIOT Database

The country sharing the highest proportion of significant edges with the EMU is Portugal (55.36%), followed by Germany (51.83%) and Belgium (45.31%). The lower limit of the range of variation is the 29.33% of similarity between EMU and Greece. Greece is also the country which shares the lowest number of significant edges with the EU27 (30.72%). In this case, the country showing the greatest similarity with EU27 is UK (50.10%), followed by Hungary (49.80%) and Denmark (48.06%). It is straightforward to notice that the first three position in this ranking are occupied by countries which are not members of the Monetary Union. The first country following is Portugal, the EMU member showing the highest similarity to EU27. It is followed by France, whose proportion of significant edges shared with EU27 (45.96%) is much higher than that shared with EMU (39.79%).[32] Germany behaves the opposite, sharing more edges with EMU (51.83%) than with EU27 (45.01%).[33]

Alongside measuring similarity among countries' clustering structures, it may prove useful to detect whether there are couples or groups of industries which are persistently clustered together *across* countries. But differently from the 'mini-clusters' of Section 3, we have 22 (one for each EU country of the sample in 2005) instead of 4 (EU27-2000, EU27-2007, EMU-2000, EMU-2007) membership vectors, which requires some refinement in the procedure followed. Take the example used at the beginning of this section as a point of departure. Figure 6b displays the membership matrix, reporting to which cluster (1, 2, 3) each industry (A, B, C, D, E, F) is associated in every country (DE, FR, IT, NL).

Departing from this membership matrix, consider building a binary matrix with all possible combinations of couples of industries (i.e. $(6 \times (6 - 1))/2 = 15$) in rows and with countries in columns (i.e. 4 countries), assigning a 1 when two industries are clustered together in the corresponding country, and 0 when they are not. For our example, Table 11 displays the rows in which a given couple of industries is clustered together in more than one country. The additional final column of the Table reports the relative frequency with which each couple of activities is associated (computed as the column-sum of each row divided by the total number of countries). This last column is interpreted as the probability of two industries being clustered together from among the country-group analysed.

Table 11: Industry pair-wise clustering probabilities across countries

Ind. 1	Ind. 2	de	fr	it	nl	Probability
E	F	1	1	1	1	1.00
A	B	1	0	1	1	0.75
C	D	1	1	0	1	0.75
B	C	0	1	1	0	0.50

The probabilities in the last column of Table 11 establish the minimum

[32] A similar relative pattern of similarities to EMU and EU27, respectively, is followed by Spain (42.96-43.10%), Italy (40.78-43.04%), Slovenia (38.52-39.63%), Greece, and Ireland (34.58-39.13%).

[33] A similar relative pattern of similarities to EMU and EU27, respectively, is followed by Portugal, Belgium (45.31-41.60%), Netherlands (42.37-39.56%), Austria (41.12-40.72%), Estonia (41.03-39.67%), Slovakia (36.96-35.74%) and, to a smaller extent, Finland (40.76-40.33%).

percentage of countries that have a given couple of industries associated. This minimum percentage is called the (inferior) 'limit' level of persistence. Then, for each 'limit' value between 0 and 1 it is possible to apply the following algorithm. Select the rows from the matrix of par-wise clustering probabilities with probability higher or equal than the chosen 'limit' value. Iteratively group those rows which share common industries, computing a corresponding joint probability as the product of the respective row-wise individual probabilities (e.g. if $limit = 0.5$, then industries A, B, C, D will be grouped together, their joint probability being $0.75 \times 0.75 \times 0.5 = 0.28$).

But given that industries are associated in *couples* and then *grouped* together, it is interesting to know to which extent countries have *each of* the industry couples that compose the group associated. Thus, for every industry group compute also: (i) the coefficient of variation of the number of couples of industries associated in each country from among those being clustered together (e.g. if $limit = 0.7$, the clustering of industries A, B has a corresponding coefficient of variation equal to 0.67, as only in one country — fr — industries A, B are not associated), and (ii) the number of countries in which all industries being clustered together are associated (e.g. if $limit = 0.7$, all countries have industries E, F associated, so the corresponding value is 4). Finally, identify those countries satisfying (ii).

For our example, the results of these computations can be summarised in Table 12, for limit values of 0.7 and 0.5.

Table 12: Persistence of clustering

Limit = 0.7					
Industries clustered	Reference cluster	(1)	(2)	(3)	Countries included in (3)
E, F	C3	1.00	0.00	4	de, fr, it, nl
A, B	C1	0.75	0.67	3	de, it, nl
C, D	C2	0.75	0.67	3	de, fr, nl
Limit = 0.5					
Industries clustered	Reference cluster	(1)	(2)	(3)	Countries included in (3)
E, F	C3	1.00	0.00	4	de, fr, it, nl
A, B, C, D	C1+C2	0.28	0.00	0	

References:
Limit Minimum % of countries that have a given couple of industries associated
(1) Joint probability that couples of industries clustered together are associated in a country
(2) Coefficient of variation of the number of couples of industries associated in each country from among those being clustered together
(3) Number of countries in which all industries being clustered together are associated

Note that by lowering the inferior limit bigger groups tend to appear (as is case of cluster $C1 + C2$ in Table 12 for $Limit = 0.5$ with respect to the individual clusters $C1$ and $C2$ for $Limit = 0.7$), but within these it may be the case that not a single country has *all* the industries being grouped together pair-wise associated (as is again the case of cluster $C1 + C2$ in Table 12 for $Limit = 0.5$). Hence, lowering the limit value has the appeal of obtaining bigger clusters with the risk of these being increasingly artificial.

A Table like 12 can be useful to have a birds' eye view of the *persistence* of industries clustered together in different countries. The interplay between columns (1) and (2) gives a summarising idea of the joint probability and dispersion of each activity-group, while column (3) checks to which extent the industry-group found is altogether part of a single cluster in every country.

Turning now to the real world, Table 13 reports the results of applying the procedure just described to the *industry* $\times$ *country* membership matrix — containing the clustering structures of 22 EU economies for 2005 — for limit values of $\{0.6, 0.7, 0.8\}$.

With an inferior limit value of 0.8 there are only four industry-groups persistently clustered together, which extends to eight and ten groups as the limit value shifts to 0.7 and 0.6, respectively. By lowering the limit value some reference clusters increase in size at the cost of reducing joint probability, increasing dispersion and fully representing less countries (e.g. the Agri-Food and Heavy Machinery reference clusters).

Observing in Table 13 industry-groups corresponding to a limit value of 0.7, for each of them at least 50% of the countries have the corresponding activities clustered together, conforming a skeleton of linkages of important persistence. Moreover, note the positive correlation between columns (1) and (3), i.e. the greater the number of countries having all industries of a group associated, the higher the joint probability that couples of industries of a group belong to the same cluster. Instead, dispersion as measured by column (2) shows an irregular pattern, as it depends on the *relative* clustering differences across countries for each industry-group. For a limit value of 0.7, reference clusters Agri-Food and Construction show the lowest dispersion.

Taken together and compared to the 'mini-clusters' reported in Table 2 of Section 3, these results show less persistence in a common core of industry-groups across all 22 EU countries. These findings suggest that applying a common template to all countries in order to identify regional clusters might be a biased procedure, in which the more a country complies with an imposed pre-defined template, the better its clustering performance score will be. Instead, it could prove more adequate to evaluate the clustering performance of individual regions in several EU countries by departing from a common (but restricted) core of industry-groups (e.g. a core set of reference clusters obtainable from Table 13), augmented by country-specific clustering patterns.

Table 13: Persistence of clustering of industries among 22 EU countries, year 2005

Limit = 0.8					
Industries clustered	Reference cluster	(1)	(2)	(3)	Countries included in (3)
A01, DA15	Agri-Food	1.000	0.000	22	at, be, cz, de, dk, ee, es, fi, fr, gr, hu, ie, it, lt, nl, pl, pt, ro, se, si, sk, uk
CB14, DI26, F45	Construction	0.669	0.299	18	at, be, cz, de, dk, ee, es, fr, gr, hu, ie, it, lt, nl, pt, ro, si, uk
J66, J67	Finance	0.864	0.407	19	at, be, cz, de, dk, ee, es, fi, fr, gr, hu, ie, it, lt, nl, pl, si, sk, uk
DB17, DB18	Dressing	0.818	0.482	18	at, de, dk, ee, es, fi, fr, gr, hu, it, lt, nl, pl, pt, ro, se, si, uk

Limit = 0.7					
Industries clustered	Reference cluster	(1)	(2)	(3)	Countries included in (3)
A01, DA15, H55	Agri-Food	0.597	0.337	17	at, be, de, dk, ee, es, gr, hu, ie, it, nl, pl, pt, ro, si, sk, uk
CB14, DI26, F45	Construction	0.669	0.299	18	at, be, cz, de, dk, ee, es, fr, gr, hu, ie, it, lt, nl, pt, ro, si, uk
J66, J67	Finance	0.864	0.407	19	at, be, cz, de, dk, ee, es, fi, fr, gr, hu, ie, it, lt, nl, pl, si, sk, uk
DB17, DB18	Dressing	0.818	0.482	18	at, de, dk, ee, es, fi, fr, gr, hu, it, lt, nl, pl, pt, ro, se, si, uk
DJ27, DJ28, DK29, DL31	Heavy Machinery	0.316	0.424	12	be, de, dk, es, fr, hu, it, pl, pt, ro, sk, uk
I60, I61, I62, I63	Transport	0.434	0.388	11	be, cz, dk, ee, es, fi, fr, it, nl, se, sk
I64, K72	ICT Services	0.773	0.555	17	be, de, dk, ee, es, fi, fr, hu, ie, it, lt, nl, pl, se, si, sk, uk
A02, DD20	Wood	0.727	0.627	16	at, cz, de, ee, es, fi, fr, hu, ie, lt, pl, pt, ro, se, si, uk

Limit = 0.6					
Industries clustered	Reference cluster	(1)	(2)	(3)	Countries included in (3)
A01, B05, DA15, H55	Agri-Food	0.154	0.448	13	at, de, dk, ee, hu, ie, it, pl, pt, ro, si, sk, uk
CB14, DI26, F45	Construction	0.669	0.299	18	at, be, cz, de, dk, ee, es, fr, gr, hu, ie, it, lt, nl, pt, ro, si, uk
J65, J66, J67	Finance	0.401	0.479	14	at, be, de, dk, ee, es, hu, ie, it, lt, pl, si, sk, uk
DB17, DB18	Dressing	0.818	0.482	18	at, de, dk, ee, es, fi, fr, gr, hu, it, lt, nl, pl, pt, ro, se, si, uk
DJ27, DJ28, DK29, DL31, DN37	Heavy Machinery	0.137	0.387	7	dk, es, hu, it, pl, pt, sk
I60, I61, I62, I63	Transport	0.188	0.452	11	be, cz, dk, ee, es, fi, fr, it, nl, se, sk
A02, DD20	Wood	0.727	0.627	16	at, cz, de, ee, es, fi, fr, hu, ie, lt, pl, pt, ro, se, si, uk
CA10, E40	Energy	0.636	0.774	14	at, be, cz, de, fi, fr, gr, hu, ie, pl, ro, si, sk, uk
G52, K70	Trade-Real Estate	0.636	0.774	14	at, cz, de, ee, es, fi, fr, hu, ie, it, lt, si, sk, uk
DE22, I64, K72, K74, O92	ICT-Media Services	0.127	0.420	6	be, ee, es, fi, lt, pl

References:
Limit Minimum % of countries that have a given couple of industries associated
(1) Joint probability that couples of industries clustered together are associated in a country
(2) Coefficient of variation of the number of couples of industries associated in each country from among those being clustered together
(3) Number of countries in which all industries being clustered together are associated

Source: Own computation based on EUROSTAT SUIOT and National Accounts Databases

6 The trade connection

Applying SB to find communities of commodities, based on the similarity of their country-clustering structure in commodity-specific intra-EU trade matrices lead to the identification of five groups (see the dendrogram in Figure 9). We can now examine these groups with the aid of Figure 10 and of Table 14.

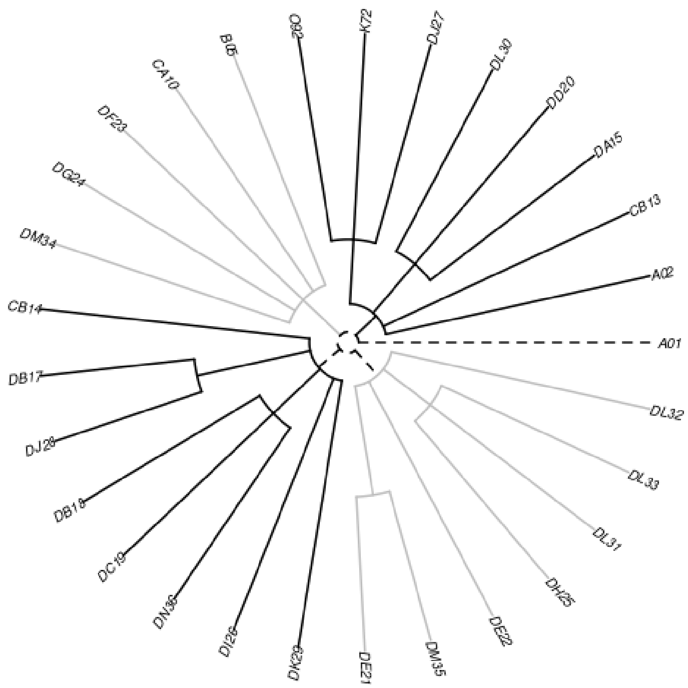
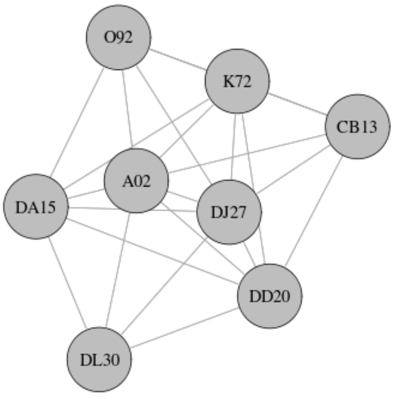


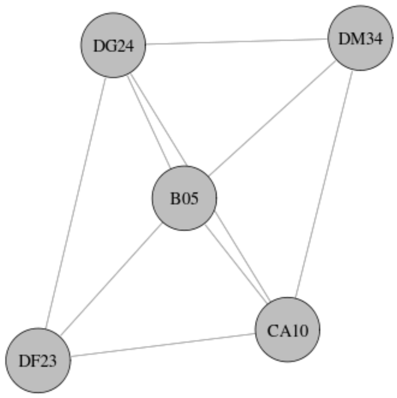
Figure 9: Dendrogram of similarity-based industry clusters, 2007
(Commodity codes reported in Table 19 of Appendix B)

Cluster 1 is made up by A01 Agriculture only.

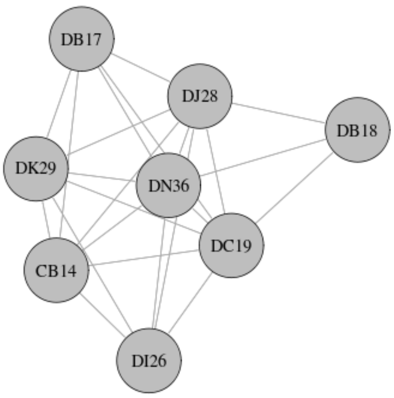
Cluster 2 includes A02 Forestry, CB13 Metal-mining, DA15 Food-beverages, DD20 Wood, DL30 Office-machinery-PC, DJ27 Iron-steel-aluminium-tub., K72 Computer-services and O92 Arts-entertainment. Looking at Figure 10a we can see that the structure of the cluster is as follows. There are three vertices — K72, DA15 and DD20 — reciprocally connected forming a triangle. Within this triangle we find two nodes, A02 and DJ27, which are connected both to each other and to all vertices of the triangle. For each edge of the triangle, we have a marginal node connected to the two corresponding vertices and to both internal nodes — O92 for the edge connecting DA15 and K72, CB13 for the edge connecting K72 and DD20, and DL30



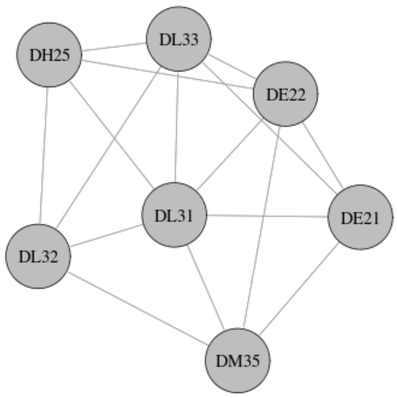
(a) Cluster 2



(b) Cluster 3



(c) Cluster 4



(d) Cluster 5

Figure 10: Graphs for similarity-based industry clusters, 2007

Table 14: Clusters breakers and enlargers, 2007

Cluster	Breakers	Broken Pieces (a)	(b)	(c)
2	A02	CB13,DJ27,K72,O92	DA15,DD20,DL30	
	DA15	A02,DD20,DL30	CB13,DJ27,K72,O92	
	DD20	A02,CB13,DJ27,K72,O92	DA15,DL30	
	DJ27	A02,DA15,DD20,DL30	CB13,K72,O92	
3	B05	CA10,DG24,DM34	DF23	
	CA10	B05,DG24,DM34	DF23	
	DG24	B05,CA10	DF23,DM34	
4	DC19	CB14,DB17,DI26,DJ28,DK29,DN36	DB18	
	DJ28	CB14,DB18,DC19,DI26,DK29,DN36	DB17	
	DN36	CB14,DB17,DC19,DI26,DJ28,DK29	DB18	
5	DE21	DE22,DH25,DL31,DL32,DL33	DM35	
	DE22	DE21,DM35	DH25,DL31,DL32,DL33	
	DL31	DE21,DE22,DL32,DL33	DH25	DM35
	DL32	DE21,DE22,DL31,DL33	DH25	DM35
	DL33	DE21,DL32,DM35	DE22,DH25,DL31	
	DM35	DE21	DE22,DH25,DL31,DL32,DL33	

Similarity-based clusters; similarity between commodities based on similarity of their international trade pattern-clustering structure. No enlargers were found.
Source: Own computation based on EUROSTAT External Trade Database

for the edge connecting DD20 and DA15. Both internal nodes, if removed, would break the cluster. The same holds for two of the three vertices of the triangle: DD20 and DA15. From a 'commodity-chain' point of view, this group includes three well defined subgroups: (i) A02 and DD20 (the Wood mini-cluster); (ii) CB13 Metal-mining and DJ27 Iron-steel-aluminium-tub.; and (iii) DL30 Office-machinery-PC and K72 Computer-services.

Cluster 3 includes B05 Fishing, CA10 Coal Mining, DF23 Petroleum-refinery, DG24 Chemicals-pharma and DM34 Motor-vehicles. The latter four nodes form a square in a pair-wise chain fashion, with an additional diagonal edge between DG24 and CA10. Moreover, node B05 is in the middle of the square, connected to all its vertices. The central node (B05) and the two connected diagonal ones (DG24 and CA10) are the only commodities which, if removed, would break the cluster. Here there is one 'commodity-chain' subgroup composed by DF23 Petroleum-refinery, DG24 Chemicals-pharma and DM34 Motor-vehicles.

Cluster 4 is made up by CB14 Stone-sand-clay-minerals, DB17 Textiles, DJ28 Structural-metal-products, DB18 Clothing, DC19 Leather, DN36 Furniture-Sports-Toys, DI26 Glass-clay-cement-ceramic and DK29 Mechanical-machinery. DN36 is the central node, connected to all the others. Around it, DK29, DJ28, DC19 and CB14 form a square with both diagonals connected (and all vertices connected to the central one). At the margin, we have DB17 which is connected with all the vertices of the square and with the central node; DI26 which is connected to all the vertices of the square with the exception of DK29 (and to the central node); and DB18 which is connected with the central node and with two of the vertices of the square: DJ28 and DC19. These latter two nodes, together with the central one, are the only ones whose removal would break the cluster. In this cluster, we can identify three 'commodity-chain' subgroups: (i) CB14 and DI26: the Construction mini-cluster; (ii) DB17, DB18 and DC19: the MC:Dressing mini-cluster; and (iii) DJ28 Structural-metal-products and DK29 Mechanical-machinery.

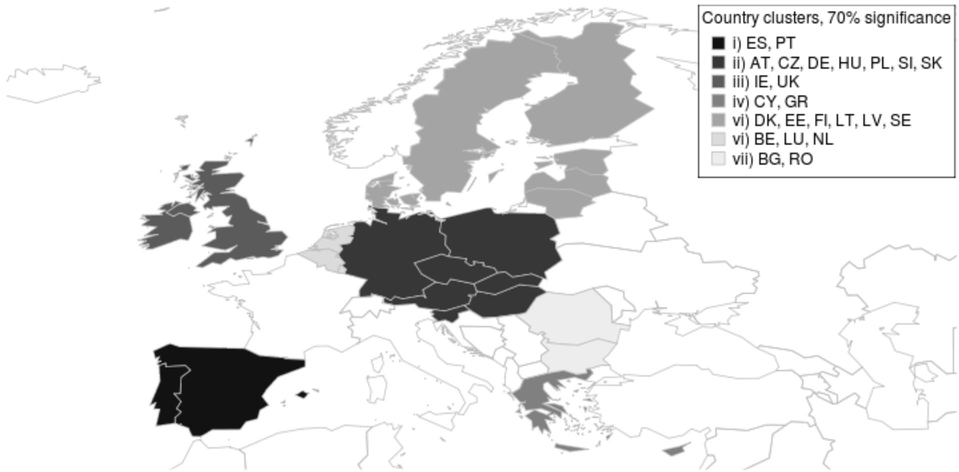
Finally, cluster 5 includes DL31 Electrical-machinery, DE21 Paper, DM35 Ships-railway-aircrafts, DL32 ICT-equipment, DH25 Rubber-plastics, DL33 Medical-precision-equip. and DE22 Publishing-printing. The first node, DL31, is connected to all the others; it is placed at the centre of an hexagon, whose vertices are the remainder six nodes in the order we gave them, connected in a pair-wise chain fashion. In addition to the edges connecting any vertex of the hexagon with the centre and its two neighbours, DH25 is connected to DE22 which is in its turn connected to DM35; DL32 is connected to DL33 which is connected to DM35. The removal of any node, with the exception of DH25, would break the cluster. Also in this last cluster, it is possible to isolate two 'commodity-chain' subgroups: (i) DE21 Paper and DE22 Publishing-printing; and (ii) DL31 Electrical-machinery, DL32 ICT-equipment and DL33 Medical-precision-equip.

Some concluding remarks can be drawn with the aid of Figures 11a and 11b, summarising the results of Table 15, which is the outcome of applying the methodology described in Section 5 (pp. 46 and ff.) starting from the country-clustering structure of each commodity-specific trade matrix for the year 2007.[34]

At the 80% significance level we found five country groups: (i) Spain, Portugal; (ii) Austria, Czech Republic, Hungary, Portugal, Slovakia; (iii) Ireland, UK; (iv) Cyprus, Greece; and (v) Denmark, Estonia, Finland, Lithuania, Latvia, Sweden. At the 70% significance level, group (ii) is larger, also including Germany and Slovenia, and two additional groups are present:(vi) Belgium, Luxembourg, Netherlands; and (vii) Bulgaria, Romania. Groups (vi) and (vii). It is immediately apparent that such groups have a strong geographical connotation, being conformed by adjacent countries; they account for 17 countries at the 80% significance level and 24 countries at the 70% one, showing that in the whole EU27 area there are strong geographical trade linkages — which are most likely independent of the EU institution.

[34]See the corresponding *country* × *commodity* membership matrix in Table 23 of Appendix B.

(a) *Limit* = 0.8 (80% significance level)



(b) *Limit* = 0.7 (70% significance level)

Figure 11: Persistence in country-clusters of intra-EU external trade (2007)

Table 15: Persistence of clustering of EU27 countries in external trade across 29 CPA commodities (year 2007)

Limit = 0.8				
Countries clustered	(1)	(2)	(3)	Products included in (3)
ES, PT	0.966	0.192	28	A01, A02, B05, CA10, CB13, CB14, DA15, DB17, DB18, DC19, DD20, DE21, DE22, DF23, DG24, DH25, DI26, DJ27, DJ28, DK29, DL30, DL31, DL32, DL33, DM34, DM35, DN36, K72
AT, CZ, HU, PL, SK	0.443	0.225	20	A02, CA10, CB14, DA15, DB17, DB18, DC19, DD20, DE22, DG24, DH25, DI26, DJ28, DK29, DL31, DL33, DM34, DN36, K72, O92
IE, UK	0.966	0.192	28	A01, A02, B05, CA10, CB14, DA15, DB17, DB18, DC19, DD20, DE21, DE22, DF23, DG24, DH25, DI26, DJ27, DJ28, DK29, DL30, DL31, DL32, DL33, DM34, DM35, DN36, K72, O92
CY, GR	0.828	0.465	24	A01, A02, CB13, CB14, DA15, DB18, DC19, DD20, DE21, DE22, DF23, DG24, DH25, DI26, DJ27, DJ28, DK29, DL30, DL31, DL32, DL33, DN36, K72, O92
DK, EE, FI, LT, LV, SE	0.409	0.204	20	A01, CB14, DA15, DB17, DB18, DC19, DD20, DE21, DE22, DF23, DH25, DI26, DJ28, DK29, DL30, DL31, DL32, DL33, DN36, O92

Limit = 0.7				
Countries clustered	(1)	(2)	(3)	Products included in (3)
ES, PT	0.966	0.192	28	A01, A02, B05, CA10, CB13, CB14, DA15, DB17, DB18, DC19, DD20, DE21, DE22, DF23, DG24, DH25, DI26, DJ27, DJ28, DK29, DL30, DL31, DL32, DL33, DM34, DM35, DN36, K72
AT, CZ, DE, HU, PL, SI, SK	0.037	0.285	16	A02, CB14, DA15, DB17, DB18, DC19, DD20, DE22, DH25, DI26, DJ28, DK29, DL31, DL33, DN36, O92
IE, UK	0.966	0.192	28	A01, A02, B05, CA10, CB14, DA15, DB17, DB18, DC19, DD20, DE21, DE22, DF23, DG24, DH25, DI26, DJ27, DJ28, DK29, DL30, DL31, DL32, DL33, DM34, DM35, DN36, K72, O92
CY, GR	0.828	0.465	24	A01, A02, CB13, CB14, DA15, DB18, DC19, DD20, DE21, DE22, DF23, DG24, DH25, DI26, DJ27, DJ28, DK29, DL30, DL31, DL32, DL33, DN36, K72, O92
DK, EE, FI, LT, LV, SE	0.061	0.319	20	A01, CB14, DA15, DB17, DB18, DC19, DD20, DE21, DE22, DF23, DH25, DI26, DJ28, DK29, DL30, DL31, DL32, DL33, DN36, O92
BE, LU, NL	0.576	0.379	16	A02, CB13, DA15, DC19, DE22, DF23, DG24, DH25, DI26, DJ27, DK29, DL31, DM34, DM35, K72, O92
BG, RO	0.759	0.574	22	A01, CB13, CB14, DA15, DB17, DC19, DD20, DE21, DE22, DF23, DG24, DH25, DJ27, DJ28, DL30, DL31, DL32, DL33, DM34, DM35, DN36, K72

Limit = 0.6				
Countries clustered	(1)	(2)	(3)	Products included in (3)
ES, FR, IT, PT	0.271	0.417	14	A01, CB14, DB18, DE21, DG24, DH25, DI26, DJ28, DK29, DL31, DL32, DL33, DM34, DN36
IE, UK	0.966	0.192	28	A01, A02, B05, CA10, CB14, DA15, DB17, DB18, DC19, DD20, DE21, DE22, DF23, DG24, DH25, DI26, DJ27, DJ28, DK29, DL30, DL31, DL32, DL33, DM34, DM35, DN36, K72, O92
DK, EE, FI, LT, LV, SE	0.061	0.319	20	A01, CB14, DA15, DB17, DB18, DC19, DD20, DE21, DE22, DF23, DH25, DI26, DJ28, DK29, DL30, DL31, DL32, DL33, DN36, O92
BE, LU, NL	0.576	0.379	16	A02, CB13, DA15, DC19, DE22, DF23, DG24, DH25, DI26, DJ27, DK29, DL31, DM34, DM35, K72, O92
AT, BG, CY, CZ, DE, GR, HU, PL, RO, SI, SK	0.000	0.259	2	DA15, DD20

References:

Limit	Minimum % of products for which a given couple of countries are associated
(1)	Joint probability that couples of countries clustered together are associated for trade in a given product
(2)	Coefficient of variation of the number of couples of countries associated for each product from among those being clustered together
(3)	Number of products in which all countries being clustered together are associated

Source: Own computation based on EUROSTAT COMEXT (External Trade) Database

Only three countries out of 27 — namely France, Italy and Malta — do not belong to any group at nor significance level, showing that they do not have special trade relations with other member countries, with their trade partners changing with the traded commodities.

In order to check the robustness of our results, we repeated the exercise with the trade matrices for the years 2008 to 2010. The same country cluster also emerge in 2009 and 2010, the only difference being the absence of Poland in the former year. Year 2008 represents a structural break, showing a different country-clustering structure, with some groups breaking in smaller ones; to give only one example, Austria and Germany leave their group and form an isolated cluster, even at the 80% significance level.

7 Summary of findings and concluding remarks

As stated in the Introduction, the aim of the present paper was that of revealing the industry clustering structure of the European Union and of singling out the similarities and dissimilarities in the structures and the trade interrelations of its constituent members. The whole analysis has been mainly carried out with the aid of two devices: spectral bisection for community detection, and the measurement of similarities between the emerging topological structures.

From a methodological point of view, one of the original contributions of the paper has been the application of SB both to flows and to similarity matrices, coupled with the measure of similarity, based on shared significant edges, we developed. We grouped countries according to the similarity of their industry clustering structure, and commodities according to the similarity of their intra-EU external trade patterns.

A second methodological contribution has been to exploit the juxtaposition between industry clusters and their associated subsystems, which in our view is a powerful device for singling out the structural features of any economic system.

We can now briefly go through some of the key empirical findings.

Section 3 lists the results of applying SB to EMU and EU27 IO networks. Besides the presence of 12 mini-clusters representing their *skeleton* (cfr. Table 2), we devised:

i) The emergence, in EU27 (2007), of a well-defined Pharma-Hi Tech cluster — including those industries producing specialised machinery, chemical products, and R&D — as opposed to the traditional Heavy Machinery cluster, which still plays a prominent role in the EU27;

ii) The tendency, both in the EMU and in the EU27, of Trade and Transport industries to converge into a specific cluster, with exactly the same composition;[35]

iii) The progressive tertiarisation of the European economies, with EU27 developing an increasingly autonomous and interconnected Services cluster. In the EMU, which does not include the UK, i.e. the European country with the most developed financial sector, this translated into a sharp separation of the MC:Finance from the rest of the Services cluster.

iv) More in general, the importance of the structure of the UK economy in shaping the development of the clustering topology of the EU27 with respect to that of the EMU.

Section 4 further inspected the above mentioned results by means of a subsystem-level and vertically integrated labour productivity analysis, leading to the following conclusions:

i) The relevance of intra-cluster intermediate consumption — see item 2, page 33;

ii) The primary role of Heavy Machinery, Specialised Machin. and Chemicals clusters in EU27 exports — see item 5, page 33;

iii) To an important extent, labour from intra-cluster industries remains in intra-cluster subsystems — see item 8, page 33;

iv) In the EMU, Manufacturing, ICT-Services, Trade and Financial Services are particularly dynamic subsystems from a labour productivity point of view; in the EU27, some Manufacturing subsystems showed a worse performance while some Services subsystems, which were sluggish in the EMU, performed better — again confirming the increased importance of services when the unit of analysis shifts from the EMU to the EU27; see Section 4.1.

Section 5 shows the results of the comparison between the industry clustering structure of 22 countries of the EU27. On the basis of the topological structure of their IO tables, we grouped them into six communities:

i) Czech Republic and Slovakia still have a very similar structure — it would be interesting to see whether Western and Eastern Germany still have a very different one, despite the re-unification;

ii) The group made up by Germany, Romania and Hungary is the only one (together with the AT-FR-UK-IE one) which does not include a Mediterranean country, though removing Germany would make Portugal (and Poland) enter the cluster;

[35] The only exception being the presence of DL30 Office-machinery-PC in the EMU.

iii) The way in which countries grouped suggests that we found five (without considering former Czechoslovakia) 'typologies' of countries, playing symmetrical roles in different geographical areas. However, more evidence would be necessary to confirm this hypothesis.

Finally, Section 6 lists the results of the search for country clusters in commodity-specific trade matrices, and of the application of SB to the resulting similarities between commodities:

i) From the first part of the exercise, seven country groups emerged as persistently appearing in the same cluster for, at least, 70% of the traded commodities. These groups account for 24 EU27 countries; only France, Italy and Malta do not have persistent trade partners;
ii) From the second part of the exercise, four[36] commodity groups emerged as sharing a similar intra-EU external trade pattern. As emerges from the detailed description of such groups (see p. 50 to the end of the section), all groups include well defined, from a 'commodity-chain' point of view, subgroups, but there is no correspondence with the industry clusters found in section 3.

[36]Actually five, but the first one includes A01 Agriculture only.

Appendices

A Glossary of Notation

Table 16: Glossary of Notation

Symbol	Dimension	Description
		Section 2
$\mathbf{F} = [f_{ij}]$	$n \times n$	flow matrix , f_{ij} being the flows going from node i to j
$\mathbf{S} = [s_{ij}]$	$n \times n$	similarity matrix , s_{ij} being the similarity between node i and j
$\mathbf{s}_{(out)} = \mathbf{We}$	$n \times 1$	$(\mathbf{W} = \mathbf{F}, \mathbf{S})$
$\mathbf{s}_{(in)}^T = \mathbf{e}^T\mathbf{W}$	$1 \times n$	$(\mathbf{W} = \mathbf{F}, \mathbf{S})$
$\mathbf{W}^e = \mathbf{s}_{(out)}\mathbf{s}_{(in)}^T/m$	$n \times n$	averaged $\mathbf{W}$ $(\mathbf{W} = \mathbf{F}, \mathbf{S})$
$\mathbf{B} = \mathbf{W} - \mathbf{W}^e$	$n \times n$	modularity matrix $\mathbf{W}$ $(\mathbf{W} = \mathbf{F}, \mathbf{S})$
$\tilde{\mathbf{B}} = \mathbf{B} + \tilde{\mathbf{B}}$	$n \times n$	generalised modularity matrix
$\mathbf{m}$	$n \times 1$	membership vector
$Q = \mathbf{m}^T\mathbf{Bm}$	scalar	modularity
$\Delta Q = \mathbf{m}_\alpha^T\mathbf{B}^{(\alpha)}\mathbf{m}_\alpha$	scalar	additional contribution to modularity
$\mathbf{B}^{(\alpha)}$	$m_\alpha \times m_\alpha$	modified modularity matrix
m_α	scalar	number of members of group α
$\mathbf{z}^\star(\mathbf{A})$		Perron eigenvector of matrix $\mathbf{A}$
		Section 4
$\mathbf{U}$	$N \times N$ $p \times i$	Use Table for domestic output at basic prices
$\mathbf{V}$	$N \times N$ $p \times i$	Make Table for domestic output at basic prices
$\mathbf{z}$	$N \times 1$ $p \times 1$	Gross output by commodity in current basic prices
$\mathbf{g}$	$N \times 1$ $i \times 1$	Gross output by industry in current basic prices
$\mathbf{f}$	$N \times 1$ $p \times 1$	Final demand by commodity in current basic prices
$\mathbf{m}^T$	$1 \times N$ $1 \times i$	Imported intermediate inputs in current prices (cif)
$\mathbf{t}_z^T$	$1 \times N$ $1 \times i$	Taxes less subsidies on products
$\mathbf{t}_g^T$	$1 \times N$ $1 \times i$	Taxes less subsidies on production
$\mathbf{w}^T$	$1 \times N$ $1 \times i$	Compensation to employees (wages plus taxes on labour)
$\boldsymbol{\pi}^T$	$1 \times N$ $1 \times i$	Gross operating surplus
$\mathbf{l}^T$	$1 \times N$ $1 \times i$	Employment by industry (in persons employed)
$\mathbf{D}$	$N \times N$ $p \times i$	Market shares matrix by commodity
$\mathbf{d} = [d_i]$	$N \times 1$ $i \times 1$	Final demand by industry in current basic prices
$\mathbf{d}_c$	$N \times 1$ $i \times 1$	Private consumption by industry in current basic prices
$\mathbf{d}_g$	$N \times 1$ $i \times 1$	Government consumption by industry in current basic prices
$\mathbf{d}_k$	$N \times 1$ $i \times 1$	Gross fixed capital formation by industry in current basic prices
$\mathbf{d}_{vk}$	$N \times 1$ $i \times 1$	Changes in inventories and valuables by industry
$\mathbf{d}_x$	$N \times 1$ $i \times 1$	Exports by industry in current basic prices
$\mathbf{X} = [x_{ij}]$	$N \times N$ $i \times i$	Inter-industry intermediate consumption at basic current prices
$\mathbf{e}$	$N \times 1$	Sum vector across columns (all components equal to one)
$\mathbf{e}_i$	$N \times 1$	Column selector vector (all components equal to zero but the i-th. one, which is equal to one)

References: p stands for product or commodity, i stands for industry or activity.

All throughout the paper, vectors are indicated by lower case boldface characters (e.g. $\mathbf{v}$), are column vectors unless explicitly transposed (e.g. $\mathbf{v}^T$), while matrices are indicated by upper case boldface characters (e.g. $\mathbf{X}$), except for lower case characters with a hat (e.g. $\widehat{\mathbf{z}}$), indicating diagonal matrices with the vector elements on the main diagonal.

B Additional Tables

Table 17: European (Monetary) Union Memberships

European Union (EU27)		Euro Area (EMU/EA17)	
EU Code	Country	EU Code	Country
AT	Austria	AT	Austria
BE	Belgium	BE	Belgium
BG	Bulgaria	CY	Cyprus
CY	Cyprus	DE	Germany
CZ	Czech Republic	EE	Estonia
DE	Germany	ES	Spain
DK	Denmark	FI	Finland
EE	Estonia	FR	France
ES	Spain	GR	Greece
FI	Finland	IE	Ireland
FR	France	IT	Italy
GR	Greece	LU	Luxembourg
HU	Hungary	MT	Malta
IE	Ireland	NL	Netherlands
IT	Italy	PT	Portugal
LT	Lithuania	SI	Slovenia
LU	Luxembourg	SK	Slovakia
LV	Latvia		
MT	Malta		
NL	Netherlands		
PL	Poland		
PT	Portugal		
RO	Romania		
SE	Sweden		
SI	Slovenia		
SK	Slovakia		
UK	United Kingdom		

Table 18: Availability of square Input-Output Tables for domestic output at basic prices (Year 2005)

EU Code	Country	Type
AT	Austria	pp
BE	Belgium	pp
CZ	Czech Republic	pp
DE	Germany	pp
DK	Denmark	ii
EE	Estonia	pp
ES	Spain	pp
FI	Finland	ii
FR	France	pp
GR	Greece	pp
HU	Hungary	pp + ii
IE	Ireland	pp
IT	Italy	pp + ii
LT	Lithuania	pp
NL	Netherlands	ii
PL	Poland	pp
PT	Portugal	pp
RO	Romania	pp
SE	Sweden	pp
SI	Slovenia	pp
SK	Slovakia	pp
UK	United Kingdom	pp + ii

Type: ii (industry × industry), pp (product × product)
Not available from EU27: Bulgaria (BG), Cyprus (CY), Luxembourg (LU), Latvia (LV), Malta (MT)
Source: Own elaboration based on EUROSTAT SUIOT Database

Table 19: NACE Rev. 1 Nomenclature (2 digits)

NACE	Short Description	Long Description
A01	Agriculture	Agriculture, hunting and related service activities
A02	Forestry	Forestry, logging and related service activities
B05	Fishing	Fishing, operating of fish hatcheries and fish farms; service activities incidental to fishing
CA10	Coal Mining	Mining of coal and lignite; extraction of peat
CA11	Petroleum-gas-extraction	Extraction of crude petroleum and natural gas; service activities incidental to oil and gas extraction excluding surveying
CA12	Uranium	Mining of uranium and thorium ores
CB13	Metal-mining	Mining of metal ores
CB14	Stone-sand-clay-minerals	Other mining and quarrying
DA15	Food-beverages	Manufacture of food products and beverages
DA16	Tobacco	Manufacture of tobacco products
DB17	Textiles	Manufacture of textiles
DB18	Clothing	Manufacture of wearing apparel; dressing and dyeing of fur
DC19	Leather	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear
DD20	Wood	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
DE21	Paper	Manufacture of pulp, paper and paper products
DE22	Publishing-printing	Publishing, printing and reproduction of recorded media
DF23	Petroleum-refinery	Manufacture of coke, refined petroleum products and nuclear fuels
DG24	Chemicals-pharma	Manufacture of chemicals and chemical products
DH25	Rubber-plastics	Manufacture of rubber and plastic products
DI26	Glass-clay-cement-ceramic	Manufacture of other non-metallic mineral products
DJ27	Iron-steel-aluminium-tub.	Manufacture of basic metals
DJ28	Structural-metal-products	Manufacture of fabricated metal products, except machinery and equipment
DK29	Mechanical-machinery	Manufacture of machinery and equipment n.e.c.
DL30	Office-machinery-PC	Manufacture of office machinery and computers
DL31	Electrical-machinery	Manufacture of electrical machinery and apparatus n.e.c.
DL32	ICT-equipment	Manufacture of radio, television and communication equipment and apparatus
DL33	Medical-precision-equip.	Manufacture of medical, precision and optical instruments, watches and clocks
DM34	Motor-vehicles	Manufacture of motor vehicles, trailers and semi-trailers
DM35	Ships-railway-aircrafts	Manufacture of other transport equipment
DN36	Furniture-Sports-Toys	Manufacture of furniture; manufacturing n.e.c.
DN37	Recycling	Recycling
E40	Electricity-gas	Electricity, gas, steam and hot water supply
E41	Water	Collection, purification and distribution of water
F45	Construction	Construction
G50	Sale-repair-vehicles	Sale, maintenance and repair of motor vehicles and motorcycles; retail sale services of automotive fuel
G51	Wholesale-trade	Wholesale trade and commission trade, except of motor vehicles and motorcycles
G52	Retail-trade	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods
H55	Hotel-restaurant	Hotels and restaurants
I60	Transport-land	Land transport; transport via pipelines
I61	Transport-water	Water transport
I62	Transport-air	Air transport
I63	Storage-travel-agencies	Supporting and auxiliary transport activities; activities of travel agencies
I64	Post-telecomm.	Post and telecommunications

(continued on next page)

(continued from previous page)

NACE	Short Description	Long Description
J65	Finance	Financial intermediation, except insurance and pension funding
J66	Insurance	Insurance and pension funding, except compulsory social security
J67	Brokerage-credit-cards	Activities auxiliary to financial intermediat.
K70	Real-estate	Real estate activities
K71	Renting-equipment	Renting of machinery and equipment without operator and of personal and household goods
K72	Computer-services	Computer and related activities
K73	R&D	Research and development
K74	Business-services	Other business activities
L75	Public-admin.	Public administration and defence; compulsory social security
M80	Education	Education
N85	Health	Health and social work
O90	Refuse-disposal	Sewage and refuse disposal, sanitation and similar activities
O91	Membership-organisations	Activities of membership organisation n.e.c.
O92	Arts-entertainment	Recreational, cultural and sporting activities
O93	Personal-services	Other service activities
P95	Household-services	Private households with employed persons

Source: EUROSTAT SUIOT Database

Table 20: Input-Output Tables by Clusters of the European Monetary Union (EMU/EA17), years 2000, 2007
industry × industry tables at basic prices built adopting the fixed product sales structure assumption (in millions of EUR)

	Year 2000	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	ic	d_c	d_g	d_k	d_{vk}	d_x	g
C01	Agri-Food	376000	9324	3318	35910	8913	48671	8408	350	4003	424	495320	778353	373351	8592	-136	77998	1733478
C02	Construction	28651	274345	14294	23468	25559	115630	9787	490	11936	540	504699	106986	4220	616352	5999	58651	1296906
C03	Energy	34470	18956	73329	25761	27184	66959	18861	441	7592	583	274136	128522	2580	6771	1167	29009	442184
C04	Transport-Dressing	107270	101107	34305	385977	104267	130058	63154	5117	43755	4236	979246	442206	22086	53924	4686	212272	1714419
C05	Heavy-HiTech Mach.	28798	109228	16645	37539	373665	60393	19560	4345	75127	1301	726600	76196	7985	243215	9718	339601	1403316
C06	Services	174950	134803	46815	239156	148592	1063860	71064	6703	77234	11491	1974668	1483061	884141	179163	1726	131397	4654155
C07	Chemicals	39429	39808	10767	32000	42073	23208	114031	870	25594	250	328031	58448	28159	4368	1180	150399	570584
C08	Office-machinery-PC	943	775	272	2344	2376	5020	634	3483	363	37	16248	4891	105	13375	728	22281	57627
C09	Motor-vehicles	8365	8556	1823	23883	11515	16235	3452	266	132545	2360	209001	213443	1883	82504	2545	129095	638470
C10	Renting-equipment	7468	13342	3265	17208	6799	22227	2881	262	3373	8315	85140	8415	85	256	7	1160	95061
ic	int. uses	806343	710244	204832	823245	750944	1552260	311833	22327	381523	29537	5593088	3300521	1324594	1208520	27618	1151860	12606201
m	imports	69689	63901	39047	163666	167194	110775	77748	24173	55192	1145	772530						
t_z	taxes on comm.	13657	13875	6326	33959	9032	76303	4960	463	3434	1209	163218						
w	wages	473387	298441	77883	377665	323837	1523444	100896	5525	130461	7716	3319255						
π	oper. surplus	372407	203343	112509	300852	144298	1328816	70869	4956	64184	54615	2656848						
t_g	taxes on prod.	-2005	7102	1588	15033	8010	62557	4277	184	3676	840	101262						
g	gross output	1733478	1296906	442184	1714419	1403316	4654155	570584	57627	638470	95061	12606201						

	Year 2007	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	ic	d_c	d_g	d_k	d_{vk}	d_x	g
C01	Agri-Food	387220	1697	2677	36848	71444	8239	8307	16696	1806	3216	538150	769644	3636	8141	10977	113437	1443986
C02	Furniture	5816	54426	5539	5728	15635	12472	28192	3352	2396	672	134227	56430	2612	29585	3562	40938	267354
C03	Energy	39627	5492	132833	32337	98080	40280	20672	32061	4902	5271	411555	133495	1139	8302	-163	38709	593038
C04	Transport-Trade	128056	28379	50269	565403	226182	163447	121715	105430	25683	13562	1428125	622198	36053	104402	-554	321433	2511658
C05	Services	146627	21800	54045	317577	1135812	177134	132311	119503	44437	138576	2287822	1979908	1668847	264178	-1027	174112	6373840
C06	Heavy Machinery	24112	16117	24580	66149	56635	689796	126948	29870	42452	1831	1078490	168209	3439	297745	1061	579118	2128061
C07	Construction	17209	3480	13937	25825	130384	24562	366569	12466	4613	5228	604274	48658	2263	844860	5926	32946	1538927
C08	Dressing-Chemicals	32389	13249	12732	42767	69428	80274	37388	212971	18129	1313	520639	151525	58592	6798	-995	284403	1020962
C09	Specialised Machin.	2344	868	1704	20515	33774	18543	5039	6741	61160	908	151595	31280	8180	69831	13868	138208	412962
C10	Finance	24244	3644	8882	62620	139579	25562	26723	13989	5047	234946	545237	266296	314	1985	11	54317	868159
ic	int. uses	807644	149152	307200	1175768	1976953	1240310	873863	553079	210625	405521	7700114	4227644	1785074	1635828	32668	1777621	17158948
m	imports	77294	25099	69603	248169	148899	267024	58017	146447	72201	30142	1142896						
t_z	taxes on comm.	12788	1222	8604	44083	112724	13199	16836	11721	3236	25141	249555						
w	wages	251047	50943	73118	523436	2166639	386247	324797	186937	83901	222319	4269384						
π	oper. surplus	301515	42848	130268	519997	1797804	241673	299955	129621	46066	149169	3658917						
t_g	taxes on prod.	-6302	-1910	4246	205	170820	-20391	-34542	-6842	-3066	35867	138082						
g	gross output	1443986	267354	593038	2511658	6373840	2128061	1538927	1020962	412962	868159	17158949						

Source: Own computation based on EUROSTAT SUIOT and National Accounts Databases

Table 21: Input-Output Tables by Clusters of the European Union (EU27), years 2000, 2007

industry × industry tables at basic prices built adopting the fixed product sales structure assumption (in millions of EUR)

	Year 2000	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	ic	d_c	d_g	d_k	d_{vk}	d_x	g
C01	Agri-Food	447581	10772	2449	12156	18443	38438	3066	64442	556	16928	1154	615984	865217	3643	10229	-420	63551	1558204
C02	Construction	23319	385591	12313	40320	15009	35726	8549	160739	753	9109	2025	693451	139430	5849	786127	8155	54988	1687999
C03	Energy	34219	29277	170042	41547	39062	61123	4904	68019	1605	9254	1718	460768	173025	1463	6462	1721	33548	676988
C04	Heavy Machinery	41187	144702	23135	654063	35217	64533	56529	82422	5376	7444	1707	1116314	375280	4004	355051	12191	331376	2194216
C05	Dressing-Chemicals	39233	62179	12402	84017	240908	38501	15494	40889	446	28369	2879	565315	214958	33408	8971	4076	170626	997354
C06	Transport-Media	129936	121374	32153	171174	106291	488422	41370	246502	6256	23901	3739	1371119	522085	28521	75812	8136	171550	2177224
C07	Specialised Machin.	4028	9534	2672	29077	5990	21766	88146	49822	587	11921	376	223920	57252	6617	123253	6686	130079	547806
C08	Services	170034	173994	53343	238974	120140	338734	69499	1428221	17800	76087	14793	2701619	1971609	1179769	242244	3803	150934	6249978
C09	Renting-equipment	6649	17913	3430	10638	4407	22315	2284	25058	8895	2113	834	104535	20283	91	380	44	1075	126407
C10	Health	4035	1209	314	1638	901	2339	531	9159	92	43786	411	64415	190002	535411	600	7	528	790964
C11	Personal-services	2352	1409	1205	3196	1607	6388	985	9383	273	2386	4121	33303	61603	2214	2909	18	2998	103046
ic	int. uses	902572	957954	313458	1286798	587973	1118285	291356	2184656	42638	231298	33756	7950744	4590744	1800989	1612039	44418	1111251	17110185
m	imports	50749	60772	120016	144758	83475	71660	89133	105624	1445	14089	1540	743262						
t_z	taxes on comm.	1816	16083	14365	12946	8233	43716	3771	118978	1494	21553	1070	244026						
w	wages	345826	459080	95795	569342	218118	636179	124133	2351082	21412	436027	26633	5283627						
π	oper. surplus	260115	189105	132190	171620	96622	296113	39020	1438092	58161	92954	39039	2813032						
t_g	taxes on prod.	-2875	5005	1164	8751	2933	11270	392	51546	1257	-4958	1008	75492						
g	gross output	1558204	1687999	676988	2194216	997354	2177224	547806	6249978	126407	790964	103046	17110185						

	Year 2007	C01	C02	C03	C04	C05	C06	C07	C08	C09	ic	d_c	d_g	d_k	d_{vk}	d_x	g
C01	Agri-Food	523916	12293	2535	48034	12821	6476	1262	86273	42183	735794	1073750	5807	11053	16722	82182	1925308
C02	Construction	27494	605901	18752	40637	66080	2349	8299	185054	32317	986884	89845	5291	1150545	13501	41077	2287143
C03	Energy	38749	27330	145253	38436	51779	5718	8871	91187	43856	451179	193316	2551	9819	170	7953	664989
C04	Transport-Trade	180133	187925	71688	820413	259517	27776	20492	330950	192720	2091615	902699	48372	128678	-2650	302072	3470787
C05	Heavy Machinery	38384	179590	28016	115281	1054950	6602	6082	111556	87185	1627645	364682	7259	500168	21409	553585	3074748
C06	Dressing	3929	3377	323	5373	11924	54585	888	10295	6997	97692	113512	393	1927	-1464	38094	250153
C07	Paper	15286	5406	898	8096	10639	1449	32925	41557	12986	129242	16515	344	481	-396	23200	169384
C08	Services	234729	234106	63345	521320	300566	28137	18084	2233480	318814	3952578	2823268	1549049	373666	2764	273478	8974802
C09	Pharma-Hi Tech	52678	72184	9121	64879	158453	13621	10465	114763	442389	938553	425210	882447	86389	13074	305462	2651136
ic	int. uses	1115298	1328113	339930	1662469	1926729	146713	107369	3205116	1179446	11011183	6002797	2501513	2262725	63130	1627102	23468451
m	imports	66447	60554	66534	242009	243025	24033	8517	157531	181090	1049742						
t_z	taxes on comm.	21346	22742	10601	66229	26618	2589	2129	166232	51326	369812						
w	wages	349840	467354	75356	750867	570070	50081	27456	2825522	851105	5967650						
π	oper. surplus	384581	438132	159437	726221	332774	29397	19855	2416853	394063	4901313						
t_g	taxes on prod.	-12205	-29752	13131	22992	-24468	-2660	4058	203548	-5893	168751						
g	gross output	1925308	2287143	664989	3470787	3074748	250153	169384	8974802	2651136	23468451						

Source: Own computation based on EUROSTAT SUIOT and National Accounts Databases

Table 22: Membership Matrix industry × country for IOT 2005

	at	be	cz	de	dk	ee	es	fi	fr	gr	hu	ie	it	lt	nl	pl	pt	ro	se	si	sk	uk
A01	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A02	2	10	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2
B05	1	11	3	1	1	1	10	3	1	3	1	1	1	3	3	1	1	1	9	1	1	1
CA10	3	2	4	2	13	9	11	2	3	4	3	3	10	9	13	3	13	3	1	3	3	3
CA11	4	12	4	2	3	3	3	10	4	9	3	11	2	10	4	8	14	3	10	3	3	3
CA12	10	13	13	10	14	10	12	11	10	10	7	12	11	11	14	9	15	11	11	11	12	9
CB13	11	3	14	11	15	11	13	4	5	4	4	13	12	12	15	3	16	4	3	12	3	10
CB14	4	4	3	3	4	4	4	4	2	5	2	2	3	4	5	10	3	5	12	4	4	2
DA15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DA16	5	1	1	4	5	12	5	12	1	1	5	1	4	13	1	1	1	6	13	13	1	1
DB17	5	5	4	2	6	3	6	5	6	3	6	4	5	5	2	2	4	3	4	2	5	4
DB18	5	14	15	2	6	3	6	5	6	3	6	5	5	5	2	2	4	3	4	2	6	4
DC19	5	15	5	2	1	3	7	6	1	6	6	5	6	6	2	2	5	1	1	2	5	1
DD20	2	4	2	2	4	2	2	2	2	1	2	2	7	2	5	2	2	2	2	2	7	2
DE21	5	1	2	2	7	13	5	2	3	3	3	1	6	2	1	2	2	7	4	5	2	1
DE22	6	6	6	5	8	5	5	7	7	7	6	6	6	3	6	4	6	7	5	5	8	5
DF23	4	2	4	2	3	3	6	8	5	3	3	3	2	14	3	5	7	3	4	6	4	3
DG24	5	2	4	2	7	3	5	2	3	3	3	6	4	6	3	2	7	3	5	1	3	4
DH25	5	2	5	2	9	4	8	5	6	3	3	4	8	7	3	2	7	8	3	7	5	4
DI26	4	4	3	3	4	4	4	5	2	5	2	2	3	4	5	3	3	5	6	4	9	2
DJ27	4	3	7	6	9	14	2	4	6	4	4	3	3	6	7	6	8	4	3	3	5	4
DJ28	4	3	7	6	9	4	2	9	6	5	4	2	3	1	7	6	8	4	3	8	5	4
DK29	5	3	7	6	9	3	2	9	6	4	4	4	3	6	7	6	8	4	3	8	5	4
DL30	5	3	8	7	9	15	5	13	6	7	5	4	6	3	1	4	6	7	14	1	6	4
DL31	5	3	8	6	9	3	2	9	6	5	4	4	3	3	6	6	8	4	5	4	5	4
DL32	5	6	5	8	9	16	2	7	6	7	5	4	9	3	6	4	6	7	15	1	5	4
DL33	5	3	9	9	9	3	5	5	6	3	3	4	4	6	8	7	7	4	6	1	5	6
DM34	5	3	5	6	9	17	8	9	6	4	4	4	2	7	9	5	8	4	3	7	10	4
DM35	5	5	5	9	9	6	2	9	6	4	6	3	2	7	9	6	8	4	7	9	5	4
DN36	2	3	2	2	10	2	2	5	6	1	6	14	7	2	2	2	2	2	3	2	5	4
DN37	4	16	7	2	9	4	2	4	11	4	4	3	3	1	7	6	8	12	3	3	5	11
E40	3	2	4	2	3	3	3	2	3	4	3	3	4	6	4	3	9	3	8	3	3	3
E41	7	7	6	2	11	3	5	5	8	8	3	3	4	6	1	3	1	3	6	3	3	4
F45	4	4	3	3	4	4	4	5	2	5	2	2	3	4	5	3	3	5	6	4	9	2
G50	5	6	5	1	10	6	8	8	6	3	6	5	2	3	9	5	6	1	4	7	10	7
G51	5	5	5	1	10	4	6	5	5	3	6	4	6	1	1	1	6	1	16	1	5	7
G52	7	6	10	4	11	7	5	5	9	3	6	7	5	3	6	1	6	6	17	6	6	8
H55	1	1	10	1	1	1	1	3	5	1	1	1	1	8	1	1	1	1	6	1	1	1
I60	5	5	11	1	10	6	6	8	5	3	6	4	2	7	9	5	7	3	4	9	4	7
I61	5	5	11	1	10	6	6	8	5	8	3	8	2	1	9	11	6	4	4	9	4	7
I62	1	5	11	9	10	6	6	8	5	8	5	8	2	7	9	1	6	9	4	6	4	8
I63	1	5	11	1	10	6	6	8	5	8	6	8	2	7	9	1	6	10	4	9	4	7
I64	6	6	12	8	11	5	5	7	9	3	5	5	9	3	6	4	6	7	5	6	6	8
J65	6	8	10	4	12	8	5	5	9	3	5	6	9	8	5	4	10	5	6	6	6	8
J66	6	8	12	4	12	8	5	7	7	7	5	6	9	8	10	4	6	7	6	6	6	8
J67	6	8	12	4	12	8	5	7	7	7	5	6	9	8	10	4	17	5	3	6	6	8
K70	7	8	10	4	4	7	5	5	9	5	6	7	5	3	5	3	10	7	6	6	6	8
K71	6	6	3	3	11	6	5	5	7	7	2	2	2	3	9	5	6	5	5	6	6	2
K72	8	6	10	8	11	5	5	7	9	7	5	5	9	3	6	4	10	5	5	6	6	8
K73	6	6	5	8	11	5	5	7	6	7	3	9	9	15	6	2	6	7	5	1	1	6
K74	6	6	3	4	11	5	5	7	7	7	6	5	5	3	6	4	6	5	18	6	6	8
L75	7	17	3	8	11	3	14	7	12	11	5	7	9	8	11	4	18	13	6	6	1	8
M80	7	8	6	8	11	3	5	7	5	2	5	5	5	6	6	4	6	7	6	1	3	8
N85	9	9	16	4	11	3	5	3	5	3	3	9	4	6	1	7	11	14	6	1	3	6
O90	7	7	3	4	11	3	9	5	4	3	3	10	4	8	11	3	12	7	6	10	1	6
O91	6	6	6	4	11	18	5	7	13	12	6	5	5	3	6	2	19	15	5	6	11	5
O92	6	6	6	5	11	5	5	7	7	7	5	5	5	3	12	4	6	16	5	6	11	5
O93	7	1	6	4	11	3	5	3	7	1	6	5	5	3	1	4	6	17	6	1	3	8
P95	12	18	17	12	16	19	15	14	14	13	8	15	13	16	16	12	20	18	19	14	13	12

References: Each cell contains the cluster number of each industry (row) for a given country (column).

Source: Own computation based on EUROSTAT SUIOT and National Accounts Databases

Table 23: Membership Matrix country × commodity for intra-EU27 external trade 2007

	A01	A02	B05	CA10	CB13	CB14	DA15	DB17	DB18	DC19	DD20	DE21	DE22	DF23	DG24	DH25	DI26	DJ27	DJ28	DK29	DL30	DL31	DL32	DL33	DM34	DM35	DN36	K72	O92
at	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
be	2	2	1	2	1	2	2	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	1	2
bg	1	3	2	3	2	3	1	3	3	3	1	3	1	3	1	1	3	2	3	1	1	1	1	1	3	1	2	1	3
cy	1	3	3	3	2	3	1	4	3	3	1	2	3	3	1	3	3	2	3	3	1	2	3	3	2	3	2	1	3
cz	1	1	4	1	2	1	1	1	1	1	1	3	1	1	1	1	1	2	1	1	2	1	1	1	1	1	1	1	1
de	3	1	1	4	1	1	1	1	1	1	1	1	1	2	2	1	1	1	1	1	2	1	4	1	1	3	1	1	1
dk	3	1	1	1	2	1	3	5	4	4	3	1	4	4	3	4	4	3	4	4	3	3	3	4	2	2	3	2	4
ee	3	4	4	3	1	1	3	5	4	4	3	1	4	4	1	4	4	3	4	4	3	3	3	4	4	4	3	3	4
es	2	5	2	3	3	3	4	4	2	2	2	2	3	3	4	3	3	4	3	3	4	4	4	2	5	2	2	2	1
fi	3	4	4	1	3	1	3	5	4	4	3	1	4	4	3	4	4	3	4	4	3	3	3	4	2	4	3	1	4
fr	2	2	3	2	1	3	2	2	2	2	2	2	2	2	4	3	3	1	3	3	2	4	4	2	5	3	2	2	1
gr	1	3	2	4	2	3	1	3	3	3	1	2	3	3	1	3	3	2	3	3	1	2	3	3	3	2	2	1	3
hu	1	1	4	1	3	1	1	1	1	1	1	3	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1
ie	4	4	3	5	3	2	5	2	5	5	3	4	3	2	2	2	5	4	2	2	5	2	2	5	2	3	2	2	1
it	2	1	2	3	3	3	1	3	2	3	1	2	2	3	4	3	3	1	3	3	1	4	4	2	5	2	2	1	1
lt	3	4	4	3	2	1	3	5	4	4	3	1	4	4	1	4	4	5	4	4	3	3	3	4	4	4	3	3	4
lu	2	2	3	2	1	1	2	1	2	2	1	2	2	2	2	2	2	1	1	2	4	2	4	2	2	2	2	1	2
lv	3	4	4	3	2	1	3	5	4	4	3	1	4	4	1	4	4	5	4	4	3	3	3	4	4	4	3	3	4
mt	2	1	3	4	1	3	5	1	5	3	1	4	3	3	4	3	3	1	2	3	1	2	4	2	2	3	2	2	1
nl	3	2	1	4	1	1	2	2	1	2	2	1	2	2	2	2	2	1	2	2	2	2	2	5	2	2	1	1	2
pl	3	1	4	1	2	1	1	1	1	1	1	1	1	4	1	1	1	2	1	1	2	1	4	1	1	2	1	1	1
pt	2	5	2	3	3	3	4	4	2	2	2	2	3	3	4	3	3	4	3	3	4	4	4	2	5	2	2	2	4
ro	1	1	4	1	2	3	1	3	5	3	1	3	1	3	1	1	1	2	3	3	1	1	1	1	3	1	2	1	4
se	3	4	4	1	3	1	3	5	4	4	3	1	4	4	3	4	4	3	4	4	3	3	3	4	2	4	3	2	4
si	1	1	2	3	1	1	1	1	1	1	1	3	1	3	1	1	1	1	1	1	1	1	1	1	5	1	1	2	1
sk	1	1	4	1	1	1	1	1	1	1	1	3	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1
uk	4	4	3	5	2	2	5	2	5	5	3	4	3	2	2	2	5	4	2	2	5	2	2	5	2	3	2	2	1

References: Each cell represents the cluster number of each country (in a row) for a given commodity (in a column).

Source: Own computation based on EUROSTAT SUIOT and National Accounts Databases

References

EC (2007). Clusters in the EU — 10 new Member States. Technical Report 1, European Commission, Luxembourg. Europe INNOVA Paper.

——— (2008). The Concept of Clusters and Cluster Policies and their Role for Competitiveness and Innovation: Main Statistical Results and Lessons Learned. Technical report, European Commission, Luxembourg. Commission Staff Working Document SEC (2008) 2637.

EUROSTAT (2006). Statistics on the trading of goods — User guide. Technical report, Eurostat, Luxembourg.

——— (2011). Creating consolidated and aggregated EU27 Supply, Use and Input-Output Tables, adding environmental extensions (air emissions), and conducting Leontief-type modelling to approximate carbon and other 'footprints' of EU27 consumption for 2000 to 2006. Technical report, Eurostat, Luxembourg.

Hoen, A.R. (2002). Identifying Linkages with a Cluster-based Methodology. *Economic Systems Research*, 14(2).

Leicht, E. & Newman, M. (2008). Community Structure in Directed Networks. *Phys. Rev. Lett.*, 100.

Leontief, W.W. (1963 [1986]). The structure of development. In *Input-Output Economics — Second Edition*, pp. 162–187. Oxford University Press.

McNerney, J. (2009). Network Properties of Economic Input-Output Networks. Technical report, International Institute for Applied Systems Analysis, Laxenburg, Austria. Interim Report.

Murtagh, F. & Contreras, P. (2011). Methods of Hierarchical Clustering. *CoRR*, abs/1105.0121.
URL: *http://arxiv.org/abs/1105.0121*

Newman, M. (2006). Modularity and Community Structure in Networks. *Proc. Natl. Acad. Sci. USA*, 103:8577–8582.

Pasinetti, L.L. (1973). The Notion of Vertical Integration in Economic Analysis. *Metroeconomica*, 25:1–29.

Piccardi, C. (2011). Finding Communities by Lumped Markov Chains. *PLoS ONE*, 6(11):1–13.

Porter, M. (2003). The Economic Performance of Regions. *Regional Studies*, 37(6&7):549–578.

Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., & Parisi, D. (2004). Defining and identifying communities in networks. *PNAS*, 101:2658–2663.

Rosvall, M., Axelsson, D., & Bergstrom, C. (2009). The map equation. *Eur. Phys. J. Special Topics*, 178:13–23.

Santos, C., Almeida, A., & Teixeira, A.A.C. (2008). Searching for clusters in tourism. A quantitative methodological proposal. Technical Report 293, Universidade do Porto, Faculdade de Economia (FEP).

Sraffa, P. (1960). *Production of Commodities by Means of Commodities.* Cambridge University Press, Cambridge.

ten Raa, T. & Rueda-Cantuche, J.M. (2009). The Choice of Model in the Construction of Industry Coefficients Matrices. *Economic Systems Research*, 21(4):363–376.