

The following are my first experiments in reimplementing and testing the FORTRAN code for regular equivalence (Smith and White 1992; see <http://eclectic.ss.uci.edu/~drwhite/links2pdf.htm>) analysis for use with PCs. This approach will be used to analyze the input-output matrices of six European countries as coded by OECD. Because regular equivalence identifies structural positions occupied by sectors within a national economy, it is oriented towards the structure of concomitant flows to and from structural positions that may be composed of heterogeneous and complementary elements. In the case of the world economy positions identified by this method by Smith and White 1992 the positional structure proved to be a 1-dimensional ordering from raw material at one pole to finished goods and machines for making machines at the other. Sectors in the national economies of developed countries were expected to have 1-factor dominant with a secondary component for interdependence of circular flows among differentiated sectors.

I took Denmark as a start. First I combined the years, 77-80-85 and 1990, but it became clear that this economy was changing rapidly and each year needed to be treated separately. The data format is simply to have two header lines as shown and then a series of single headers followed by comma-separated entries for each row of the matrix, in this case 33 lines each with 33 entries.

```

4 matrices
33 nodes
1977 title
3201,0,20861,3,269,33,1,14,0,1,1,0,0,0,1,0,
0,0,6,0,0,0,0,0,0,352,11,134,3,0,2,4,2
23,19,25,1,0,0,6,1,341,0,169,3,1,0,1,0,0,0,
1,0,0,0,0,0,0,65,0,4,0,0,0,0,0,0
Etc

```

This file goes on to include all four matrices. Strip the headers and substitute blanks for commas to import to UCInet as a raw data file. An SVD of the raw network data in UCInet shows most nodes pile in the center, which is also what Pajek shows (output from UCInet to Pajek). The outliers don't tell you much about the economy because similar categories are not grouped. That is the purpose of structural equivalence and regular equivalence analysis, the latter being the more general method.

The two original regular equivalence and distance algorithms, REGGE.FOR and REGDI.FOR, were rewritten and debugged for the first time for the PC. The original CRAY routines were not changed, but a normalization procedure was adopted for both: 15 iterations of taking row and column sums and dividing each entry by the square root of its row and its column sum. For distance the zero in the diagonal does not affect normalization, but regular equivalence diagonals are set temporarily to zero during normalization and then reset to the highest value in the matrix when done. All row and column totals will equal .5 plus the highest value in the equivalence matrix. Hence marginals do not enter into the pattern of results. In SVD analysis, the first factor will always have constant loadings, so it is only factors 2 and 3 (and possibly others) that account for variance.

The distance and similarity methods give similar results, and output files are named REGDNORMtitle and REGSNORMtitle, where 'title' is the first four digits of the last title line in the input data. The Danish results for 1977 show a nearly 1-factor structure corresponding to agriculture vs. industry. Sectors migrate along the principal component over time.

The new programs are now available at the url <http://eclectic.ss.uci.edu/~drwhite/REGGE/> (THIS LINK IS CLICKABLE)

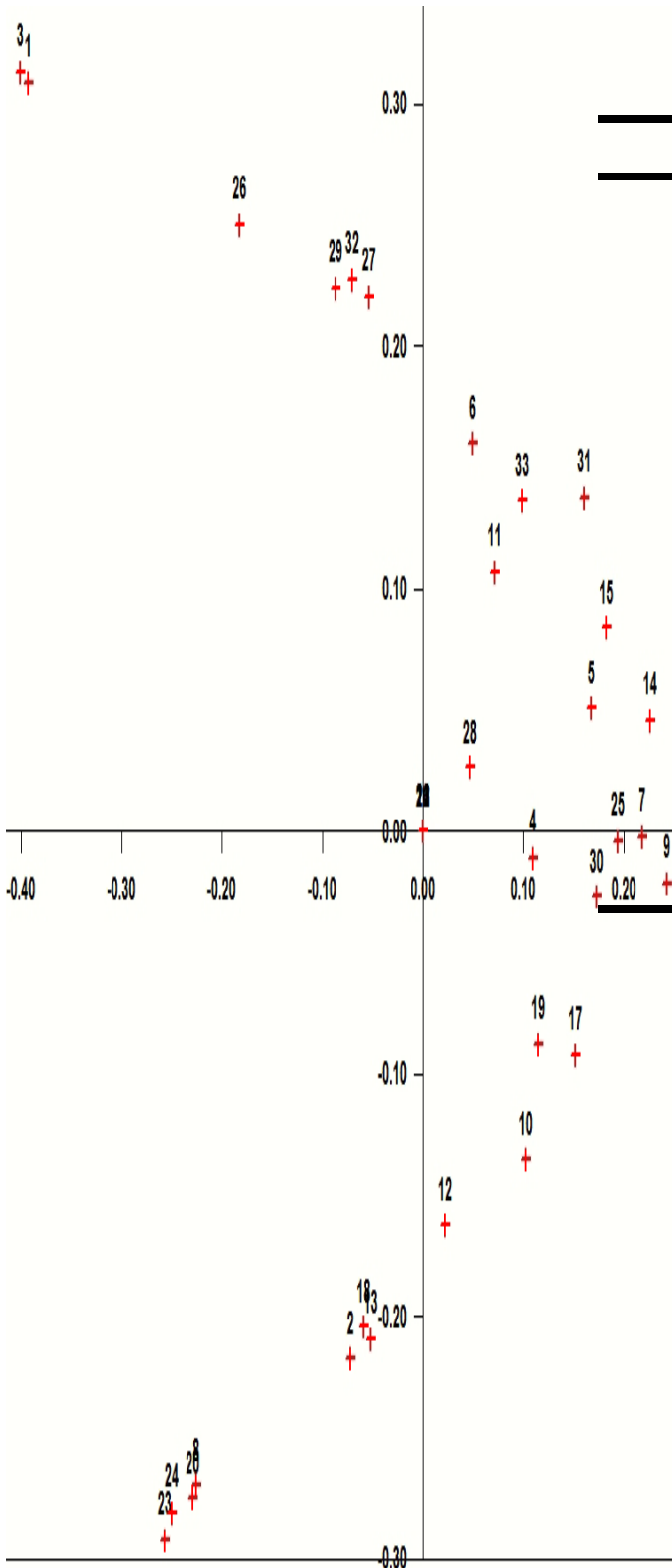
For the input-output matrices in European countries, the relevant background is found in The OECD Input-Output Database available at url

<http://www.oecd.org/dataoecd/48/43/2673344.pdf>

The actual data, however, is in a database format that is somewhat outmoded and difficult to parse. Colleagues of Dirk Helbing, University of Technology, Desden, reformatted these data in excel spreadsheet format that are not publicly available. I can provide them under a non-sharable basis to students in my seminar. The easiest way to prepare databases for analysis with RegSim is to highlight the matrix for row/col numbers 1-33 and past into a word processor, e.g., word, convert table to text, global replace all blanks to single blanks, and global replace the blanks by commas, add the appropriate headers, and save to ASCII format.

The following pages are my graphics, notes and results for Denmark, 1977.

Denmark77: Regular Distance analysis, 1977
 D.R. White @2004 Regdis software, SIMILARITY
 data courtesy Dirk Helbing, UT Dresden



Sector – SVD analysis components			#2 vert	#3 horiz
33	33	Community, social & personal services	-0.236	-0.179
6	6	Paper, paper products & printing	-0.231	0.161
31	31	Finance & insurance	-0.221	0.025
27	27	Wholesale & retail trade	-0.214	0.061
32	32	Real estate & business services	-0.193	0.270
15	15	Non-electrical machinery	-0.186	-0.247
14	14	Metal products	-0.181	-0.228
26	26	Construction	-0.175	0.267
29	29	Transport & storage	-0.170	0.300
11	11	Non-metallic mineral products	-0.144	0.071
5	5	Wood products & furniture	-0.119	-0.218
4	4	Textiles, apparel & leather	-0.101	-0.102
9	9	Petroleum & coal products	-0.067	-0.099
2	2	Mining & quarrying	-0.058	-0.065
3	3	Food, beverages & tobacco	-0.052	0.277
1	1	Agriculture, forestry & fishing	-0.050	0.276
28	28	Restaurants & hotels	-0.041	-0.074
30	30	Communication	-0.023	-0.244
25	25	Electricity, gas & water	0.038	-0.257
7	7	Industrial chemicals	0.064	-0.307
17	17	Electrical machinery & components	0.180	-0.183
13	13	Non-ferrous metals	0.182	0.144
19	19	Shipbuilding & repairing	0.213	-0.149
24	24	Other manufacturing	0.246	0.102
23	23	Professional goods	0.249	0.145
12	12	Iron & steel	0.254	0.000
18	18	Radio, TV & communication equipment	0.256	0.048
10	10	Rubber & plastic products	0.259	-0.017
8	8	Drugs & medicines	0.261	0.111
20	20	Other transport	0.261	0.112

The input-output economy has
 linear order from agricultur
 (top) & industry (bottom)

a 1-factor outcome evident;
 ratio of factors: 1 to ~2.1;
 Horiz X axis shrunk accordingly

The three 0,0 industries lack data

22	22	Aircraft	0
21	21	Motorvehicles	0
16	16	Office&computingmachinery	0

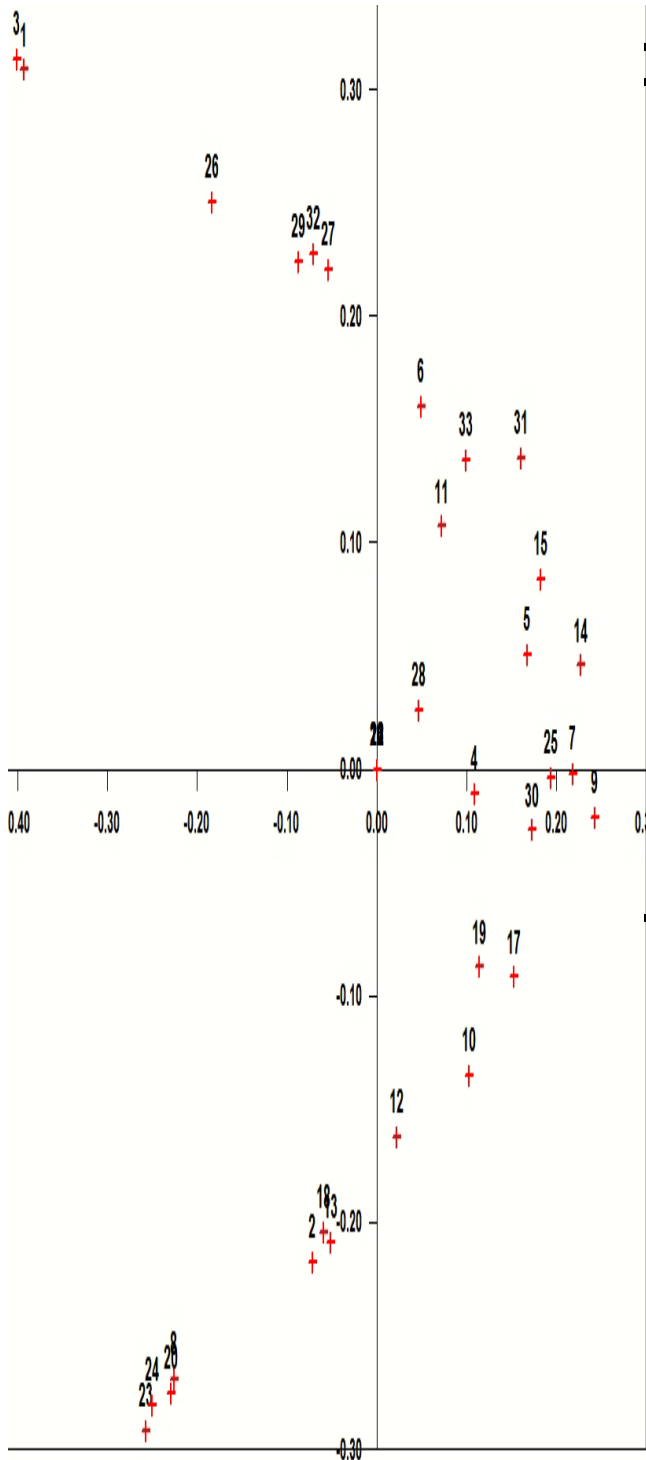
SVD factors RegNorm77

1:	0.583	21.3	21.3	1.984
2:	0.294	10.7	32.0	2.133
3:	0.138	5.0	37.0	1.526

Horiz = svd pc #3 Vert = svd pc #2
 principal component #1 57% variance; w/ #2 86%

1:	17.161	57.2	57.2	1.978
2:	8.677	28.9	86.1	3.765

D.R. White @2004 RegSim software, SIMILARITY
data courtesy Dirk Helbing, UT Dresden



orig	new	Sector – SVD components	#2	#3
1	1	1 Agriculture, forestry & fishing	0.309	-0.393
2	2	2 Mining & quarrying	-0.217	-0.072
3	3	3 Food, beverages & tobacco	0.313	-0.401
4	4	4 Textiles, apparel & leather	-0.011	0.109
5	5	5 Wood products & furniture	0.051	0.168
6	6	6 Paper, paper products & printing	0.160	0.049
7	7	7 Industrial chemicals	-0.003	0.218
8	8	8 Drugs & medicines	-0.270	-0.226
9	9	9 Petroleum & coal products	-0.021	0.243
10	10	10 Rubber & plastic products	-0.135	0.103
11	11	11 Non-metallic mineral products	0.107	0.072
12	12	12 Iron & steel	-0.162	0.022
13	13	13 Non-ferrous metals	-0.209	-0.052
14	14	14 Metal products	0.046	0.227
15	15	15 Non-electrical machinery	0.084	0.183
17	17	16 Electrical apparatus, nec	-0.092	0.152
18	18	17 Radio, TV & communication equipment	-0.204	-0.059
19	19	18 Shipbuilding & repairing	-0.088	0.115
20	20	19 Other transport	-0.275	-0.229
23	23	20 Professional goods	-0.292	-0.257
24	24	21 Other manufacturing	-0.281	-0.250
25	25	22 Electricity, gas & water	-0.004	0.194
26	26	23 Construction	0.250	-0.183
27	27	24 Wholesale & retail trade	0.221	-0.054
28	28	25 Restaurants & hotels	0.026	0.047
29	29	26 Transport & storage	0.224	-0.087
30	30	37 Communication	-0.027	0.173
31	31	28 Finance & insurance	0.138	0.161
32	32	29 Real estate & business services	0.227	-0.071
33	33	30 Community, social & personal services	0.136	0.099

The input-output economy has a linear order from agriculture (top) & industry (bottom)

But many industries have migrated downwards!

a 1-factor outcome evident;
ratio of factors is 1 to ~4.8;
Horiz X axis shrunk accordingly

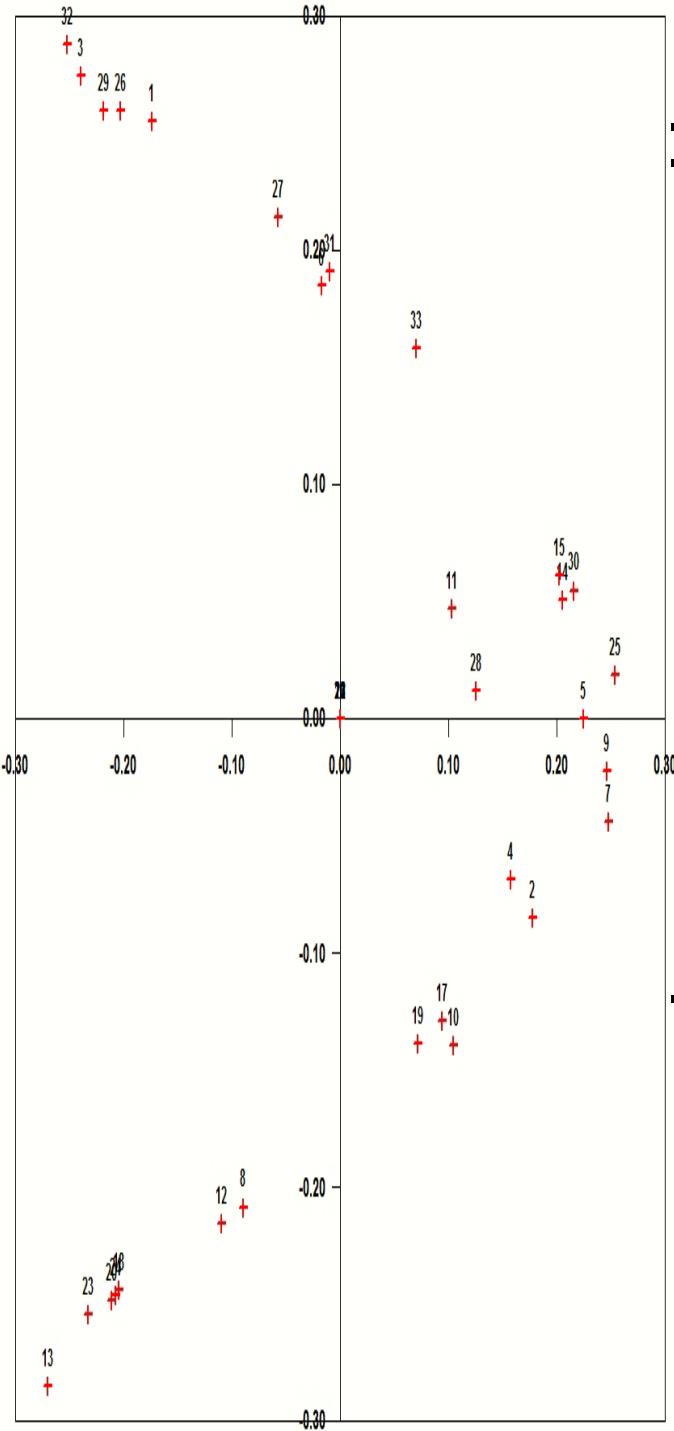
The three 0,0 industries lack data

22	22	22 Aircraft	0	0
21	21	21 Motorvehicles	0	0
16	16	16 Office&computingmachinery	0	0

→ SVD components (33)	Ratios:			
1: 0.583	21.3	21.3	1.984	
2: 0.294	10.7	32.0	2.133	
3: 0.138	5.0	37.0	1.526	

ert Y= svd pc #2 Horiz X= svd pc #3
principal component #1 46% variance; w/ #2 10% 30
1: 13.930 46.4 46.4 4.800
2: 2.902 9.7 56.1 2.144

Denmark90: Regular Similarities analysis, 1990
D.R. White @2004 RegSim software, SIMILARITY.
Data courtesy Dirk Helbing, UT Dresden



orig	new	Sector – principal components	#2	#3
1	1	1 Agriculture, forestry & fishing	0.255	-0.183
2	2	2 Mining & quarrying	-0.078	0.176
3	3	3 Food, beverages & tobacco	0.274	-0.246
4	4	4 Textiles, apparel & leather	-0.051	0.162
5	5	5 Wood products & furniture	-0.001	0.227
6	6	6 Paper, paper products & printing	0.186	-0.021
7	7	7 Industrial chemicals	-0.044	0.243
8	8	8 Drugs & medicines	-0.212	-0.091
9	9	9 Petroleum & coal products	-0.018	0.237
10	10	10 Rubber & plastic products	-0.142	0.098
11	11	11 Non-metallic mineral products	0.045	0.110
12	12	12 Iron & steel	-0.217	-0.114
13	13	13 Non-ferrous metals	-0.289	-0.275
14	14	14 Metal products	0.052	0.209
15	15	15 Non-electrical machinery	0.064	0.199
17	17	16 Electrical apparatus, nec	-0.130	0.098
18	18	17 Radio, TV & communication equipment	-0.243	-0.192
19	19	18 Shipbuilding & repairing	-0.141	0.071
20	20	19 Other transport	-0.252	-0.215
23	23	20 Professional goods	-0.256	-0.226
24	24	21 Other manufacturing	-0.244	-0.198
25	25	22 Electricity, gas & water	0.016	0.250
26	26	23 Construction	0.260	-0.208
27	27	24 Wholesale & retail trade	0.213	-0.055
28	28	25 Restaurants & hotels	0.011	0.128
29	29	26 Transport & storage	0.258	-0.220
30	30	37 Communication	0.052	0.223
31	31	28 Finance & insurance	0.189	-0.012
32	32	29 Real estate & business services	0.285	-0.250
33	33	30 Community, social & personal services	0.158	0.074

The input-output economy has a linear order from agriculture (top) & industry (bottom)

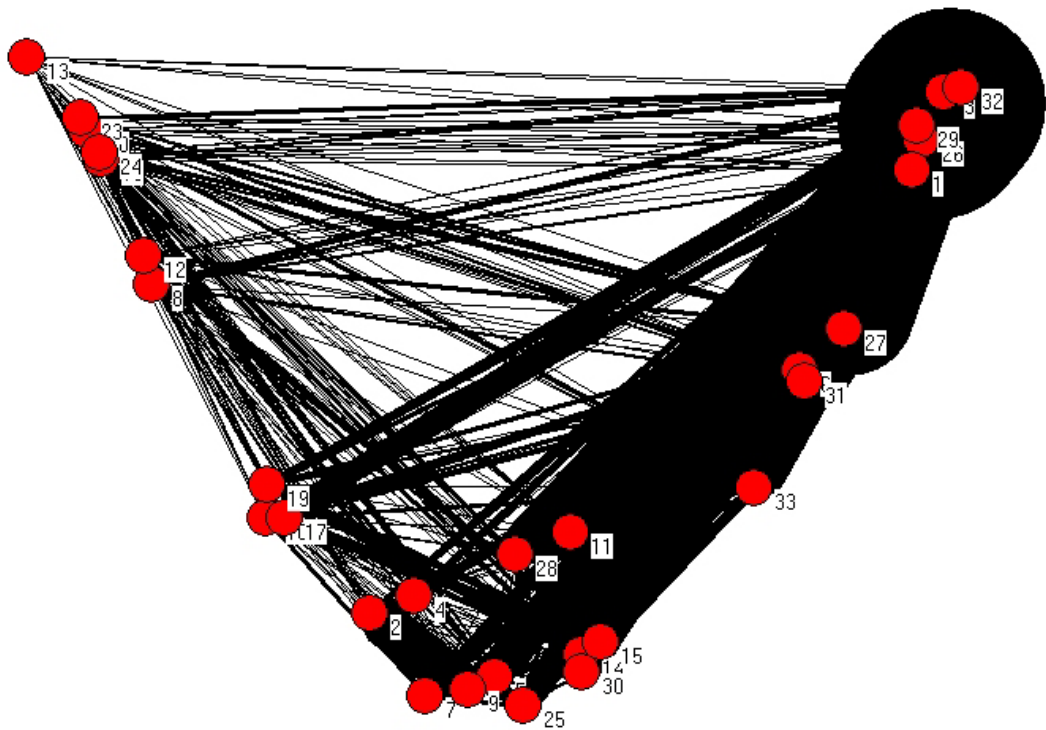
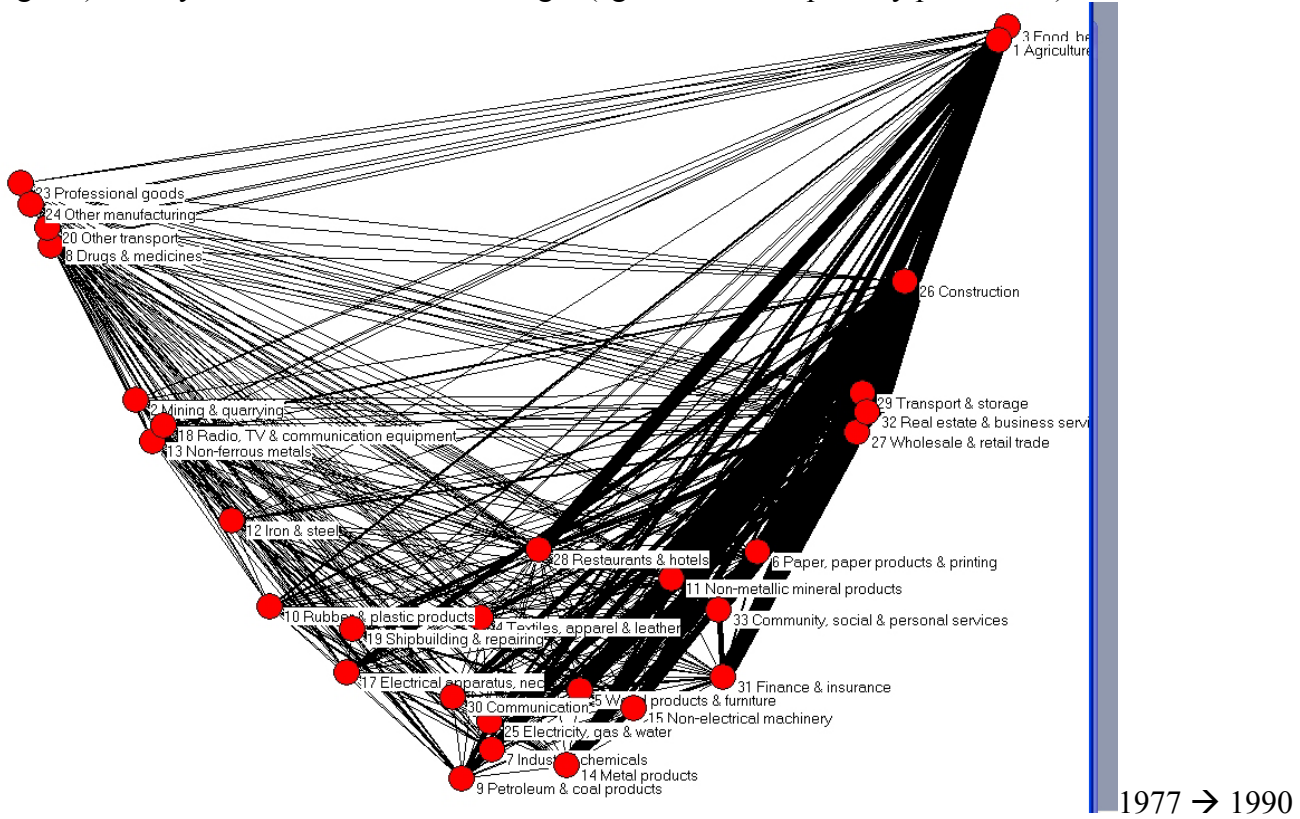
The three 0,0 industries lack data

a 1-factor outcome evident;
ratio of factors is 1 to 2.84;
vertical axis stretched accordingly

22	22	22	Aircraft	0
21	21	21	Motorvehicles	0
16	16	16	Office&computingmachinery	0

Vert Y= svd pc #2 Horiz X= svd pc #3
1: 0.546 35.2 35.2 2.148
2: 0.254 16.4 51.7 2.842
3: 0.090 5.8 57.4 1.925
→ principal components on 30
1: 9.663 32.2 32.2 2.545
2: 3.797 12.7 44.9 1.665

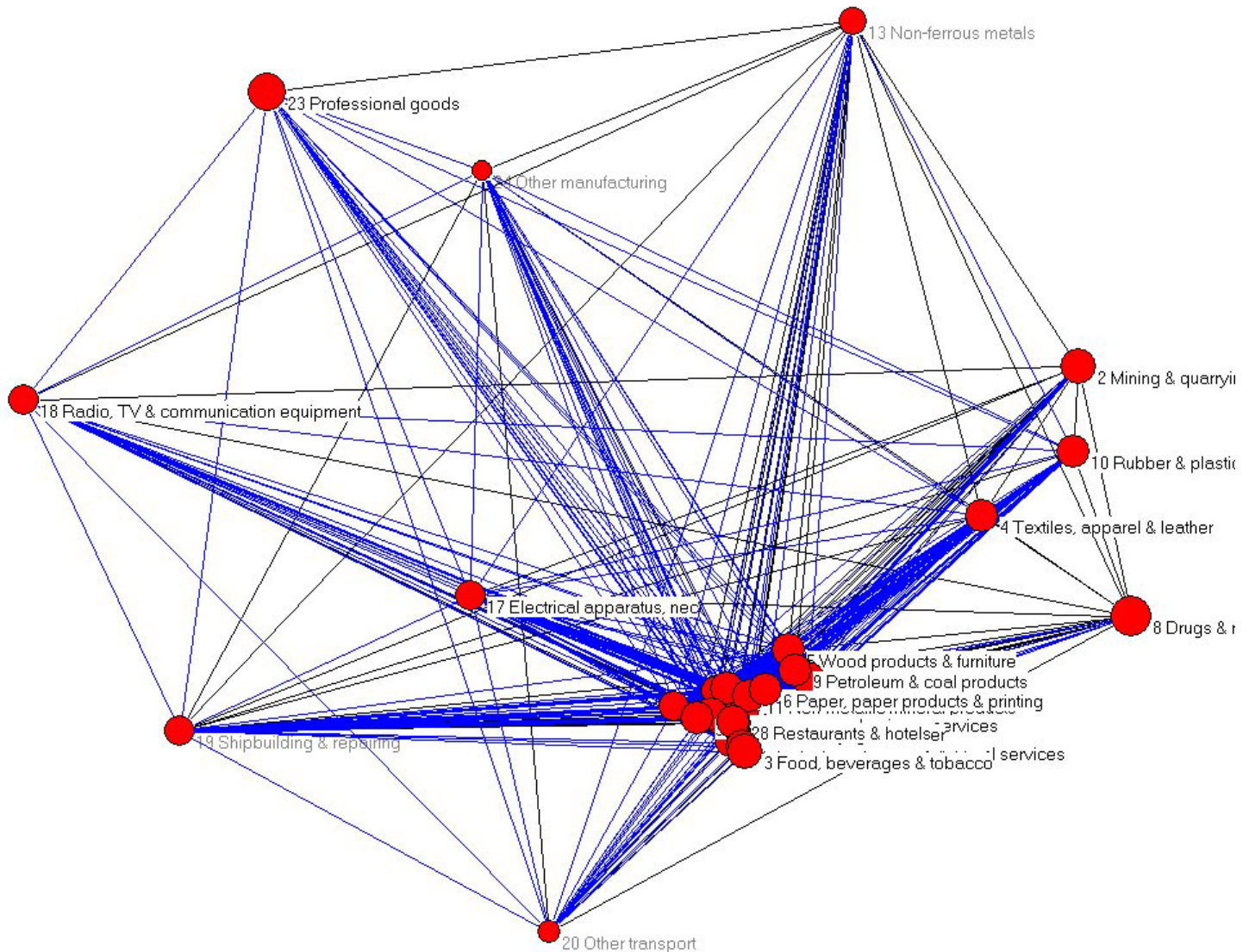
These are scalings of the raw network data for 1977 and 1990 using SVD coordinates 2 and 3 (rotated 90 degrees). Heavy flows are evident on the right (agricultural and primary production) side.



That regular equivalence has done exactly what it is supposed to do is evident: similarly-sized flows have been matched across pairs of nodes and positions that are similarly situated in the inter-positional flow pattern.

The semicircular arrangement of the regular equivalence nodes on SVD components 2 (horiz) and 3 (vertical) is also as it ought to be: while there is one independent component of variance that orders nodes according to an industrial dominance chain, with flow magnitudes dominant in this case from right to left and both growing in magnitude and nucleating on the right side as well, there are also residual similarities created by the smaller-magnitude cycling of flows that connects the left and right poles of the primary scale. This is more or less what Smith and White (1992) identified as the pattern in international trade.

Below is a pajek SVD of the raw network data for 1977: it differentiates shipbuilding but not agriculture, which is in the core, and it does not get at equivalence of position. Note how heavy flows fail to differentiate (as with equivalence analysis) but pull to the center.



1977 SVD

Agriculture, forestry & fishing	1	-0.183	0.309	-0.393
Mining & quarrying	2	-0.183	-0.217	-0.072
Food, beverages & tobacco	3	-0.183	0.313	-0.401
Textiles, apparel & leather	4	-0.183	-0.011	0.109
Wood products & furniture	5	-0.183	0.051	0.168
Paper, paper products & printing	6	-0.183	0.160	0.049
Industrial chemicals	7	-0.183	-0.003	0.218
Drugs & medicines	8	-0.183	-0.270	-0.226
Petroleum & coal products	9	-0.183	-0.021	0.243
Rubber & plastic products	10	-0.183	-0.135	0.103
Non-metallic mineral products	11	-0.183	0.107	0.072
Iron & steel	12	-0.183	-0.162	0.022
Non-ferrous metals	13	-0.183	-0.209	-0.052
Metal products	14	-0.182	0.046	0.227
Non-electrical machinery	15	-0.183	0.084	0.183
Electrical apparatus, nec	17	-0.183	-0.092	0.152
Radio, TV & communication equipment	18	-0.183	-0.204	-0.059
Shipbuilding & repairing	19	-0.183	-0.088	0.115
Other transport	20	-0.183	-0.275	-0.229
Professional goods	23	-0.183	-0.292	-0.257
Other manufacturing	24	-0.183	-0.281	-0.250
Electricity, gas & water	25	-0.183	-0.004	0.194
Construction	26	-0.183	0.250	-0.183
Wholesale & retail trade	27	-0.183	0.221	-0.054
Restaurants & hotels	28	-0.183	0.026	0.047
Transport & storage	29	-0.183	0.224	-0.087
Communication	30	-0.183	-0.027	0.173
Finance & insurance	31	-0.183	0.138	0.161
Real estate & business services	32	-0.183	0.227	-0.071
Community, social & personal services	33	-0.183	0.136	0.099
Office & computing machinery	16	-0.000	-0.000	-0.000
Motor vehicles	21	-0.000	-0.000	-0.000
Aircraft	22	0.000	0.000	0.000

1990 SVD

1	-0.183	0.255	-0.183
2	-0.183	-0.078	0.176
3	-0.183	0.274	-0.246
4	-0.183	-0.051	0.162
5	-0.183	-0.001	0.227
6	-0.182	0.186	-0.021
7	-0.183	-0.044	0.243
8	-0.183	-0.212	-0.091
9	-0.183	-0.018	0.237
10	-0.183	-0.142	0.098
11	-0.183	0.045	0.110
12	-0.182	-0.217	-0.114
13	-0.183	-0.289	-0.275
14	-0.183	0.052	0.209
15	-0.183	0.064	0.199
16	-0.000	-0.000	-0.000
17	-0.183	-0.130	0.098
18	-0.183	-0.243	-0.192
19	-0.183	-0.141	0.071
20	-0.183	-0.252	-0.215
21	0.000	0.000	-0.000
22	0.000	0.000	0.000
23	-0.183	-0.256	-0.226
24	-0.183	-0.244	-0.198
25	-0.183	0.016	0.250
26	-0.182	0.260	-0.208
27	-0.183	0.213	-0.055
28	-0.183	0.011	0.128
29	-0.183	0.258	-0.220
30	-0.183	0.052	0.223
31	-0.183	0.189	-0.012
32	-0.183	0.285	-0.250
33	-0.183	0.158	0.074