IDEA paper

IDEA PAPER SERIES

IDEA
14
1998

UNDERSTANDING R&D INTENSITY INDICATORS

- Effects of differences in industrial structure and country size

Tore Sandven
and Keith Smith

Tore Sandven is a researcher at the STEP group. Keith Smith is research director at the STEP group. STEP group, Storgaten 1, N-0155 Oslo. Tel: +47 22 47 73 10, Fax: +47 22 42 95 33, email: tore.sandven@step.no, keith.smith@step.no.

STEP group

Studies in technology, innovation and economic policy
Studier i teknologi, innovasjon og økonomisk politikk

IDEA

This report is a background paper for Project 1.2, ‘Science and Technology Indicators: a Guide for Policymakers’, of the IDEA (Indicators and Data for European Analysis) Project. IDEA is Project No. PL951005 under the Targeted Socio-Economic Research Programme, Area 1 (Evaluation of Science and Technology Policy Options in Europe), Theme 1.3: *Methodologies, Tools and Approaches Relevant for the Preparation, Monitoring and Evaluation of Science and Technology Policies.*

An overview of the project as a whole, covering objectives, work programme, and results, including downloadable reports, can be found on the IDEA Web-site:

http://www.sol.no/step/IDEA/

ABSTRACT

This report explores a range of problems associated with the interpretation of research and development (R&D) intensities across countries. Our argument is that this commonly-used indicator contains a number of subtle problems which must be taken into account when using it to compare innovation performance.

Expenditure on R&D as a proportion to some measure of total economic activity is a frequently used measure for comparing the extent of innovation activities across different kinds of units, be they firms, industries, national economies, etc. This ratio is often referred to as *R&D intensity*. At the level of national economies, total R&D expenditures to GDP (gross domestic product) in a given year is frequently used to compare the innovative efforts of different countries. Often countries are simply compared on the basis of this ratio. For instance, R&D expenditures of a given country may be judged insufficient by a reference to data showing that the R&D intensity of the country in question is substantially smaller than the R&D intensity of certain other countries.

There are two basic problems in using this indicator to compare innovation performance. The first is that, as an empirical regularity, large countries have higher R&D-intensities than small countries. The second problem is that R&D-intensity is obviously affected by the industrial structure. For example, if one country has a large proportion of its output in R&D-intensive industries, it will have a higher overall R&D intensity, even if R&D/output ratios are equal in every industry.

What is the relative importance of country size and industrial structure in determining the value of the R&D inensity indicator? We argue that size, in itself, offers no real explanation of inter-country R&D intensity differences. However, when we decompose R&D intensity in manufacturing into, on the one hand, a component expressing the industrial structure of the country in question, and, on the other hand, a component expressing how the country in general compares with the other countries in terms of R&D intensity inside each industry, we find a clear and

strong positive association between economy size and the *structure component*. The larger the size of the economy, as measured by GDP, the higher the R&D intensity in manufacturing we would predict from knowledge only of the industrial structure, or, in other words, the more the industrial structure is favourable to a high R&D intensity in manufacturing.

This report takes up a number of methodological issues in the use and interpretation of R&D data. It discusses and criticises the so-called STIBERD indicator developed by OECD to take account of industrial structure differences, and offers guidelines for new approaches in understanding comparative R&D intensities.

TABLE OF CONTENTS

ABSTRACT II
TABLE OF CONTENTS IV
TABLES V
FIGURES VI
INTRODUCTION 1
The data *2*
DIFFERENCES IN R&D INTENSITY ACROSS COUNTRIES 3
Evolution over time in differences in R&D intensity across countries *8*
Differences in R&D intensity across industries *15*
R&D intensity in manufacturing and industrial structure *25*
THE QUESTION OF PERFORMANCE 51
The OECD STIBERD indicator *55*
The problem of statistical interaction *64*
An alternative indicator: weighted mean of ranks *73*
CONCLUSION 90

TABLES

Table 1: R&D intensities in the manufacturing sector, per cent, 1991. ... 3
Table 2: R&D intensities in the manufacturing sector, per cent, differences from average, 1991 ... 4
Table 3: Classification of countries into large and small economies ... 4
Table 4: Mean and median R&D intensity in manufacturing, per cent, for large and small economies, 1991. ... 5
Table 5: Analysis of variance: R&D intensity in manufacturing for large and small economies, 1991. ... 5
Table 6: Correlation matrix of cross-country manufacturing R&D intensity profiles between 1979, 1983, 1987 and 1991. ... 10
Table 7: Mean R&D intensity in manufacturing, per cent, large and small economies, difference between large and small economies, percentage points. R^2, p-value and statistical significance from ANOVA. 1979, 1983, 1987 and 1991. ... 13
Table 8: Statistical association between GDP and R&D intensity in manufacturing, 1979, 1983, 1987 and 1991. Correlation coefficient (r), R^2 and statistical significance. ... 14
Table 9: Statistical association between Ln GDP and R&D intensity in manufacturing, 1979, 1983, 1987 and 1991. Correlation coefficient (r), R^2 and statistical significance. ... 14
Table 10: The 22 manufacturing industries distinguished in this study. ... 16
Table 11: Correlation of the profile of R&D intensities across industries in each country with the typical profile. Correlation coefficients (r). ... 23
Table 12: R&D intensity in manufacturing, per cent (1), decomposed into sum of structure component (2) and industry intensity component (3), 1991. ... 29
Table 13: R&D intensity in manufacturing, difference from average, percentage points (1), decomposed into sum of structure component (2) and industry intensity component (3), 1991. ... 30
Table 14: Mean and median industry intensity component value, large and small economies, 1991. 33
Table 15: Analysis of variance. Industry intensity component for large and small economies, 1991. 33
Table 16: Mean and median structure component value, large and small economies, 1991. ... 35
Table 17: Analysis of variance. Structure component for large and small economies, 1991 ... 36
Table 18: Correlation matrix of cross-country industry intensity component profiles between 1979, 1983, 1987 and 1991. ... 41
Table 19: Correlation matrix of cross-country structure component profiles between 1979, 1983, 1987 and 1991. ... 45
Table 20: Correlation matrix of inter-industry typical (median) R&D intensity profiles between 1979, 1983, 1987 and 1991. ... 47
Table 21: Mean industry intensity component value, large and small economies, difference between large and small economies. R^2, p-value and statistical significance from ANOVA. 1979, 1983, 1987 and 1991. ... 48

Table 22: Statistical association between GDP and industry intensity component, 1979, 1983, 1987 and 1991. Correlation coefficient (r), R^2 and statistical significance. 48
Table 23: Statistical association between Ln GDP and industry intensity component, 1979, 1983, 1987 and 1991. Correlation coefficient (r), R^2 and statistical significance. 49
Table 24: Mean structure component value, large and small economies, difference between large and small economies. R^2, p-value and statistical significance from ANOVA. 1979, 1983, 1987 and 1991. 50
Table 25: Statistical association between GDP and structure component, 1979, 1983, 1987 and 1991. Correlation coefficient (r), R^2 and statistical significance. 50
Table 26: Statistical association between Ln GDP and structure component, 1979, 1983, 1987 and 1991. Correlation coefficient (r), R^2 and statistical significance. 51
Table 27: Countries ranked according to the weighted mean of ranks indicator. 76
Table 28: Ranks of the different countries on the industry intensity component (IIC) and the weighted mean of ranks (WMR) indicators, 1991... 80
Table 29: The weighted mean of ranks indicator values, original and transformed expression, 1991. 81
Table 30: Correlation, for each country, between the typical R&D intensity of each industry and the country's relative R&D performance in each industry as measured by modified z scores, 1991. 83
Table 31: Correlation matrix of distribution of weighted mean of ranks indicator values across countries between 1979, 1983, 1987 and 1991. 89
Table 32: Correlation coefficient (r) between, on the one hand, the weighted mean of ranks indicator, standardized version, and, on the other hand, GDP and Ln GDP for the years 1979, 1983, 1987 and 1991. 90

FIGURES

Figure 1: GDP, thousand million US dollars (x-axis), and R&D intensity in the manufacturing sector, per cent (y-axis), 1991. 6
Figure 2: Ln GDP, US dollars (x-axis), and R&D intensity in the manufacturing sector, per cent (y-axis), 1991. 7
Figure 3: R&D intensity in the manufacturing sector, 1979, 1983, 1987 and 1991. 8
Figure 4: R&D intensity in the manufacturing sector, 1979, 1983, 1987 and 1991. 9
Figure 5: Correlation matrix of cross-country manufacturing R&D intensity profiles between all years in the period 1979 - 1991. Graphical presentation.. 11
Figure 6: Correlation of the profile of relative manufacturing R&D intensities across countries in each year with the profile in each succeeding year. Number of years before the correlation coefficient gets below 0.95, and year in which this happens. 12
Figure 7: R&D intensity, per cent, and share of total manufacturing value added, per cent, for 22 industries, USA, 1991. 17

Figure 8: R&D intensity, per cent, and share of total manufacturing value added, per cent, for 22 industries, Norway, 1991 18
Figure 9: Typical (median) R&D intensity, per cent, and typical (mean) share of total manufacturing value added, per cent, for 22 industries, among 13 countries, 1991. 20
Figure 10: R&D intensities in each industry, per cent, in the USA and median across all countries. 21
Figure 11: R&D intensities in each industry, per cent, in the USA (along the y-axis) and median across all countries (along the x-axis). 22
Figure 12: R&D intensities in each industry, per cent, in the Netherlands and median across all countries. 24
Figure 13: R&D intensity in manufacturing, difference from average, percentage points (figures in parenthesis), decomposed into sum of structure component (x-axis) and industry intensity component (y-axis), 1991. 31
Figure 14: GDP, thousand million US dollars (x-axis), and industry intensity component value (y-axis), 1991. 34
Figure 15: Ln GDP, US dollars (x-axis), and industry intensity component value (y-axis), 1991 35
Figure 16: GDP, thousand million US dollars (x-axis), and structure component value (y-axis), 1991. 37
Figure 17: Ln GDP, US dollars (x-axis), and structure component value (y-axis), 1991. 38
Figure 18: Industry intensity component values, 1979, 1983, 1987 and 1991. 39
Figure 19: Industry intensity component values, 1979, 1983, 1987 and 1991. 40
Figure 20: Correlation matrix of cross-country industry intensity component profiles between all years in the period 1979-1991. Graphical presentation 42
Figure 21: Structure component values, difference from average version, 1979, 1983, 1987 and 1991. 43
Figure 22: Structure component values, difference from average version, 1979, 1983, 1987 and 1991. 44
Figure 23: Correlation matrix of cross-country structure component profiles between all years in the period 1979-1991. Graphical presentation. 45
Figure 24: Typical (median across all countries) R&D intensity in each industry (per cent) in 1979, 1983, 1987 and 1991. 46
Figure 25: Correlation matrix of inter-industry typical (median) R&D intensity profiles between all years in the period 1979-1991. Graphical presentation 47
Figure 26: Decomposition of industry intensity component into main term and interaction term, 1991. 70
Figure 27: Decomposition of industry intensity component into main term and interaction term, excluding Australia and Netherlands, 1991. 71
Figure 28: Finland's R&D intensity ranking in different industries (y-axis). Industries weighted by share of value added and ranked according to performance (x-axis). 1991. 77
Figure 29: Italy's R&D intensity ranking in different industries (y-axis). Industries weighted by share of value added and ranked according to performance (x-axis). 1991. 78

Figure 30: Japan's, USA's and UK's R&D intensity ranking in different industries (y-axis). Industries weighted by share of value added and ranked according to performance (x-axis). 1991. . 79
Figure 31: Z scores weighted mean of ranks indicator, standardized version (y-axis), z scores industry intensity component (x-axis), relative performance correlation coefficient (in parentheses), 1991. 85
Figure 32: Weighted mean of ranks indicator values (standardized version), 1979, 1983, 1987 and 1991. 88

INTRODUCTION

Expenditure on research and development (R&D) as a proportion to some measure of total economic activity is a frequently used measure for comparing the extent of innovation activities across different kinds of units, be they firms, industries, national economies, etc. This ratio is often referred to as *R&D intensity*. At the level of national economies, total R&D expenditures to GDP (gross domestic product) in a given year is frequently used to compare the innovative efforts of different countries. Often countries are simply compared on the basis of this ratio. For instance, R&D expenditures of a given country may be judged insufficient by a reference to data showing that the R&D intensity of the country in question is substantially smaller than the R&D intensity of certain other countries.

However, there are some problems connected to making these kinds of comparisons. One problem is that the ratio R&D to GDP appears to be positively correlated with absolute size of the economy as measured by GDP. This association has been accounted for as reflecting scale effects.[1] Thus, when the economies involved are of different sizes, just comparing the ratio of R&D expenditures to GDP without taking this relationship into account may be problematic.

Another complicating factor is that there are large differences in *industrial structure* across countries. A problem with this is that the differences in R&D intensity across countries to some extent, large or small, may reflect differences in industrial structure. If this to a significant extent is the case, how should we take it into account when we compare R&D expenditures across different countries?

The present paper is intended to contribute to clarifying these issues. Thus, the paper is very limited in scope. For instance, it does not take account of innovation expenditures other than R&D expenditures. Thus, the question of the variation in the composition of R&D expenditures across industries, and what consequences this may have for the problems addressed in the paper, is not studied. More generally, the

1 Cf. J.A.D. Holbrook, 'The influence of scale effects on international comparisons of R&D expenditures', Science and Public Policy, 1992.

level of innovation in different economies is influenced by a number of broader social factors like the mode of organization prevalent in business enterprises, cultural and ideological factors, the relationship between social classes and groups, the nature of the macroeconomic regime, etc. The paper does not address any of these broader issues.

The data

In this paper we will use data from the OECD STAN and ANBERD databases for 12 countries: Australia, Canada, Denmark, Finland, France, Germany, Italy, Japan, the Netherlands, Sweden, the UK and the USA. In addition, for Norway we will use data from Statistics Norway. The Norwegian data should be fully compatible with the OECD data and will probably soon be included in STAN/ANBERD.

The data used in this paper cover the *manufacturing* sector only. Thus, in what follows the *R&D intensity* of each country is not the R&D intensity of the whole economy but of the *manufacturing* sector only, i.e. it is total R&D expenditures of the manufacturing sector expressed as a percentage of total *value added* of the manufacturing sector (whereas GDP is the value added of the economy as a whole). Similarly, the industrial structure of each country in this paper refers to the composition of industries of the manufacturing sector only. Thus, for instance, the share accounted for by manufacturing, primary industries, services, etc. in the economy as a whole is not taken into account.

STAN has industry level data on sales, employment, exports, value added, etc. for each year. In the following only the value added figures will be used. The data allow us to break down total value added in manufacturing on 22 different industries. ANBERD has data on R&D expenditures in manufacturing broken down on the same 22 industries. The figures are total expenditures on R&D in each industry, irrespective of source of funding (whether private or public, by individual firms or industry associations, etc.).

There is a limited number of missing values in the data. In the analysis below the missing values have been replaced by estimates made by the present authors.

DIFFERENCES IN R&D INTENSITY ACROSS COUNTRIES

Let us start looking at the data. If we divide total R&D expenditures in the manufacturing sector by total value added in the manufacturing sector in a given time period, we get what we will here call the *R&D intensity* of the manufacturing sector. If we multiply this quotient by 100, the intensity will be expressed in per cent.

Table 1, below, shows R&D intensities in manufacturing for 1991 for the 13 countries included in the study, ranked according to these intensities.

Table 1: R&D intensities in the manufacturing sector, per cent, 1991.

Country	*R&D instensities*
Sweden	9.6
USA	8.6
Japan	7.1
France	6.6
UK	6.2
Germany	6.0
Finland	5.8
Netherlands	4.9
Norway	4.7
Denmark	4.3
Canada	3.6
Italy	3.2
Australia	2.6
Mean	5.6
St. dev	2.0

As we see, there is substantial variation in the R&D intensity of the manufacturing sector across countries, even among the countries included in this study, which are all advanced OECD economies. To get a better picture of these differences, let us express all R&D intensities in terms of differences from the average intensity, which is 5.6 per cent. This is shown in Table 2, below.

Table 2: R&D intensities in the manufacturing sector, per cent, differences from average, 1991.

Country	R&D intensity
Sweden	3.9
USA	2.9
Japan	1.5
France	0.9
UK	0.6
Germany	0.4
Finland	0.2
Netherlands	-0.7
Norway	-0.9
Denmark	-1.3
Canada	-2.0
Italy	-2.4
Australia	-3.0

There seems to be a certain tendency here for manufacturing R&D intensity to increase with the absolute size of the economy. Let us divide the countries into large and small economies, with the six largest defined as large and the remaining seven defined as small, as shown in Table 3, below.

Table 3: Classification of countries into large and small economies

Large	*Small*
USA	Canada
Japan	Australia
Germany	Netherlands
France	Sweden
Italy	Denmark
UK	Finland
	Norway

If we think of the USA, Japan, Germany, France, the United Kingdom and Italy as large economies and the other seven countries as small economies, we see that in general the large economies have an R&D intensity above average while the small ones are below average. This is confirmed by the average and the median for each group, shown in Table 4, below.

Table 4: Mean and median R&D intensity in manufacturing, per cent, for large and small economies, 1991.

	Mean	Median
Large	6.3	6.4
Small	5.1	4.7

There are three exceptions to this pattern, however. One of these is quite minor, namely that Finland has an R&D intensity which is 0.2 percentage points above average. The other two are major exceptions, however. Sweden in fact has the highest intensity of all the countries, while Italy has an intensity far *below* average. Thus, the variation inside each of these categories is considerable. This is confirmed by Analysis of variance (ANOVA), as reported in Table 5, below.

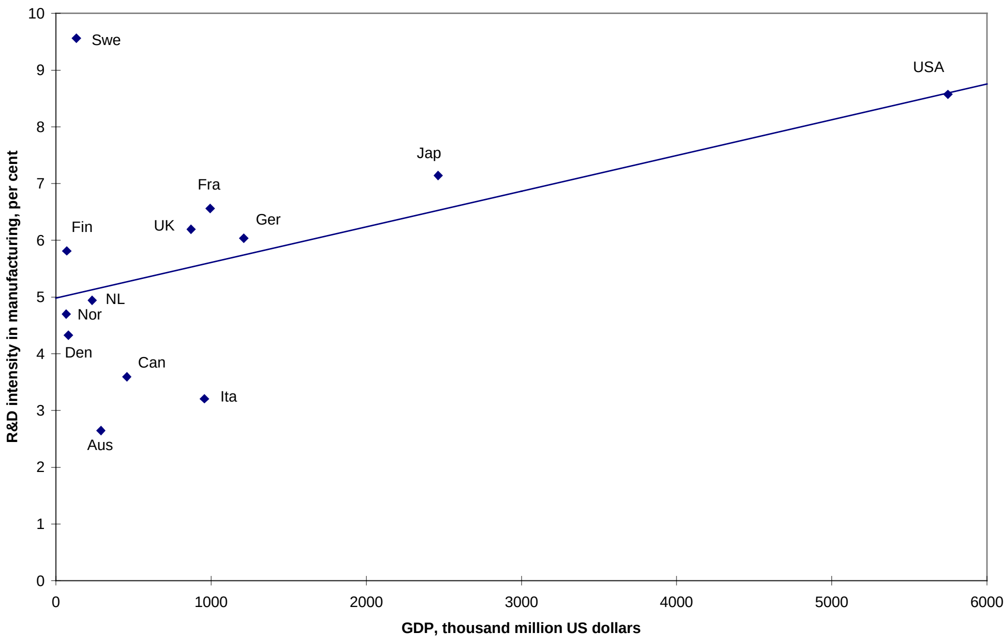
Table 5: Analysis of variance: R&D intensity in manufacturing for large and small economies, 1991.

Source of variation	*Sum of squares*	*DF*	*Mean square*	*F*	*P-value*
Between	4.662993	1	4.662993	1.137904	0.3090
Within	45.07665	11	4.097878		
Total	49.73965	12			
R^2	0.09				

As we see, the difference between the groups is *not* statistically significant, with a p-value of 0.309. R^2 is only 0.09.

Let us look closer into the relationship between size of economy and R&D intensity in the manufacturing sector. Figure 1, below, shows the relationship between GDP in US dollars and manufacturing R&D intensity for the 13 countries in 1991. Purchasing power parities (PPP) have been used to transform local currencies into US dollars.

Figure 1: GDP, thousand million US dollars (x-axis), and R&D intensity in the manufacturing sector, per cent (y-axis), 1991.



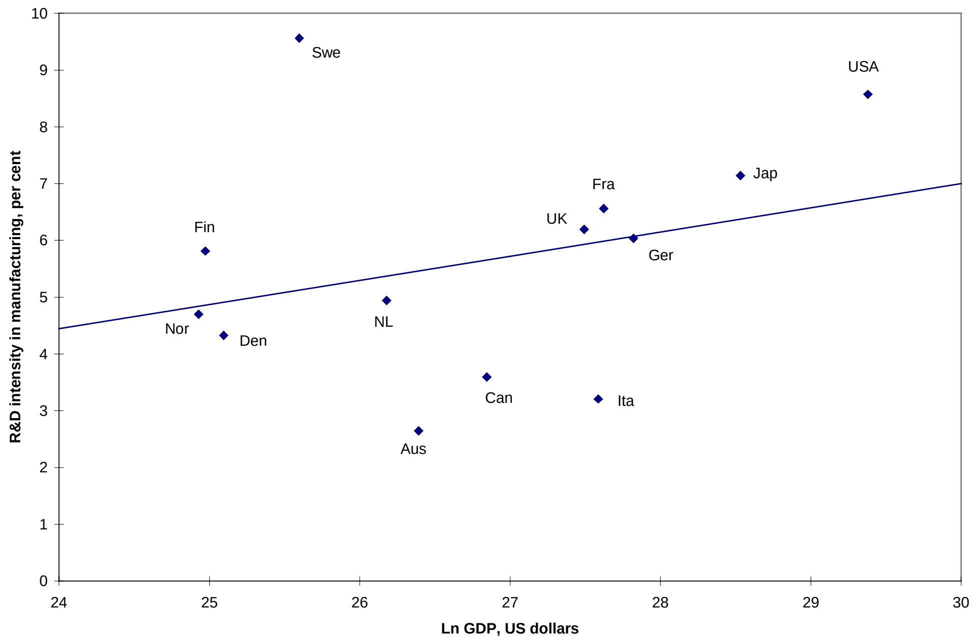
There is to a certain degree a positive linear relationship between these two variables, as shown by the regression line. However, the association is not a strong one. We have a correlation of $r = 0.48$, with $R^2 = 0.23$. This is *not* significant at the 5 per cent level, although it is significant at the 10 per cent level.

However, the distribution of GDP is a very uneven one, with the USA becoming a very influential observation. The mean of the distribution is 1044 thousand million dollars, while the median is only 456 thousand million dollars. With a standard deviation of 1567 thousand million dollars, this means that the difference between the mean and the median equals 37.5 per cent of the standard deviation. The skewness statistic is 2.62. Now, if the distribution is perfectly symmetrical, this statistic becomes 0. Furthermore, there is a “rule of thumb” that the skewness statistic must exceed 0.8 in absolute value before the distribution gets “noticeably skewed”.[2] Clearly, the present distribution is “noticeably skewed”.

2 See Michael S. Lewis-Beck, Data Analyses: An Introduction, Sage University Paper series on Quantitative Applications in the Social Sciences, 07-103, Thousand Oaks: Sage, 1995, p. 16.

I have therefore chosen also to look at the relationship between size of economy and R&D intensity in the manufacturing sector using the *natural logarithm* of GDP as a measure of economy size in place of the GDP itself. This relationship between the Ln GDP and manufacturing R&D intensity in 1991 is shown in Figure 2, below.

Figure 2: Ln GDP, US dollars (x-axis), and R&D intensity in the manufacturing sector, per cent (y-axis), 1991.



We see that we get a much more even distribution when we measure economy size by *Ln GDP* instead of GDP itself. Here the mean and the median are very close, 26.80 against 26.85; the difference is only 3 per cent of the standard deviation of the distribution (which is 1.42). The skewness statistic is only 0.17, hence the distribution is not noticeably skewed.

Also in this case we find a certain positive association between economy size and manufacturing R&D intensity, as witnessed by the regression line. However, the association is a weak one, with $r = 0.30$, $R^2 = 0.09$. This relationship is *not* statistically significant.

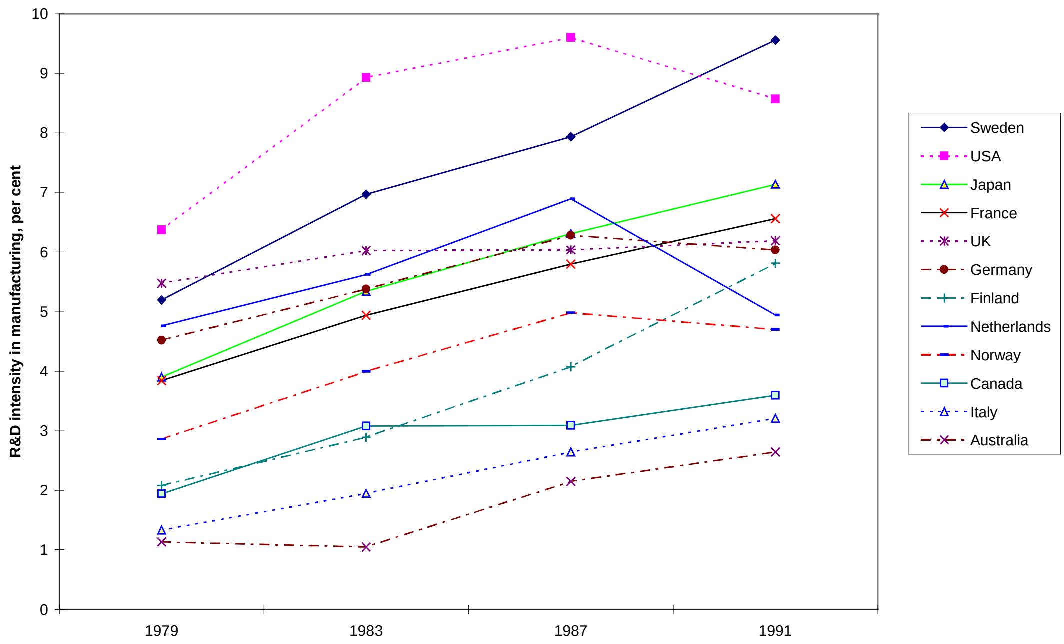
Thus, while we see that there is a certain tendency for the manufacturing R&D intensity to increase with economy size, we cannot be sufficiently certain that the increase in R&D intensity has *anything to do with* increasing economy size. The larger countries might just happen, on average, to have higher R&D intensity.

Evolution over time in differences in R&D intensity across countries

Let us see if we get the same results when we look at other years than 1991. We have these data for all years from 1979 to 1991, 1991 being the most recent year. We have chosen to single out four different years for more detailed analysis. In addition to 1991, they are 1979 (the first year), 1983 and 1987, i.e. every fourth year from 1979 to 1991.

Figure 3, below, shows R&D intensity in the manufacturing sector for the 13 countries in the four years mentioned.

Figure 3: R&D intensity in the manufacturing sector, 1979, 1983, 1987 and 1991.

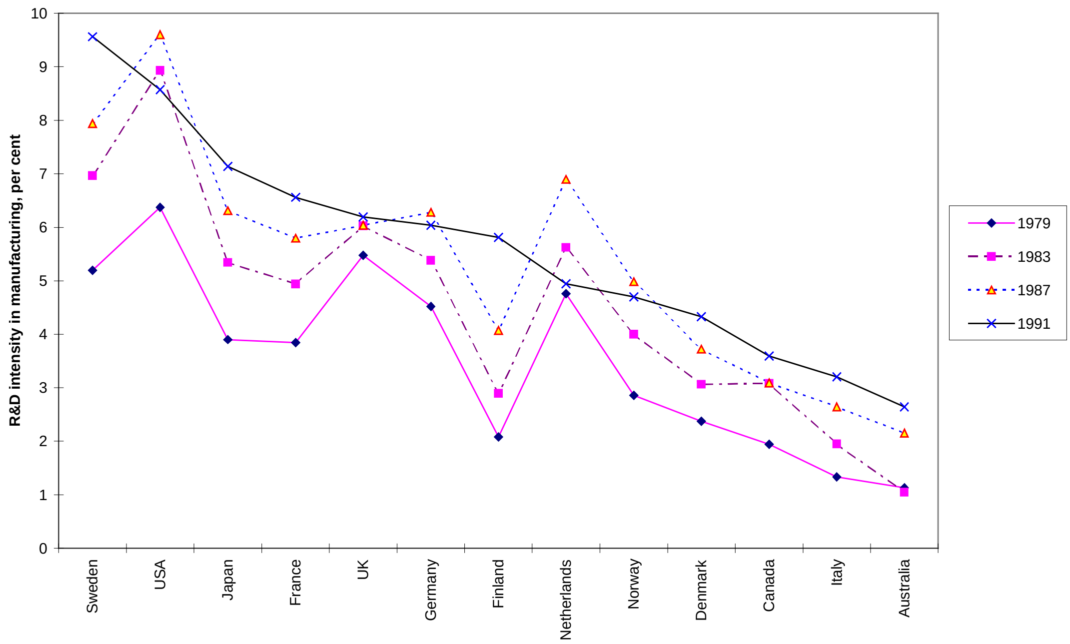


We see that from 1979 to 1983 and from 1983 to 1987 there seems to be roughly the same development in all countries. With only a couple of exceptions, there R&D

intensity increases both from 1979 to 1983 and from 1983 to 1987. From 1987 to 1991 the development seems less uniform across countries, however. While the R&D intensity in most of the countries continues to increase, in some countries, namely the USA, Germany, Netherlands and Norway, it drops. In addition, there is virtually no change for the UK, which was also the case from 1983 to 1987.

To give an additional perspective of these changes, the same information as in the previous figure is presented in Figure 4, below, but this time in the form of "profiles" across the countries for each of the four years.

Figure 4: R&D intensity in the manufacturing sector, 1979, 1983, 1987 and 1991.



In this figure, the countries are ranked along the x-axis according to manufacturing R&D intensity in 1991. For each year we get a "profile" of how the countries compare to each other in this regard. We see that the changes we noted in the previous figure in the development from 1987 to 1991 compared to the developments from 1979 to 1983 to 1987, have the implication that compared to the cross-country profiles for 1979, 1983 and 1987, which look quite similar to each other, the profile for 1991 is noticeably different.

To try to give a numerical expression of these differences, we have correlated the profiles in each year with each other. Here each country is the unit of analysis and the variables are the R&D intensity in each year. The set of R&D intensities for each of the 13 countries in one year is thus correlated with the set of R&D intensities for the same countries in another year.

Of course, the profile in one year correlated with itself gives $r = 1$. If the profiles in two different years are exactly the same, i.e. if there are no changes from one year to another, this also gives a correlation coefficient of 1, as will also be the case if all countries have the same rate of growth from one year to another. However, if the manufacturing R&D intensities of different countries have different rates of growth from one year to another, the correlation coefficient will be less than 1, and the more the growth rates of different countries diverge from one year to another, the less the correlation coefficient between the profiles of these two years.

The results of correlating these four profiles with each other are shown in the correlation matrix below (Table 6).

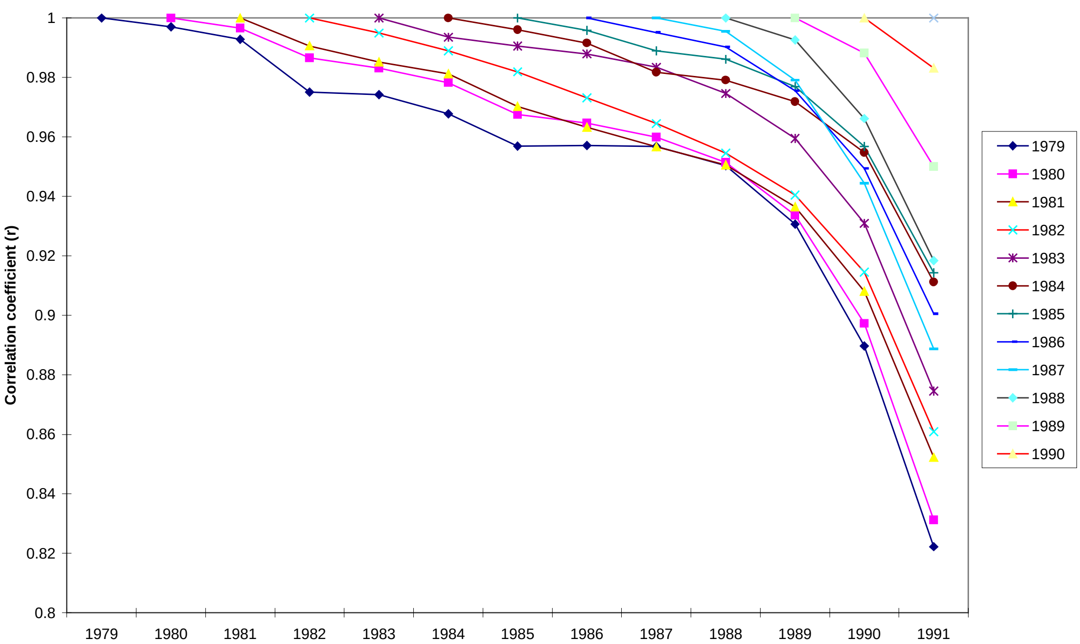
Table 6: Correlation matrix of cross-country manufacturing R&D intensity profiles between 1979, 1983, 1987 and 1991.

Year	*1979*	*1983*	*1987*	*1991*
1979	1			
1983	0.97	1		
1987	0.96	0.98	1	
1991	0.82	0.87	0.89	1

This correlation matrix seems to confirm the impression that 1991 is noticeably different from 1979, 1983 and 1987, while there are not so large differences among these latter three. Thus, while the correlation between 1979 and 1983 is 0.97 and between 1983 and 1987 is 0.98, the correlation between 1987 and 1991 is only 0.89. Similarly, while the correlation between 1979 and 1987 is 0.96, the correlation between 1983 and 1991 is only 0.87. (There is nothing to compare to the correlation between 1979 and 1991, of course.)

It thus seems evident that the most recent of these four years, 1991, is atypical compared to the other three. An important question is whether the 1991 deviation is pure anomaly, not fitting into any recognizable pattern (perhaps expressing inaccuracies in the data for this particular year), or whether it can be explained by the onset of a process of more rapid change somewhere between 1987 and 1991. To get some indications of the answer to this question, we have correlated the profiles of each year with each other in the same way as in the correlation matrix in Table 6, above, but instead of using only the four years included there, we have correlated the profiles of *all* years in the period 1979-1991 with each other. The resulting correlation matrix is shown, graphically, in Figure 5, below.

Figure 5: Correlation matrix of cross-country manufacturing R&D intensity profiles between all years in the period 1979 - 1991. Graphical presentation.



This graph strongly suggests that the single year 1991 does *not* represent an inexplicable aberration compared to earlier years. Rather, compared to the period 1979-1987, in the period from 1987 to 1991 there seems to be occurring more rapid change in the relative manufacturing R&D intensities among the countries. From 1979 to 1988 the correlation coefficients fall moderately from one year to the next, and there is no evidence of any break in the trend. However, when we go from 1988 to 1989 the fall seems to be noticeably larger, and this seems to accelerate as we go from 1989 to 1990 and then from 1990 to 1991. Thus, around 1988/1989 we seem to have entered a period of accelerating change in the pattern of relative manufacturing R&D intensities across countries, and there is no sign that there has emerged a new relatively stable pattern by 1991, quite the contrary.

That the changes in the profile of relative manufacturing R&D intensities across the countries started to become more rapid around 1989/1990 is also indicated by the following table (Table 7), based on the correlation matrix presented in Figure 5, above. It shows, when we take the profile of each year and correlate it successively with each of the succeeding years, how many years it takes before the correlation coefficient gets below 0.95, and the year in which this happens.

Figure 6: Correlation of the profile of relative manufacturing R&D intensities across countries in each year with the profile in each succeeding year. Number of years before the correlation coefficient gets below 0.95, and year in which this happens.

year of departure	*no. of years before r gets below 0.95*	*First year where r is below 0.95*
1979	10	1989
1980	9	1989
1981	8	1989
1982	7	1989
1983	7	1990
1984	7	1991
1985	6	1991
1986	4	1990
1987	3	1990
1988	3	1991
1989	2	1991
1990	-	-

For instance, we see that if we take the profile in 1979 and correlate it with the profile in each of the succeeding years, it takes 10 years before the correlation coefficient gets below 0.95, and this then happens in 1989. Similarly, if we take the profile of 1986 as the point of departure, it takes only four years, and this then happens in 1990.

We see that as the point of departure moves from 1979 through to 1989, the time this takes gets shorter and shorter, until we are down to only two years when starting in

1989. Moreover, irrespective of year of departure, this happens either in 1989, 1990 or 1991. Only for the profile of the year 1990 do we not get an *r* below 0.95 by 1991. (For 1991 we do not have any subsequent years in our data series, of course.)

I will not speculate about the causes of these changes here in this paper or about whether they are real at all or for instance reflect changes in data quality or the like.

In any case, we should note that the lowest correlation we find here, that between 1979 and 1991, the first and the last year of the series, is still as high as 0.82. This means that the profiles at the start and at the end of the series still are *roughly* similar.

With this information of changes within a nevertheless roughly stable pattern between 1979 and 1991, let us look at the relationship between economy size and manufacturing R&D intensity also in 1979, 1983 and 1987.

Let us first look at the difference between the rough categories large and small economies, as defined in Table 3, above. The following table (Table 8) shows, for each of the four years, the mean R&D intensity in manufacturing for large and small economies (in per cent), the difference between the two (in percentage points), and R^2 and p-value from ANOVA (cf. Table 5, above). Lastly, the p-value is translated into a statement of statistical significance at conventional levels (10 per cent, 5 per cent or 1 per cent).

Table 7: Mean R&D intensity in manufacturing, per cent, large and small economies, difference between large and small economies, percentage points. R^2, p-value and statistical significance from ANOVA. 1979, 1983, 1987 and 1991.

Year	*Mean large economies*	*Mean small economies*	*Difference*	*R^2*	*p-value*	*statistical significance*
1979	4.2	2.9	1.3	0.17	0.1644	not significant
1983	5.4	3.8	1.6	0.15	0.1908	not significant
1987	6.1	4.7	1.4	0.11	0.2578	not significant
1991	6.3	5.1	1.2	0.09	0.3089	not significant

We see that the results from 1979, 1983 and 1987 are quite similar to those from 1991. On average, the large economies have somewhat higher R&D intensity in manufacturing than the small ones, but the difference is not statistically significant and R^2 is low.

Next, let us look at the correlation between GDP and R&D intensity in manufacturing in the different years. This is shown in Table 8, below.

*Table 8: Statistical association between GDP and R&D intensity in manufacturing, 1979, 1983, 1987 and 1991. Correlation coefficient (*r*), R^2 and statistical significance.*

Year	*r*	R^2	*statistical significance*
1979	0.56	0.31	.05 level
1983	0.64	0.41	.05 level
1987	0.61	0.37	.05 level
1991	0.48	0.23	.10 level

As we see, generally there is a moderate positive correlation between GDP and R&D intensity in manufacturing. The relationship is significant at the 5 per cent level for the first three years, but only at the 10 per cent level for 1991. However, the USA is here an outlier which seems to create problems. Let us compare these results to the results we get when we use *Ln GDP* instead of GDP as a measure of economy size, reported in Table 9, below.

*Table 9: Statistical association between Ln GDP and R&D intensity in manufacturing, 1979, 1983, 1987 and 1991. Correlation coefficient (*r*), R^2 and statistical significance.*

Year	*r*	R^2	*statistical significance*
1979	0.47	0.22	not significant
1983	0.48	0.23	.10 level
1987	0.41	0.17	not significant
1991	0.30	0.09	not significant

We see that the correlation here is low and in general not statistically significant.

To sum up, consideration of 1979, 1983 and 1987 does not give results which differ in any important way from what we found for 1991. Although we find a slightly stronger association between economy size and R&D intensity in manufacturing in the earlier years than in 1991, it is still the case that the association is not strong and not statistically significant. We do find a certain tendency for the manufacturing R&D intensity to increase with economy size, but we cannot be sufficiently certain that the increase in R&D intensity has *anything to do with* increasing economy size.

On the other hand, if a substantial positive association between economy size and R&D intensity in manufacturing is considered to be a result established by previous result, our results would not constitute strong evidence to the contrary, either. For instance, if the null hypothesis is that the correlation coefficient is at least 0.50, we would not be able to refute this hypothesis. If a 5 per cent significance level is applied, even the lowest correlation coefficient reported above, 0.30 for the correlation between Ln GDP and R&D intensity in 1991, would not be significantly different from the null hypothesis. With a correlation coefficient of 0.30 and 13 observations, we can only say that with a probability of 95 per cent the 'true' correlation coefficient is less than 0.68. This is still quite high, however.

Differences in R&D intensity across industries

Up till now, we have only looked at the manufacturing sector in each country as one undifferentiated quantity. Now, of course, the manufacturing sector in any country is made up of a multitude of different kinds of activities and may in principle be divided into any number of distinct industries. The data we have on R&D expenditures and value added permit us to divide the manufacturing sector into 22 different industries, defined by ISIC numbers at the 2, 3 and 4 digit level. Thus, we may not only compute an R&D intensity for the manufacturing sector as a whole in each country, but for each of the 22 industries as well. Furthermore, we may compute the share of total manufacturing value added which each industry accounts for. This makes it possible to compare across countries the R&D intensity not only of the manufacturing sector as a whole, but also of the individual industries. Secondly, it makes it possible to for us to take account of *industrial structure* when comparing R&D intensities in manufacturing across countries. The fundamental fact which makes industrial structure of importance in this connection is that, after variation in R&D intensity across countries within particular industries has been accounted for, there still remains enormous variation in R&D intensity across industries.

Let us first look at which the 22 industries distinguished in the data are. This is shown in Table 10, below.

Table 10: The 22 manufacturing industries distinguished in this study.

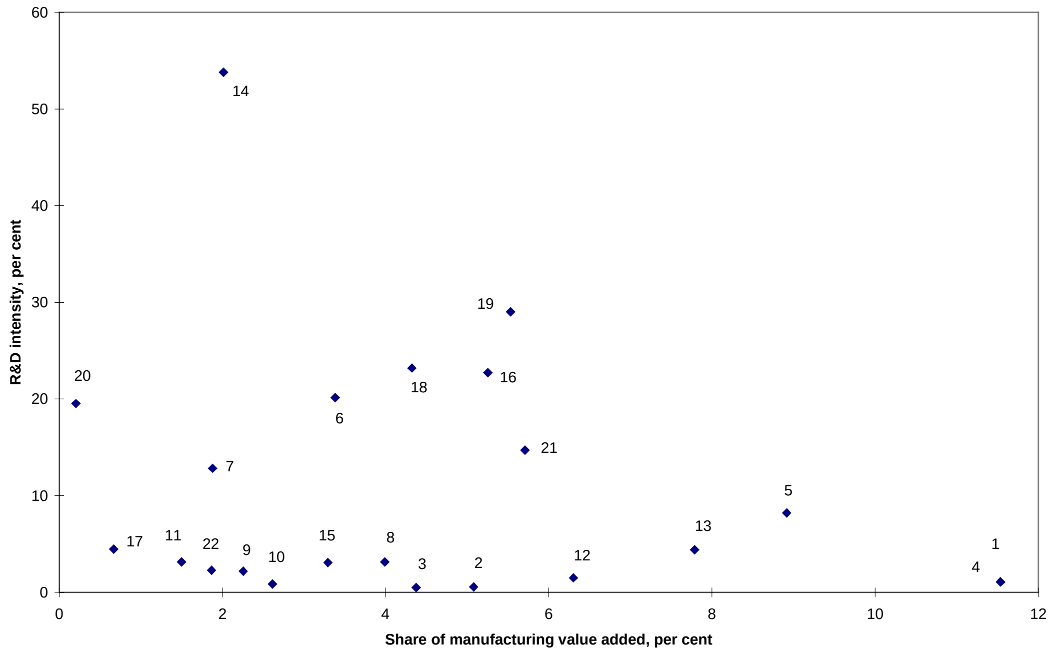
ref no.	ISIC code	Industry
1	31	Food, beverages and tobacco
2	32	Textiles, apparel and leather
3	33	Wood products and furniture
4	34	Paper, paper products and printing
5	351+352-3522	Chemicals excluding drugs
6	3522	Drugs and medicines
7	353+354	Petroleum refineries and products
8	355+356	Rubber and plastic products
9	36	Non-metallic mineral products
10	371	Iron and steel
11	372	Non-ferrous metals
12	381	Metal products
13	382-3825	Non-electrical machinery
14	3825	Office and computing machinery
15	383-3832	Electrical machinery excluding communication equipment
16	3832	Radio, TV and communication equipment
17	3841	Shipbuilding and repairing
18	3843	Motor vehicles
19	3845	Aircraft
20	3842+3844+3849	Other transport equipment
21	385	Professional goods
22	39	Other manufacturing

In addition to the ISIC codes of each of the industries, the table also gives each industry a reference number, from 1 to 22. These reference numbers are used in some of the figures below.

To get a rough impression of the variation in R&D intensities across industries and of the point about the importance of industrial structure, let us look more closely at a couple of selected countries. we have chosen the USA and Norway as examples. As we see from Table 1, above, USA has an R&D intensity of 8.6 per cent in manufacturing in 1991, while Norway has only 4.7 per cent.

Figure 7, below, shows R&D intensity and share of total manufacturing value added for each of the 22 industries in the USA in 1991

Figure 7: R&D intensity, per cent, and share of total manufacturing value added, per cent, for 22 industries, USA, 1991.

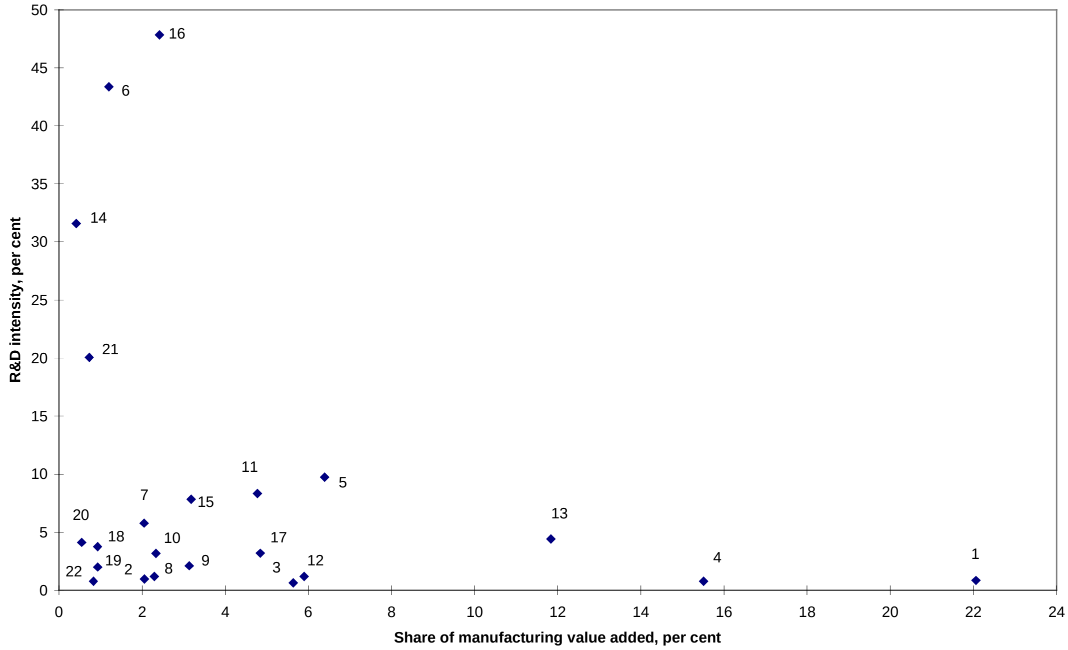


In this figure, the numbers identifying the industries are the same as the reference numbers from Table 10, above.

The enormous variation in R&D intensity across industries is apparent. We see that some industries have an R&D intensity of more than 20 per cent, other industries have only 1 per cent, a couple of them even less than 1 per cent. Furthermore, we get the impression that the industries with really high R&D intensities together make up a much smaller share of total value added than the industries with low R&D intensities. This would have shown even clearer had it not been for the curious coincidence that the two industries which rank highest in terms of value added, *Food, beverages and tobacco* (no. 1) and *Paper, paper products and printing* (no. 4) both are so close both on R&D intensity and value added that they cannot be distinguished in the diagram. Both have very low R&D intensity (about 1.1 per cent), but together they account for 23.1 per cent of value added in manufacturing.

Let us now look at Norway. R&D intensity and share of value added for the 22 industries are shown in Figure 8, below.

Figure 8: R&D intensity, per cent, and share of total manufacturing value added, per cent, for 22 industries, Norway, 1991.



First, notice that the scales along the two axes are not the same as in the previous figure. In the US case, no industry accounts for more than 12 per cent of value added, while in the Norwegian case *Food, beverages and tobacco* account for more than 22 per cent. Similarly, along the y-axis we had to use 60 per cent R&D intensity as maximum value to cover all the industries in the US case, while in the Norwegian case 50 per cent is sufficient. This should be kept in mind when one compares the two charts.

Now, we saw that the USA has almost twice as high R&D intensity in manufacturing as Norway. It seems quite evident, however, that the USA in general does not have twice the R&D intensity as Norway in each individual industry. Rather, what is striking when one compares the two charts is that in Norway the industries lie closer to the two axes than in the USA. Specifically, in Norway the industries with really high R&D intensity seem to account for a much smaller share of total manufacturing production than in the USA.

This illustrates a general point. We see that there may be two quite different reasons why a country may have a high R&D intensity in manufacturing (and vice versa for a low R&D intensity). The most straightforward reason, one might say, is simply that the country quite generally has a high R&D intensity in each industry, or at least in most industries. However, it might also have a high R&D intensity in manufacturing without having any particularly high intensity in any industry, namely if the industrial structure of the country is such that it has a relatively high share of manufacturing production in industries with high R&D intensity and a relatively low share of production in industries with low R&D intensity. Indeed, it is quite possible for one country to have a lower R&D intensity than another in every single industry and still have a higher R&D intensity in manufacturing as a whole. Now, of course, a high R&D intensity may also reflect a combination of these two reasons. In the following we will look more closely into the relationship between R&D intensity and industrial structure. It has already been indicated that a large part of the difference in R&D intensity in manufacturing between USA and Norway may be explained by differences in industrial structure (but in no way necessarily all of it).

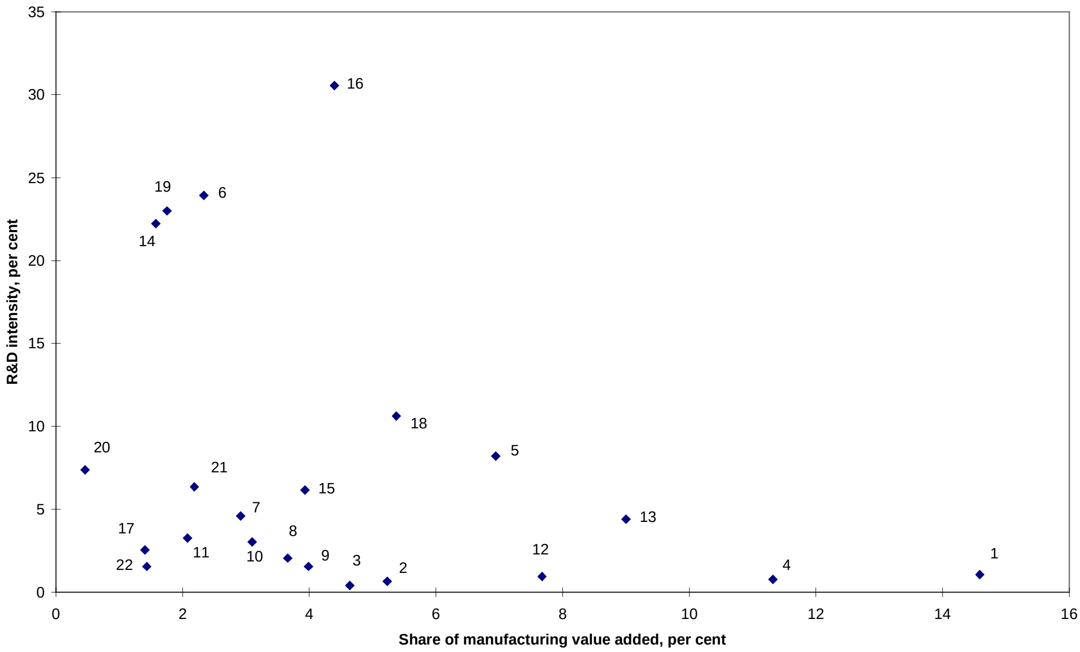
First, let us try to find an expression which might reasonably be said to capture a *typical* R&D intensity for each industry and a *typical* industrial structure, when R&D intensities in each industry and the industrial structure of *all* the 13 countries are taken into account. These *typical* values will be central in the analysis which follows. This especially applies to the typical R&D intensities of each industry.

For the typical industrial structure, i.e. the typical share of total manufacturing value added of each industry, we have simply chosen the *mean* (simple, unweighted average) for each industry across all the 13 countries.

For the typical R&D intensities, we have chosen *not* to use the mean value. This is because in a couple of cases the presence of extreme values makes the mean unreasonable as a measure of what is typical. Instead, we have chosen to use the *median* value across the 13 countries as the typical R&D intensity for each industry. This solution would not have been reasonable in the case of the typical share of manufacturing value added accounted for by each industry, because if the median had been used here, the shares would in general not have summed to 1 (or 100 per cent). However, in the case of typical R&D intensities, no such problem is involved.

Figure 9, below, shows the typical R&D intensity and the typical share of total manufacturing value added for the 22 industries.

Figure 9: Typical (median) R&D intensity, per cent, and typical (mean) share of total manufacturing value added, per cent, for 22 industries, among 13 countries, 1991.

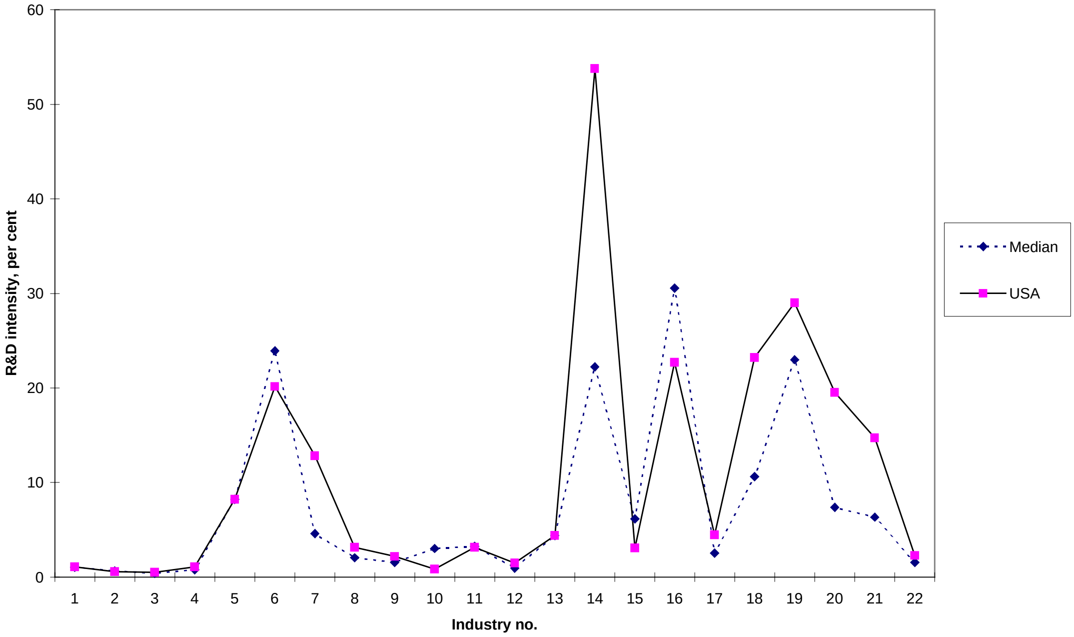


Note again that the scales along the two axes are not the same as in the previous figure.

We see that also when we take median R&D intensities across the 13 countries for each industry, we find enormous variation across industries in R&D intensity. Also the mean share of manufacturing production varies very much across industries.

To get an idea of ‘how typical’ these ‘typical values are, let us now compare these typical R&D intensities in each industry, defined by the median across all countries, to the R&D intensities of each industries in the individual countries. As an example, in Figure 10, below, the median R&D intensities in each industry are compared to the R&D intensities in the USA.

Figure 10: R&D intensities in each industry, per cent, in the USA and median across all countries.

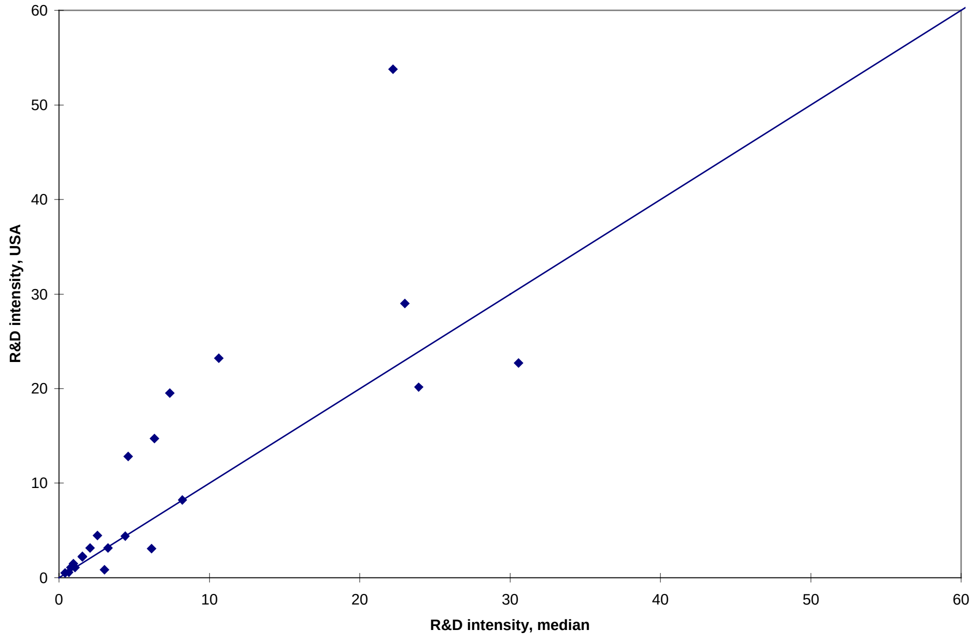


Again, the numbers identifying the industries are the reference numbers from Table 10, above.

The impression we get here is that the US profile of R&D intensity by industry is roughly similar to the typical profile. By and large it is the same industries which have high, respectively low, R&D intensities in both cases.

The same information may also be presented in the form of a scattergram, with each industry as units of observation and its R&D intensity in the USA plotted along one axis, its typical R&D intensity along the other. This is done in Figure 11, below.

Figure 11: R&D intensities in each industry, per cent, in the USA (along the y-axis) and median across all countries (along the x-axis).



The line drawn through the diagram is the 45 degrees line, where the R&D intensity in the USA is equal to the typical R&D intensity. For industries which are found to the left of this line, the US intensity is higher than the typical, and for industries on the right of this line, the US intensity is lower than the typical intensity. For industries which are placed exactly *on* the line, the US intensity value *is* the median value.

In a similar way as we did above with the profiles of relative manufacturing R&D intensities in manufacturing among countries in different years, we may use each industry as units of observation and correlate the US intensity values with the median intensity values. Correlating the US profile with the typical profile in this way, we get a correlation coefficient of 0.80. In the same way we can correlate each of the other country profiles with the typical profile. The results are shown in Table 12, below.

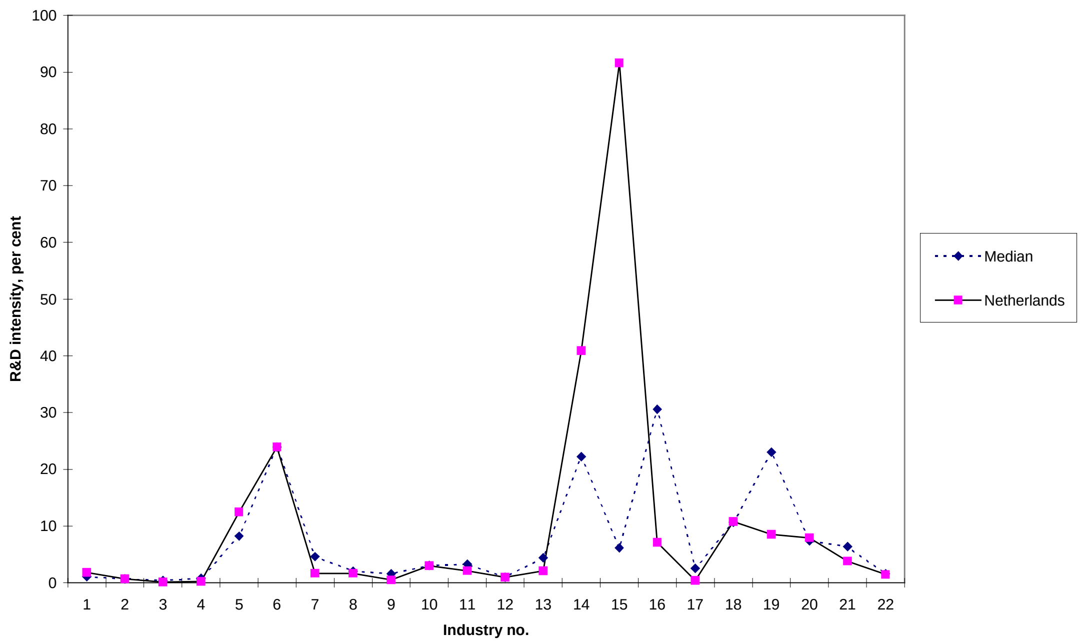
Table 11: Correlation of the profile of R&D intensities across industries in each country with the typical profile. Correlation coefficients (r).

Country	*Correlation*
UK	0.99
Sweden	0.91
Italy	0.90
France	0.85
Canada	0.84
Norway	0.82
Japan	0.81
USA	0.80
Denmark	0.73
Germany	0.72
Finland	0.66
Australia	0.62
Netherlands	0.29

Note that Denmark is special here, as the data do not register any production, and therefore no R&D intensity, in *Aircraft* (industry no. 19). Consequently, the correlation with the typical profile in the Danish case is made only on the basis of the 21 other industries.

We see that in general the correlations seem quite high, and certainly high enough to disconfirm any suspicion that the typical values defined might be quite arbitrary. Eight out of 13 correlation coefficients are higher than 0.80, and only one is lower than 0.60. However, this one, the Dutch case, is very low, only 0.29. Let us therefore look closer at the Dutch profile, again comparing it to the typical profile. This is done in Figure 12, below.

Figure 12: R&D intensities in each industry, per cent, in the Netherlands and median across all countries.



Perhaps surprisingly, we see that also the Dutch profile looks quite similar to the typical profile. We note one very large exception, however. In *Electrical machinery excluding communication equipment* (industry no. 15), the Netherlands has an R&D intensity which is extremely much higher than in any of the other countries. In this industry the R&D intensity is 91.6 per cent in the Netherlands, while the median is 6.2 per cent and no other country has higher than 12.5 per cent. In fact, this is one of the extreme values mentioned above (perhaps the most extreme) which motivated the use of the median instead of the mean as the definition of the typical value. Apart from this major deviation, the Dutch profile does not look too deviant. We get this confirmed if we run the correlation without industry no. 15, keeping only the 21 other industries. The correlation coefficient then immediately rises to 0.69, which is not particularly low.

Thus, it seems that the *typical* R&D profile of intensities across industries defined above on the whole is sufficiently similar to the profiles of each of the countries to be acceptable as expressing a characteristic of the industries as such, over and above the intra-industry variation we find across countries. Consequently, we should have no serious objections to using these typical R&D intensities for each industry as rough standards of reference in the following analysis.

R&D intensity in manufacturing and industrial structure

Let us return to the point made above that a high R&D intensity in manufacturing in a given country may reflect *either* on average a comparatively high R&D intensity inside each individual industry *or* an industrial structure characterized by a relatively high share accounted for by industries with high R&D intensity, or some combination of these two factors (and vice versa for a low R&D intensity in manufacturing). The underlying idea here is that the R&D intensity in the manufacturing sector as a whole in a given country may be conceived as the result of a combination of two analytically distinct components, on the one hand, the relative R&D intensity level inside the individual industries compared to other countries, on the other hand, the industrial structure of the country compared to other countries. In the following this idea is developed in a more formal way.

The point of departure is the R&D intensity in the manufacturing sector of a given country (i.e. any of the countries which we examine), which will be denoted by I_m , where *I* stands for R&D intensity and *m* stands for manufacturing. This is defined by

$$I_m = \frac{R_m}{V_m} \qquad (1)$$

where R_m denotes total R&D expenditures and V_m denotes total value added in the manufacturing sector of the country in question. Now,

$$R_m = R_1 + R_2 + \ldots\ldots\ldots + R_n \qquad (2)$$

where R_1, R_2, etc., denotes R&D expenditures in, respectively, industry no. 1, industry no. 2, etc., up to industry no. *n* (i.e. no. 22 in our case). By substituting this into equation (1), we can express the R&D intensity in the manufacturing sector by:

$$I_m = \frac{R_m}{V_m} = \frac{R_1}{V_m} + \frac{R_2}{V_m} + \ldots\ldots\ldots + \frac{R_n}{V_m} \qquad (3)$$

This expression we can transform by multiplying each of the components of the sum by $\frac{V_i}{V_i}$, i.e. by 1, where V_i is value added in industry no. i:

$$I_m = \left(\frac{R_1}{V_1}\cdot\frac{V_1}{V_m}\right)+\left(\frac{R_2}{V_2}\cdot\frac{V_2}{V_m}\right)+\ldots\ldots\ldots\ldots+\left(\frac{R_n}{V_n}\cdot\frac{V_n}{V_m}\right)$$

or

$$I_m = \sum_{i=1}^{n}\frac{R_i}{V_i}\cdot\frac{V_i}{V_m} \qquad (4)$$

Now, the first expression, $\frac{R_i}{V_i}$, gives the R&D intensity in industry no. i, which we will be denoted by I_i. The second expression, $\frac{V_i}{V_m}$, is industry *i*'s share of total manufacturing value added, which will be denoted by w_i (where *w* stands for *weight*). We then get:

$$I_m = \sum_{i=1}^{n} I_i \cdot w_i \qquad (5)$$

That is to say, the overall R&D intensity of the manufacturing sector is the weighted sum of the R&D intensities of all industries making up the manufacturing sector, when the weights are defined by each industry's share of total manufacturing value added.

Now, we have seen above that there are very large differences in R&D intensity across industries. However, for any given industry there are also differences in R&D intensity across countries. Let us now, in spite of these differences across countries, define a *typical* R&D intensity for each industry. As this typical value we have chosen the *median* R&D intensity across the 13 countries in the industry concerned. These median values are the ones depicted along the y-axis of Figure 8, above. Other definitions of the typical R&D intensity of each industry could have been chosen, of course, for instance the *mean* (the average) value across the 13 countries. However, as mentioned above, the presence of a small number of extreme, atypical values makes the mean less attractive as an indicator of the typical. The median is more resistant to atypical outliers.[3]

Let us denote the typical (in this case, the median) R&D intensity of each industry by $\bar{I}_i$. Now, expression (5)

$$I_m = \sum_{i=1}^{n} I_i \cdot w_i$$

3 Cf. Herbert F. Weisberg, Central Tendency and Variability, Sage University Paper series on Quantitative Applications in the Social Sciences, 07-183, Newbury Park: Sage, 1992, p. 30.

we can transform by adding *and* subtracting $\bar{I}_i$, that is to say by adding 0, to the factor I_i in each of the components of the sum:

$$I_m = \sum_{i=1}^{n} (I_i + \bar{I}_i - \bar{I}_i) \cdot w_i$$

which gives

$$I_m = \sum_{i=1}^{n} \left[\left(\bar{I}_i \cdot w_i \right) + \left(I_i - \bar{I}_i \right) \cdot w_i \right]$$

and

$$I_m = \sum_{i=1}^{n} \bar{I}_i \cdot w_i + \sum_{i=1}^{n} \left(I_i - \bar{I}_i \right) \cdot w_i \qquad (6)$$

The above equation (6) expresses the R&D intensity of the manufacturing sector in a given country as the sum of two different components. The first of these components, $\sum_{i=1}^{n} \bar{I}_i \cdot w_i$, is the sum over all industries of the product, for each industry, of the *typical* (median across the 13 countries) R&D intensity of the industry and the industry's share of manufacturing value added in the country concerned. This expression says what the R&D intensity in manufacturing in the country concerned would have been *if*, given the country's actual industrial structure, in each industry the R&D intensity had been equal to the *typical* R&D intensity of the industry in question. By holding the R&D intensity in each industry constant in this way, this expression may reasonably be thought of as a measure of the effect of industrial structure on the R&D intensity in manufacturing. It is, of course, absolutely crucial here that the reference values termed *typical* actually *are* reasonably typical for each industry.

The second component is in essence a residual. If the *actual* R&D intensity in manufacturing of a given country is *higher* than what is 'predicted' by the industrial structure, this second component will be positive (and vice versa if the actual R&D intensity is lower than 'predicted'). This means that 'on average' the R&D intensity in each industry is *higher* (respectively, lower) than the typical R&D intensity. More precisely, the R&D intensity in each industry must be higher (lower) on *weighted* average, as can be seen from the formula for the second component, $\sum_{i=1}^{n}\left(I_i - \bar{I}_i\right)\cdot w_i$.

For each country, it gives the weighted sum over all industries of the difference between the R&D intensity of the country concerned and the typical (median) R&D intensity in each industry, when the weights that are attached to each industry are defined by the industry's share of total manufacturing value added of the country concerned. This second component may consequently be thought of as a rough and aggregate measure of the R&D intensity *inside* the individual industries, compared to other countries.

Thus, we have decomposed the R&D intensity in manufacturing in a given country into a sum of two components. The first component expresses the manufacturing R&D intensity we would expect from the industrial structure of the country, while the second expresses how high the R&D intensity in general is inside each individual industry

Now, in the expression $I_m = \sum_{i=1}^{n}\bar{I}_i \cdot w_i + \sum_{i=1}^{n}\left(I_i - \bar{I}_i\right)\cdot w_i$, the two components of the sum are not, so to speak, 'symmetrical'. The first component says what the overall R&D intensity would have been if the country in question had had the median industry-specific R&D intensity in each single industry; it gives a hypothetical overall R&D intensity value. The second component, by contrast, is defined as a weighted difference from an average. Let us now construct the formula in a more 'symmetric' way, by letting both components be expressed differences from an average.

To do this, we define the mean R&D intensity in the manufacturing sector across the 13 countries (for 1991, the mean value reported in Table 1, above), denoted by $\bar{I}_m$. This mean value we now subtract from both sides of equation (6) to get

$$I_m - \bar{I}_m = \sum_{i=1}^{n}\bar{I}_i \cdot w_i + \left[\sum_{i=1}^{n}\left(I_i - \bar{I}_i\right)\cdot w_i\right] - \bar{I}_m \qquad (7)$$

Rearranging, we get

$$I_m - \bar{I}_m = \left[\left(\sum_{i=1}^{n} \bar{I}_i \cdot w_i\right) - \bar{I}_m\right] + \left[\sum_{i=1}^{n}\left(I_i - \bar{I}_i\right)\cdot w_i\right] \qquad (8)$$

This expression takes the difference of the overall manufacturing R&D intensity of the country in question from the average overall manufacturing R&D intensity as the point of departure, and it roughly expresses this difference as a sum of, on the one hand, how much of this difference can be attributed to the industrial structure of the country (the expression in the first brackets), and, on the other hand, how much of this difference can be attributed to how high R&D intensities are within the different industries (the expression in the second brackets).

Let us now see what results get when we apply the above method of decomposition, expressing the R&D intensity in manufacturing of each country as sum of a component expressing the industrial structure of the country and a component expressing the R&D intensities inside the individual industries. Let us call these two components the *structure component* and the *industry intensity component*, respectively.

Table 12, below, shows this decomposition with the actual manufacturing R&D intensities as the point of departure, i.e., it is based on expression (6), above.

Table 12: R&D intensity in manufacturing, per cent (1), decomposed into sum of structure component (2) and industry intensity component (3), 1991.

Country	*R&D intensity in manufacturing (1)*	*Structure component (2)*	*Industry intensity component (3)*
Sweden	9.6	5.5	4.1
USA	8.6	6.9	1.6
Japan	7.1	7.1	0.0
France	6.6	6.0	0.5
UK	6.2	6.5	-0.3
Germany	6.0	6.7	-0.7
Finland	5.8	3.9	2.0
Netherlands	4.9	6.5	-1.6
Norway	4.7	3.8	0.9
Denmark	4.3	4.1	0.2
Canada	3.6	5.6	-2.0
Italy	3.2	4.4	-1.2
Australia	2.6	4.1	-1.5
Mean	5.6	5.5	0.2

Here, (1) = (2) + (3). Discrepancies are due to rounding error. Notice that the industry intensity component is *not* standardized in such a way as to make the mean across all countries equal to 0.

Let us now present the same results with differences from the average as the point of departure (expression 8, above). This is shown in Table 13, below.

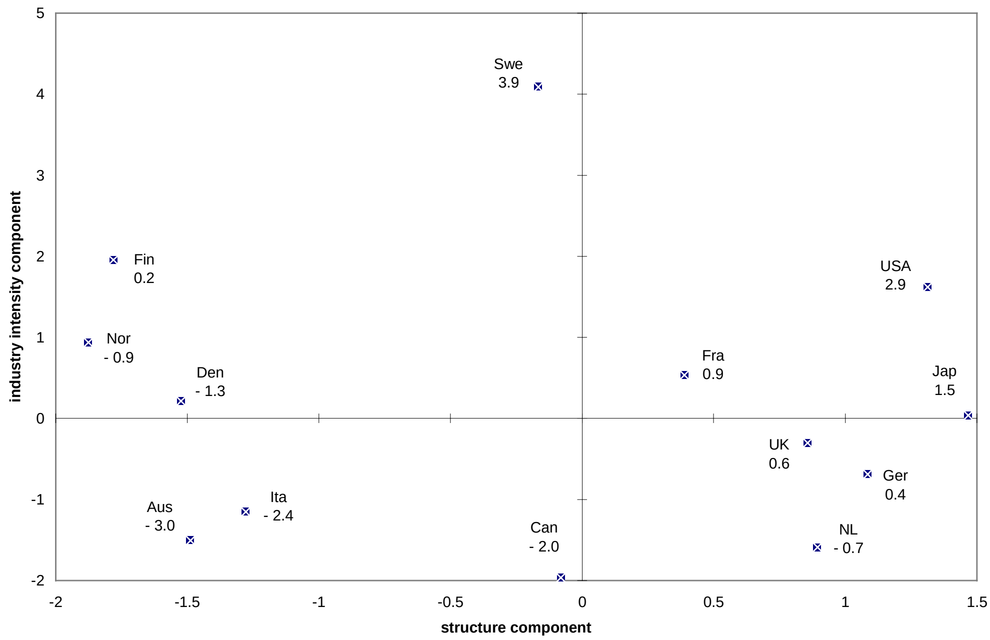
Table 13: R&D intensity in manufacturing, difference from average, percentage points (1), decomposed into sum of structure component (2) and industry intensity component (3), 1991.

Country	*R&D intensity in manu-facturing (1)*	*Structure component (2)*	*Industry intensity component (3)*
Sweden	3.9	-0.2	4.1
USA	2.9	1.3	1.6
Japan	1.5	1.5	0.0
France	0.9	0.4	0.5
UK	0.6	0.9	-0.3
Germany	0.4	1.1	-0.7
Finland	0.2	-1.8	2.0
Netherlands	-0.7	0.9	-1.6
Norway	-0.9	-1.9	0.9
Denmark	-1.3	-1.5	0.2
Canada	-2.0	-0.1	-2.0
Italy	-2.4	-1.3	-1.2
Australia	-3.0	-1.5	-1.5

To repeat, the information contained in this table is exactly the same as in Table 13, the only difference being that in Table 14 the average R&D intensity in manufacturing has been subtracted from both column 1 and column 2.

This decomposition, based on the difference from average expression of Table 14 and expression (8), is also shown graphically in Figure 13, below.

Figure 13: R&D intensity in manufacturing, difference from average, percentage points (figures in parenthesis), decomposed into sum of structure component (x-axis) and industry intensity component (y-axis), 1991.



In the figure, the structure component is represented along the x-axis and the industry intensity component along the y-axis, while the numbers in parentheses are the R&D intensities in manufacturing. Of course, if both component are above average, R&D intensity in manufacturing also has to be above average, and vice versa if both components are negative. If one component is positive and the other negative, the overall result depends on the which is the larger in absolute value.

We see that almost all combinations are present, and that there is very little correlation between the two components. In fact, the correlation coefficient is - 0.08, i.e. in practice 0. Three countries are above average on both components, namely USA, France and Japan. Three countries are above average on the structure component but below average on the industry intensity component, namely Germany, UK and the Netherlands. Three countries are below average on both components, namely Canada, Italy and Australia. Lastly, in the quadrant defined by below average on the structure component but above average on the industry intensity component we find the four Nordic countries, Sweden, Finland, Norway and Denmark.

By comparing, for instance, Sweden, Japan and the USA, we can also see how, in this case, a high intensity in manufacturing may reflect different combinations of the structure and intensity components. In the Swedish case, the industrial structure does not indicate any particularly high R&D intensity in manufacturing, rather the contrary. The high R&D intensity in Swedish manufacturing is thus solely due to very high intensities inside the individual industries. For Japan it is the other way around, with the intensity inside the individual industries in general being only on the average, but with an industrial structure which very much indicates a high R&D intensity in manufacturing as a whole. The USA is high on *both* components.

Lastly, let us look a bit closer at the example with the difference between the USA and Norway above and verify the impression we got when we looked at the diagrams representing R&D intensity and share of manufacturing production of each industry in the two countries. We have seen that the difference in R&D intensity in manufacturing between the two countries was 3.9 percentage points (with the USA, of course, having the higher intensity). We can now verify that most of this difference, namely 3.2 percentage points, is due to the difference in industrial structure, while 0.7 percentage points is due to the USA on average having higher R&D intensity inside the individual industries.

Let us now look at the relationship between size of economy and each of the two components separately.

We start with the industry intensity component. In Table 14, below, we see the mean and median industry intensity component value for large and for small economies, as defined in Table 3, above.

Table 14: Mean and median industry intensity component value, large and small economies, 1991.

	Mean	*Median*
Large	0.0	-0.1
Small	0.3	0.2

It is in fact the small economies who have the higher value on the industry intensity component, both mean and median. The difference between large and small economies appears very small however. This is confirmed by an analysis of variance, reported in Table 15, below.

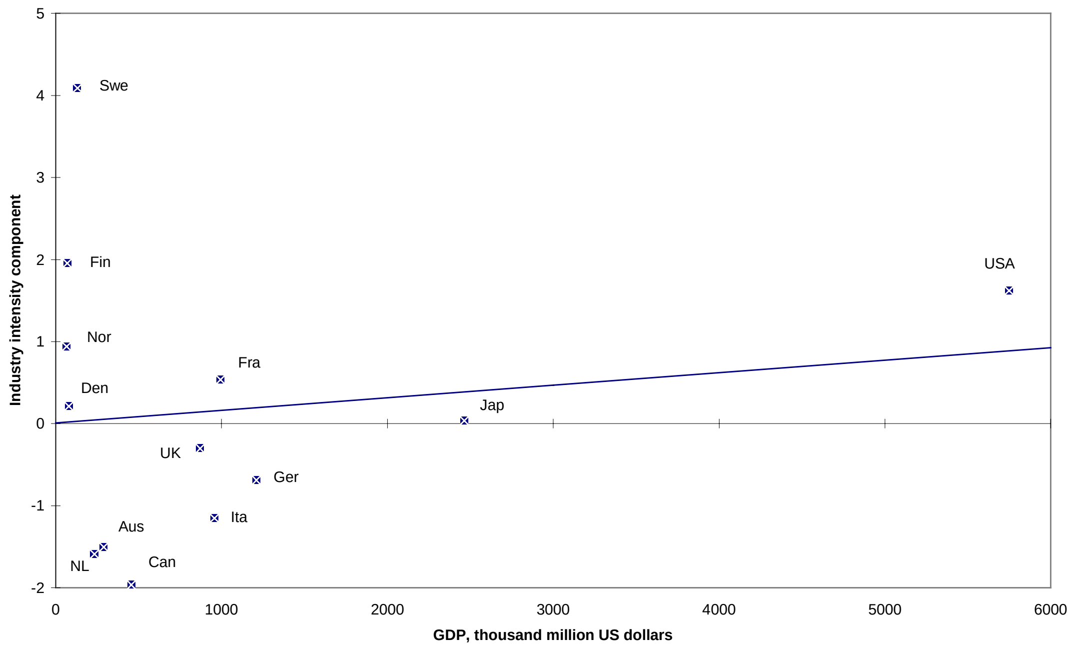
Table 15: Analysis of variance. Industry intensity component for large and small economies, 1991.

Source of variation	*Sum of squares*	*DF*	*Mean square*	*F*	*P-value*
Between	0.285281	1	0.285281	0.091514	0.767904
Within	34.29069	11	3.117335		
Total	34.57597	12			
R^2	0.01				

We see that the difference between the groups is not significant (F is very low, the p-value is 0.768), and R^2 is only 0.01.

Let us see what the relationship looks like if we use economy size as a continuous variable, measured by GDP. The relationship between GDP and the industry intensity component is shown in Figure 14, below.

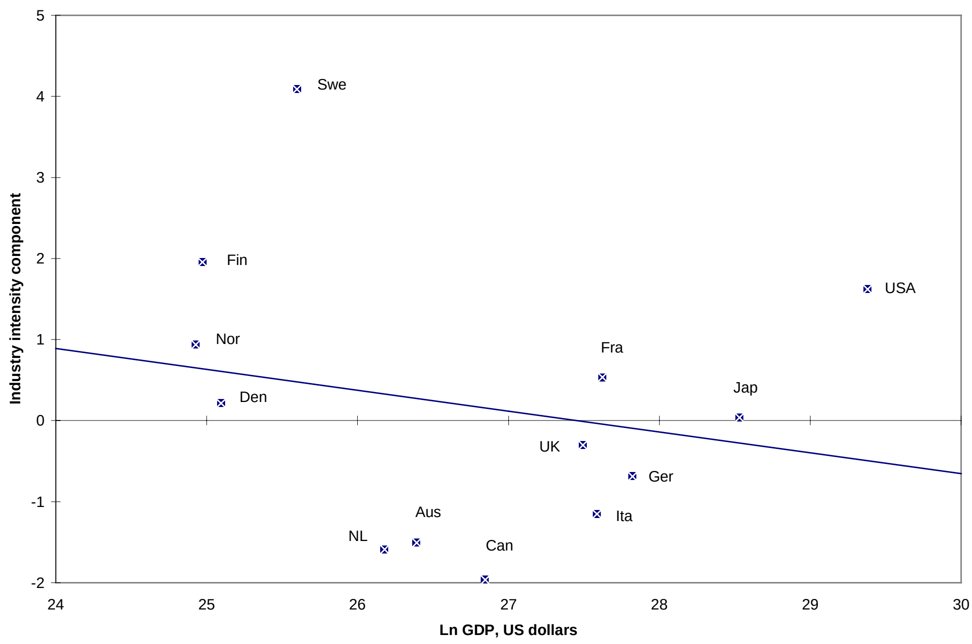
Figure 14: GDP, thousand million US dollars (x-axis), and industry intensity component value (y-axis), 1991.



The figure shows a very slight, not significant association between the two variables. We have $r = 0.14$ and $R^2 = 0.02$. However, again the outlier USA seems to influence the correlation heavily.

Let us see what result we get when we use the alternative measure of economy size, namely Ln GDP. Figure 15, below, shows the relationship between Ln GDP and the industry intensity component.

Figure 15: Ln GDP, US dollars (x-axis), and industry intensity component value (y-axis), 1991.



With this measure of economy size, the correlation gets slightly negative. We have r = - 0.22 and $R^2 = 0.05$. The relationship is *not* significant at any conventional level (not significant at 10 per cent level).

In conclusion, we do *not* find any relationship between economy size and the industry intensity component. Thus, it does not seem that large economies in general have higher R&D intensities inside the individual industries than do small economies.

We now turn to the relationship between economy size and the *structure component*. In Table 16, below, the mean and median structure component value for large and small economies are shown.

Table 16: Mean and median structure component value, large and small economies, 1991.

	Mean	*Median*
Large	6.3	6.6
Small	4.8	4.1

Here the large economies have higher both mean and median value than the small economies, and the differences here appear to be more substantial. The difference in mean value is 1.5 percentage points, in median value 2.5 percentage points.

The impression that the differences here are more substantial is confirmed by the analysis of variance reported in Table 17, below (which relates to the difference in *means*, of course).

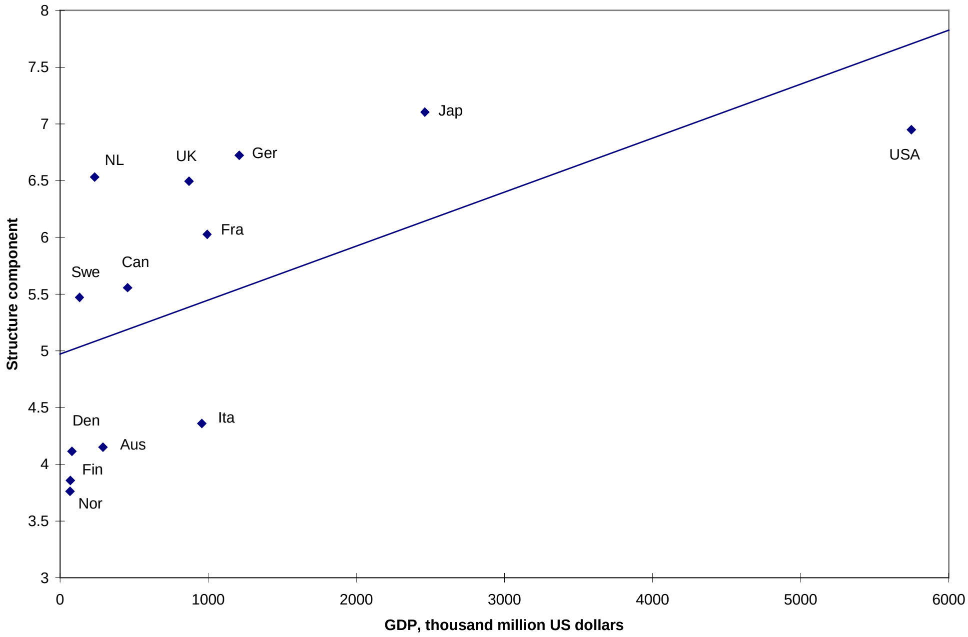
Table 17: Analysis of variance. Structure component for large and small economies, 1991.

Source of variation	*Sum of squares*	*DF*	*Mean square*	*F*	*P-value*
Between	7.255013	1	7.255013	6.648222	0.025663
Within	12.00398	11	1.091271		
Total	19.259	12			
R^2	0.38				

We see that $R^2 = 0.38$, and that the difference between the groups is significant at the 5 per cent level (p-value is 0.0257).

Next, we look at the relationship between economy size and the structure component when economy size is measured continuously. Figure 16, below, shows the relationship between GDP and the structure component.

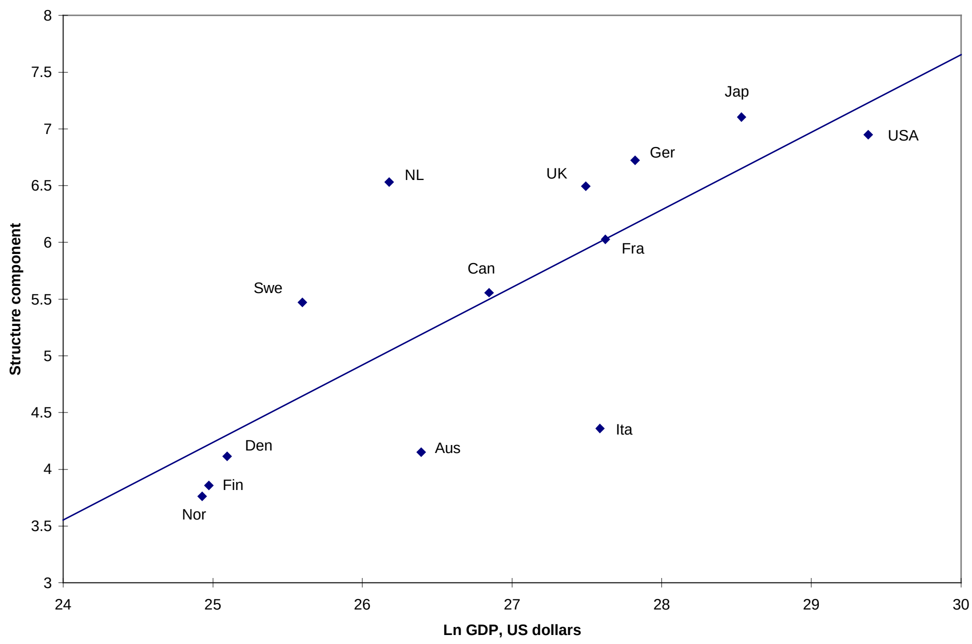
Figure 16: GDP, thousand million US dollars (x-axis), and structure component value (y-axis), 1991.



Also here we get a clear positive association between economy size and the structure component. The correlation coefficient is 0.59, R^2 is 0.35. This relationship is significant at the 5 per cent level.

Let us lastly look at the association between economy size and the structure component when economy size is measured by Ln GDP instead of GDP. This relationship is shown in Figure 17, below.

Figure 17: Ln GDP, US dollars (x-axis), and structure component value (y-axis), 1991.



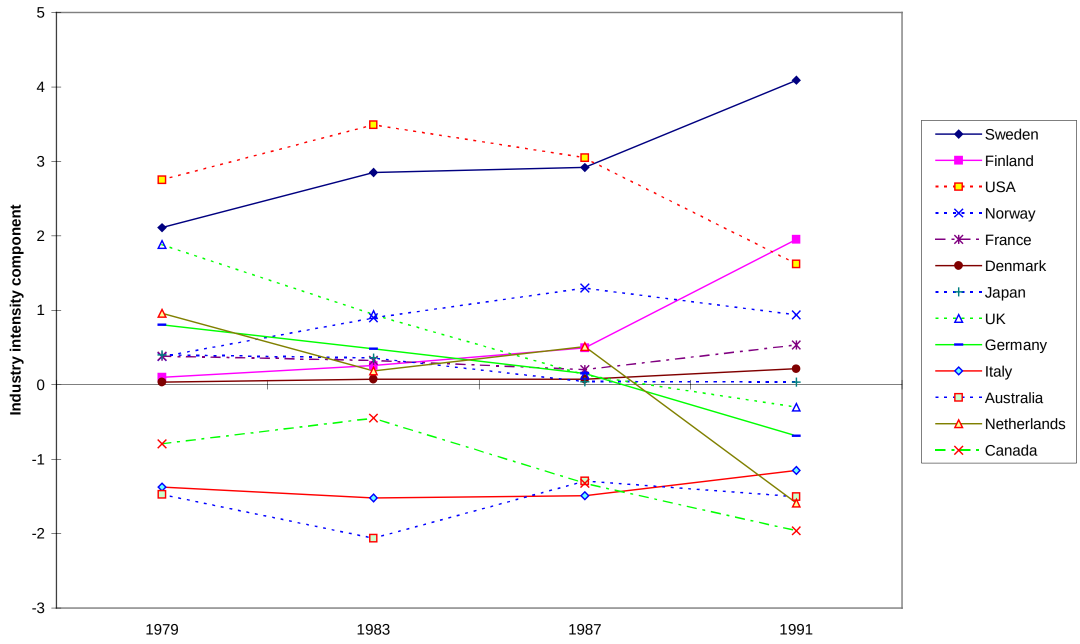
Here, the positive association is even clearer than in the case where GDP itself was used to measure economy size. The correlation coefficient is quite high, 0.77, with $R^2 = 0.59$. This relationship is statistically significant at the 1 per cent level.

Thus, to summarize, we find a very clear association between economy size and the structure component. The larger the economy, the more the industrial structure is favourable to a high R&D intensity in the manufacturing sector as a whole. Put another way, the larger the economy, the higher the share of industries which typically have relatively high R&D intensity, the lower the share of industries which typically have relatively low R&D intensity.

Let us now see if we get roughly the same results when we look at the earlier years three years singled out, namely 1979, 1983 and 1987. Before going on to looking at the relationship between economy size and each of the two components in these earlier years, we will first look at the evolution in the values for each of the countries on the two components for the 13 countries.

We start with the evolution of country values on the industry intensity component. These values are shown in Figure 18, below.

Figure 18: Industry intensity component values, 1979, 1983, 1987 and 1991.

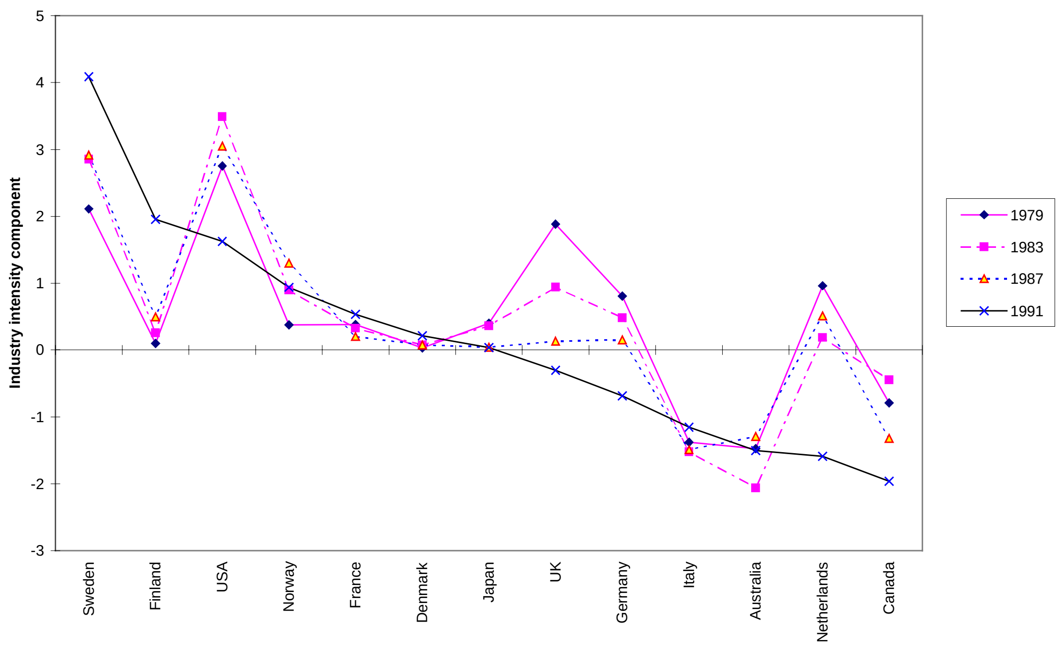


We see that the change over time on the industry intensity components varies across countries. Especially, there seems to be variation in the rate of change across countries from 1987 to 1991.

Notice that the basis for the decomposition in each year are the typical (median) R&D intensities in each industry *in the year concerned*. This means that overall we should not expect any trend either upwards or downwards in these component values: when some countries increase their value, there must be a roughly corresponding decrease for other countries. This means that we have, for instance, *not* used the typical R&D intensities in one single year, or the mean across several years, or the like, as the basis for the comparison. If we had proceeded in this way, we would have expected a general, namely upwards, trend in the values of the industry intensity component over time. Then we could have looked at the change in R&D intensity over time for each of the countries (which is in general positive, cf. Figure 3, above) and decomposed this change into what is due to change in industrial structure, on the one hand, and what is due to change in R&D intensity inside each industry, on the other. This kind of analysis may be the subject of a later paper.

Let us present the information from Figure 17 on the development over time in the industry intensity component value for each country in another form, as profiles across countries in each of the four years, like we did for R&D intensity in manufacturing in Figure 4, above. These profiles are shown in Figure 19, below.

Figure 19: Industry intensity component values, 1979, 1983, 1987 and 1991.



This chart of industry intensity component profiles across countries for different years gives much the same impression the corresponding chart of R&D intensity in manufacturing profiles across countries in Figure 4, above. As we would expect from the presentation in Figure 18, above, the 1991 profile seems relatively deviant compared to the profiles of the three earlier years, which look more similar. This impression is confirmed by the correlation matrix presented in Table 18, below.

Table 18: Correlation matrix of cross-country industry intensity component profiles between 1979, 1983, 1987 and 1991.

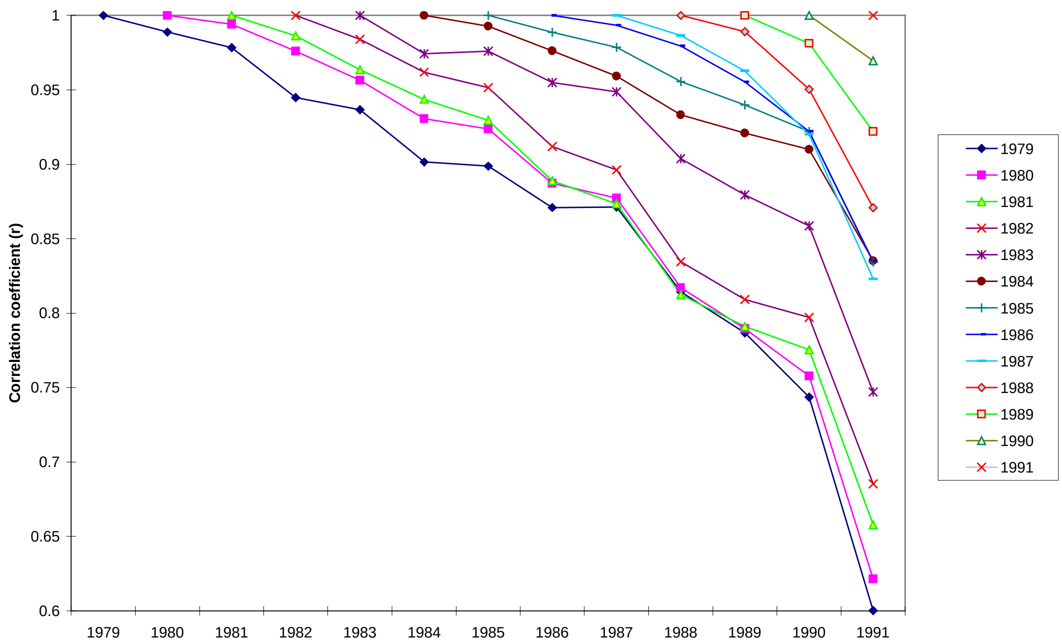
Year	1979	1983	1987	1991
1979	1			
1983	0.94	1		
1987	0.87	0.95	1	
1991	0.60	0.75	0.82	1

We see that while the correlation of these profiles between 1979 and 1983 is 0.94 and the correlation between 1983 and 1987 is 0.95, the correlation between 1987 and 1991 is only 0.82. Likewise, while the correlation between 1979 and 1987 is 0.87, the correlation between 1983 and 1991 is only 0.75. Indeed, the correlation between 1979 and 1987 (8 years) is higher than the correlation between 1987 and 1991 (4 years), 0.87 and 0.82, respectively.

The second thing to notice here is that compared to the corresponding correlation matrix concerning R&D intensities in manufacturing presented in Table 6, above, the correlation coefficients we are dealing with here are noticeably lower. Thus, while in the R&D intensity in manufacturing case the correlation between 1979 and 1991 is 0.82, in the industry intensity component case it is only 0.60. Similarly, all other corresponding correlation coefficients are lower in the industry intensity component case than in the R&D intensity in manufacturing case. Thus, the relative values among countries change more rapidly on the industry intensity component than for actual R&D intensity in manufacturing.

Lastly, let us look at the development of the correlation coefficient year by year, as we did in the R&D intensity in manufacturing case in Figure 5, above. The corresponding figure for the industry intensity component is Figure 20, below.

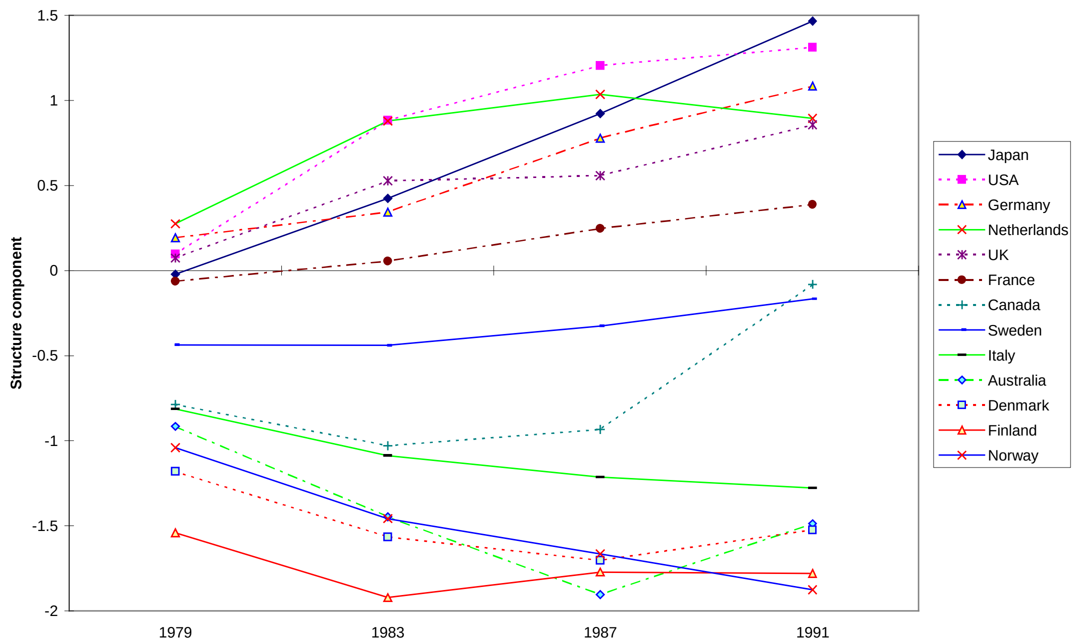
Figure 20: Correlation matrix of cross-country industry intensity component profiles between all years in the period 1979-1991. Graphical presentation.



Here we get a clear impression that the rate of change accelerates after 1987. It may also seem that, in this case more than in the R&D intensity in manufacturing case, the changes are especially large between the two last years, 1990 and 1991. We will not go further into this question here. Notice that the scale along the x-axis is not the same as in Figure 5. There the minimum value represented was 0.80, here it is 0.60.

Let us now turn to a corresponding presentation of the development of the *structure component* values for the different countries. These values are shown for the years 1979, 1983, 1987 and 1991 in Figure 21, below.

Figure 21: Structure component values, difference from average version, 1979, 1983, 1987 and 1991.

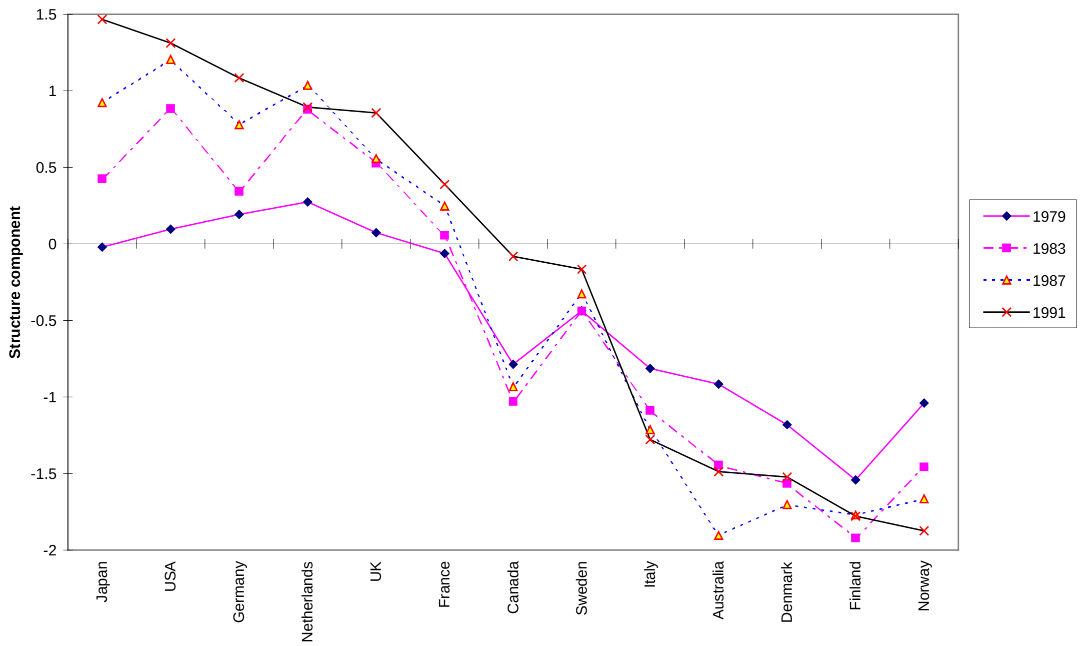


I have chosen to use the difference from average version of the structure component (expression 8, above) to emphasize that the development over time which it is most meaningful to look at here is the relative development among countries. This again is connected to the fact that here we use the typical R&D intensities in each industry in the year concerned as the standard of reference. If we had used the expected R&D intensity in manufacturing version of the structure component (expression 6), we would have seen that the structure component was increasing over time quite generally for all countries. This might have created the impression that there was a uniform development over time in all countries for the industrial structure to change in the direction of a higher share of production being accounted for by industries with relatively high R&D intensity and a lower share by industries with relatively low R&D intensity. This impression would not be correct, however, because included in the increase in the expected R&D intensity in manufacturing version of the structure component value is not only the change in industrial structure but also the fact that the typical (median) R&D intensities in the individual industries in general have tended to increase over time, i.e. it partly reflects that the standard of reference itself changes over time. If we were to say something about structural change over time we would, again, have to use the typical values of *one single year* (or an average of several years, or the like) as a standard of measurement. Again, this may be the subject of a later paper.

Remember again that the industry intensity component, and therefore also the difference from average version of the structure component, is not standardized in such a way that its mean across all countries equals 0. This is evident in Figure 20, especially for 1979.

The impression we get from Figure 21 is that the variation across countries in the rate of change in the structure component value is less than in the case of both the industry intensity component and in the R&D intensity in manufacturing itself. This impression we also get from the profile version of the same information, presented in Figure 22, below.

Figure 22: Structure component values, difference from average version, 1979, 1983, 1987 and 1991.



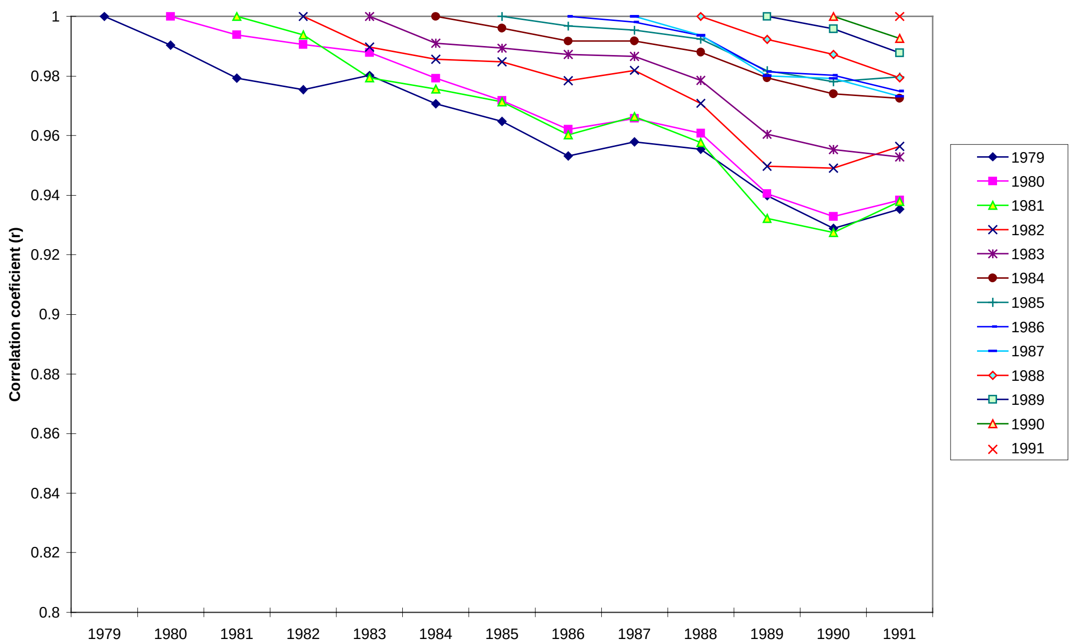
This impression is confirmed by the correlation matrix where the profiles across countries of each year are correlated with each other. This matrix is presented in Table 19, below.

Table 19: Correlation matrix of cross-country structure component profiles between 1979, 1983, 1987 and 1991.

Year	*1979*	*1983*	*1987*	*1991*
1979	1			
1983	0.98	1		
1987	0.96	0.99	1	
1991	0.94	0.95	0.97	1

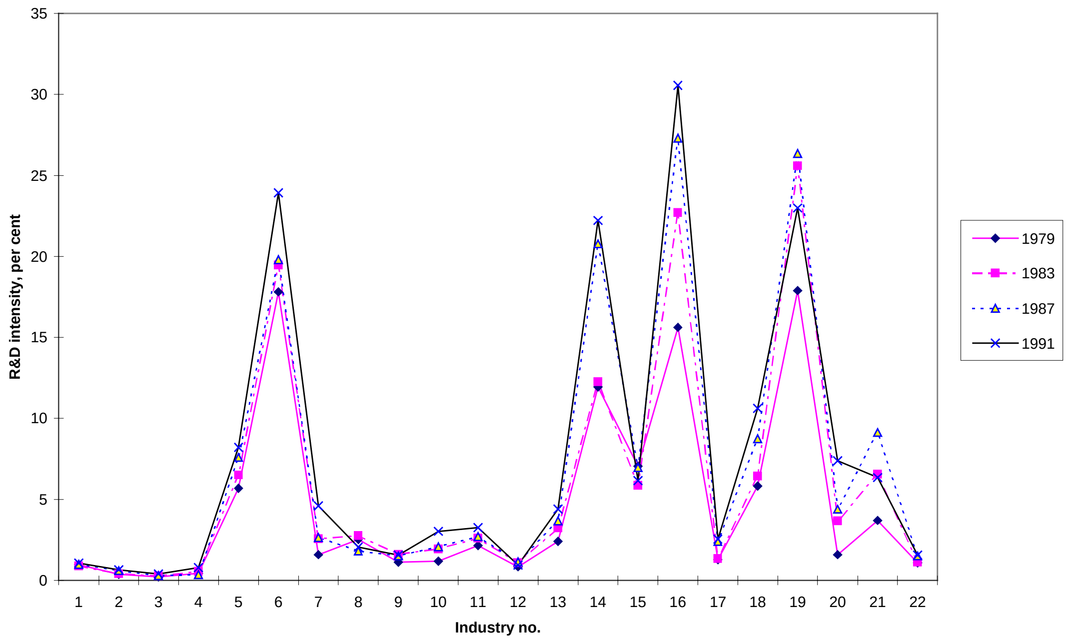
We see that the correlation coefficients are considerably higher in the structure component case than in both the industry intensity component and the R&D intensity in manufacturing cases. Whereas for the two former cases the correlation between 1979 and 1991 was 0.60 and 0.82, respectively, here it is as high as 0.94. Furthermore, in this case we find no evidence of any acceleration of change after 1987. This is further confirmed by inspection of the correlation matrix involving all the years in the period 1979-1991, presented graphically in Figure 23, below.

Figure 23: Correlation matrix of cross-country structure component profiles between all years in the period 1979-1991. Graphical presentation.



Thus, the profile across countries of the structure component value is much more stable than the profile of the industry intensity component value. That a measure describing industrial structure should show far more stability than a measure describing R&D expenditures is, of course, not surprising, so this result seems very reasonable indeed. However, to confirm this impression, let us just check whether the standard of measurement itself, the typical (median) R&D intensities in the individual industries, have a reasonable degree of stability over time. Figure 24, below, shows the typical (median across all countries) R&D intensity in each industry in 1979, 1983, 1987 and 1991.

Figure 24: Typical (median across all countries) R&D intensity in each industry (per cent) in 1979, 1983, 1987 and 1991.



These profiles look very similar. We also see that there seems to be a tendency for the typical R&D intensities in the individual industries to increase over time.

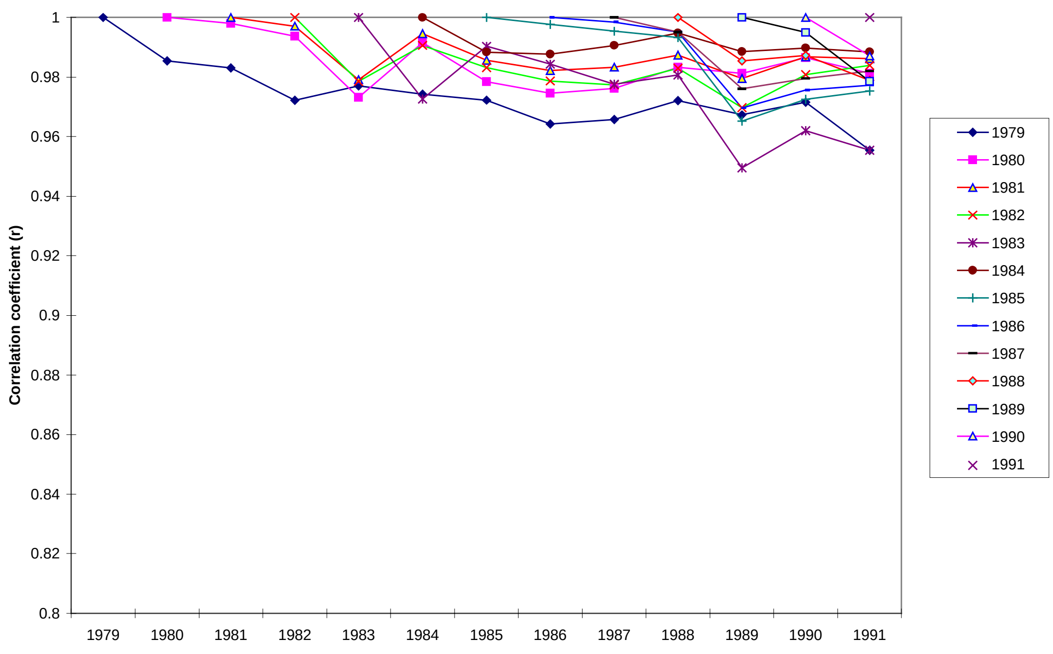
The correlation matrix presented in Table 20, below, confirms the impression of the stability over time of the standard of measurement.

Table 20: Correlation matrix of inter-industry typical (median) R&D intensity profiles between 1979, 1983, 1987 and 1991.

Year	1979	1983	1987	1991
1979	1			
1983	0.98	1		
1987	0.97	0.98	1	
1991	0.96	0.96	0.98	1

We see that the lowest correlation here, between 1979 and 1991, is as high as 0.96. The stability also emerges from the graphical presentation of the correlation matrix involving all the years in the period 1979-1991 in Figure 25, below.

Figure 25: Correlation matrix of inter-industry typical (median) R&D intensity profiles between all years in the period 1979-1991. Graphical presentation.



Indeed, from this figure it is not even evident that there is any clear tendency for any change in the profile over time at all, apart from random variation from one year to the other.

Having looked at these developments over time, let us now see if we find the same kinds of relationship between economy size and the two components of manufacturing R&D intensity in 1979, 1983 and 1987 as we did in 1991.

We start with the industry intensity component. First we look at the contrast between large and small economies (as defined in Table 3, above). In Table 21, below, we show mean values of the industry intensity component for large and small economies and the difference between these mean values, as well as R^2, p-value and statistical significance from ANOVA, for all the four years.

Table 21: Mean industry intensity component value, large and small economies, difference between large and small economies. R^2, p-value and statistical significance from ANOVA. 1979, 1983, 1987 and 1991.

Year	Mean large economies	Mean small economies	Difference	R^2	p-value	statistical significance
1979	0.8	0.2	0.6	0.06	0.4031	not significant
1983	0.7	0.3	0.4	0.02	0.6270	not significant
1987	0.3	0.4	0.0	0.00	0.9661	not significant
1991	0.0	0.3	-0.3	0.01	0.7680	not significant

As we see, we here get the same result for the three earlier years as for 1991. In all cases the difference between large and small economies in mean industry intensity component value is very small and far from statistically significant

Turning to a continuous measure of economy size, the correlation coefficient between GDP and industry intensity component value, with R^2 and statistical significance, is shown for all four years in Table 22, below.

Table 22: Statistical association between GDP and industry intensity component, 1979, 1983, 1987 and 1991. Correlation coefficient (r), R^2 and statistical significance.

Year	r	R^2	statistical significance
1979	0.51	0.26	.10 level
1983	0.54	0.29	.10 level
1987	0.44	0.19	not significant
1991	0.14	0.02	not significant

Again, we find no clear association between GDP and industry intensity component, even though *r* for the earlier years is not as low as for 1991. For the two earliest years the relationship is significant at the 10 per cent level, but the relationship is again probably heavily dependent on the influential USA observation. Thus suspicion is strengthened when we look at the corresponding correlations between *Ln GDP* and industry intensity component value for each of the four years, shown in Table 23, below.

*Table 23: Statistical association between Ln GDP and industry intensity component, 1979, 1983, 1987 and 1991. Correlation coefficient (*r*), R^2 and statistical significance.*

Year	*r*	R^2	*statistical significance*
1979	0.28	0.08	not significant
1983	0.20	0.04	not significant
1987	0.02	0.00	not significant
1991	- 0.22	0.05	not significant

Here the lack of association is very clear. Correlation coefficients are very small for all four years, and not statistically significant.

Thus, we conclude that we find no clear relationship between economy size and industry intensity component. The data do not support any claim that large economies tend to have higher R&D intensity inside each individual industry than small economies, nor is there any evidence for the opposite hypothesis.

Next, let us turn to the *structure component*. In the same way as for the industry intensity component above, Table 24, below, shows the mean value of the structure component for large and small economies and the difference between these mean values, as well as R^2, p-value and statistical significance from ANOVA, for all the four years.

Table 24: Mean structure component value, large and small economies, difference between large and small economies. R^2, p-value and statistical significance from ANOVA. 1979, 1983, 1987 and 1991.

Year	Mean large economies	Mean small economies	Difference	R^2	p-value	statistical significance
1979	3.4	2.7	0.7	0.38	0.0255	.05 level
1983	4.7	3.6	1.2	0.37	0.0272	.05 level
1987	5.8	4.3	1.5	0.39	0.0224	.05 level
1991	6.3	4.8	1.5	0.38	0.0257	.05 level

Here we get the same results in the three earlier years as in 1991. There is a clear tendency for the structure component value to be higher among the large economies than among the small, the difference in mean values being significant at the 5 per cent level for all years. R^2 is practically the same in all four cases, varying between 0.37 and 0.39.

Turning to a continuous measure of economy size, Table 25, below, shows the correlation between GDP and the structure component for the four years.

Table 25: Statistical association between GDP and structure component, 1979, 1983, 1987 and 1991. Correlation coefficient (r), R^2 and statistical significance.

Year	r	R^2	statistical significance
1979	0.48	0.23	.10 level
1983	0.57	0.33	.05 level
1987	0.58	0.34	.05 level
1991	0.59	0.35	.05 level

Also in this case we get roughly the same results for the earlier years as for 1991. The association is somewhat weaker in 1979 than in the other years, where the correlation coefficient is only 0.48 and significant only at the 10 per cent level. For the three later years it is between 0.57 and 0.59, and significant at the 5 per cent level. However, in the GDP case there is the outlier problem, so let us again also look at the relationship between *Ln GDP* and the structure component. This is shown in Table 26, below.

*Table 26: Statistical association between Ln GDP and structure component, 1979, 1983, 1987 and 1991. Correlation coefficient (*r*), R^2 and statistical significance.*

Year	*r*	R^2	*statistical significance*
1979	0.72	0.51	.01 level
1983	0.72	0.52	.01 level
1987	0.72	0.52	.01 level
1991	0.77	0.59	.01 level

Here, the positive association between economy size and the structure component is very clear. The correlation coefficient is above 0.70 for all years, varying between 0.72 and 0.77. In all four cases it is significant at the 1 per cent level.

In conclusion, we find a very clear relationship between economy size and the structure component. The variation in industrial structure across countries is such that the larger the absolute size of the national economy, the higher the R&D intensity in manufacturing which we would expect given the industrial structure. Or, in other words, the larger the absolute size of the economy, the higher the share of production accounted for by industries which typically have relatively high R&D intensity and the lower the share of production accounted for by industries which typically have relatively low R&D intensity.

THE QUESTION OF PERFORMANCE

I will now turn to an explicit consideration of an issue which is often closely connected to discussions of R&D expenditures and R&D intensities, whether this issue is explicitly stated or simply implicitly assumed, namely the question of *performance*.

The level of R&D expenditures relative to production, i.e. R&D intensity, is quite generally thought of as saying something about innovativeness and competence of the firms in question and their personnel, quality and sophistication of production, etc. Now, of course, R&D intensity is an imperfect and partial measure of innovativeness. For one thing, it is an input measure, saying something about effort regarding innovation, sophistication, etc., but not saying what comes out of this effort in terms of new products and processes, etc. Furthermore, there are other types of innovation costs in addition to R&D costs which may be important in many cases, such as costs connected to product design, acquisition of licences, trial production, market analysis, etc. Taking also these other costs into account may give a different picture of differences in innovativeness among firms and industries than if one exclusively focuses on R&D costs. Innovative capability in a firm or an industry or region or even nation may also to a large extent reside in skills an capabilities not closely connected to any specific type of costs, in organizational and institutional characteristics, etc.

Nevertheless, R&D expenditures, or more precisely R&D intensity, is much used as an indicator of innovativeness etc. When one does this, one should of course be aware of and take into account that it is an imperfect and partial indicator.

When one compares countries in terms of R&D intensities, it is common simply to use R&D expenditures in a given year in proportion to GDP. In the present context, where we only have detailed data on the manufacturing sector, the corresponding measure to R&D expenditures as a proportion of GDP is simply the R&D intensity of the manufacturing sector as a whole, i.e. total R&D expenditures as a proportion of total value added in manufacturing.

However, we have seen above that the R&D intensity in manufacturing as a whole in a country may be understood as the result of the values of two quite distinct components. (Indeed, we saw that for our 1991 data the two components were not even correlated.) A given R&D intensity in manufacturing may express a wide variety of combinations of values on the two components. In the following we will discuss the issue of *performance* on the background of this decomposition.

From one perspective the *structure component* may be thought of as an indicator of performance in relation to innovativeness and competence, the sophistication of production, etc. This will be so if one thinks that it is important for a country to engage substantially in the type of production characterized by high R&D intensity, to restructure the economy towards the high R&D intensity industries. For instance, Charles Edquist and Bengt-Åke Lundvall adopt this perspective in an analysis of the

Danish and Swedish systems of innovation.[4] Discussing the *performance* of these innovation systems, they find that ‘both Denmark and Sweden have a relatively weak position in R&D-intensive products.’[5] They note that this is not so surprising in the Danish case, given Denmark’s ‘weak R&D effort.’ However, noting Sweden’s ‘very substantial investments in R&D,’ as well as ‘its high number of patents per million inhabitants in the United States, and its strong multinationals in engineering’, they find it remarkable that Sweden ‘has been so slow in absorbing R&D-intensive products.’[6] One of their main conclusions is that ‘the average low-R&D character of Swedish production is a severe problem for the Swedish system of technological change.’[7] Here they very clearly distinguish between an effort component and a structure component: in spite of a very substantial R&D effort, Sweden has a ‘low-R&D *character of production*.’ An indicator they explicitly use for this structural dimension is the share of production (and exports) accounted for by industries defined as having high R&D intensity. The structure component used in the present paper is a more generalized representation of this idea (of course, the question of exports is not treated here). Indeed, we have seen that while the R&D effort inside each industry in Sweden generally is very strong (the industry intensity component), the industrial structure as such, holding everything else equal, indicates a less than average R&D intensity in manufacturing (the structure component). It is this structural dimension which is one of the main issues Edquist and Lundvall emphasize in their discussion of the *performance* of the Danish and Swedish innovation systems.

From an alternative or even opposite perspective, the dimension of performance which one has in mind would be one which the *industry intensity component,* rather than the *structure component,* says something about. In this perspective the industrial structure of the country in question is taken as given, and then one asks how well the country performs in terms of R&D effort given the industrial structure that it actually

4 Charles Edquist and Bengt-Åke Lundvall, ‘Comparing the Danish and Swedish Systems of Innovation’, in Richard R. Nelson (ed.), National Innovation Systems. A Comparative Analysis, New York: Oxford University Press, 1993, pp. 265-298.

5 ibid., p. 287.

6 ibid.

7 ibid., p. 290.

has. This is the perspective underlying, for instance, OECD's 'STIBERD' indicator,[8] to be discussed below. This perspective may be *opposite* from the one just discussed, in that one explicitly holds the industrial structure in question to be by and large rational, given the resources and preconditions of the country. But the two perspectives do not necessarily preclude one another. One may believe that a given country *should* change its industrial structure in the direction of a higher share of production accounted for by high R&D intensity industries, and still be interested in how well the country on average performs in terms of R&D expenditures given the industrial structure it actually has at the present. Thus, these perspectives may also be complementary.

The industry intensity component adjusts R&D intensity in manufacturing for industrial structure in a quite straightforward manner. As pointed out above, it is in essence a residual. The *structure component* says how high R&D intensity in manufacturing we would have expected in a given country *given* its industrial structure and *if* in each industry it had the *typical* (median) R&D intensity. The *industry intensity component* is simply the difference between the actual R&D intensity of the country and its structure component. If this difference is positive, the R&D intensity inside the individual industries must on average have been higher than the typical R&D intensity, and vice versa if the difference is negative. Thus, the industry intensity component seems to adjust R&D intensity in manufacturing for industrial structure in a quite understandable and reasonable way.

However, if R&D intensity inside individual industries is to be considered from the point of view of *performance* relative to innovation, etc., there is another aspect, in addition to adjusting for industrial structure in the above way, which should be taken into account. A high industry intensity component value means that on *weighted* average the country in question has an R&D intensity inside the individual industries which is substantially higher than the typical. However, as applies to all averages, even weighted ones, this average may be an average of very different values. Thus, the industry intensity component does not take into account the *distribution* of R&D expenditures across industries. For instance, a high industry component value may

8 Cf. Manufacturing Performance: A Scoreboard of Indicators, OECD Documents, Paris: OECD, 1994, pp. 51-57.

express particularly high R&D intensities in a few high R&D intensity industries which together account for a modest share of total manufacturing production in the country in question, while industries which together account for the major share of manufacturing production have quite low R&D intensities compared to other countries. If the distribution of R&D resources across industries in a country is very skewed, i.e. substantially more skewed than what is normal, only adjusting for industrial structure as does the industry intensity component will not give an accurate picture of how well the industries in a given country in general perform in terms of R&D intensity. Thus, if we want a measure of how well the industries in a country in general perform in terms of R&D intensity, we should take into account *both* the industrial structure *and* the distribution of R&D expenditures across industries. The industry intensity component takes account only of industrial structure, not of the distribution of R&D expenditures across industries.

We will bring this distribution of R&D expenditures across industries into the discussion. We will first look at an indicator proposed by the OECD, the STIBERD indicator.

The OECD STIBERD indicator

The reason for discussing the OECD STIBERD indicator is that it is presented as adjusting R&D intensity in manufacturing for both industrial structure and the allocation of R&D expenditures across industries. We will, however, show that it does not in fact do this, and that it has some quite undesirable properties. Nevertheless, a discussion of this attempt at controlling for both industrial structure and the allocation of R&D intensity across industries may illuminate some of the problems involved here.

STIBERD stands for *Structurally Adjusted Business Enterprise R&D Intensity*.[9] It is presented as an indicator which is 'used to examine R&D performance' and which makes 'adjustments for a country's industrial structure and thus places the countries

9 ibid., p. 51.

on a more comparative basis,'[10] i.e. relative to just comparing the R&D intensity of the manufacturing sector as a whole across different countries.

The definition of STIBERD is that it is 'calculated as the sum of the R&D intensities of each sector of a country divided by the average R&D intensity of the sector for the 12 OECD countries, weighted by the value added of the sector for the country divided by the value added for total manufacturing for the country.'[11] (The 12 OECD countries of course refer to the countries compared in this particular OECD document.) In the notation used earlier in this paper, we thus have

$$STIBERD = \sum_{i=1}^{n} \frac{I_i}{\bar{I}_i} \cdot \frac{V_i}{V_m}$$

or

$$STIBERD = \sum_{i=1}^{n} \frac{I_i}{\bar{I}_i} \cdot w_i$$

The reason the OECD gives for introducing the STIBERD indicator is precisely the inadequacy of using the R&D intensity of the manufacturing sector as a whole as a measure of R&D performance. This measure is, according to the OECD, 'an imperfect one, as it takes no account of the differences in industrial structure from one country to another. If the natural endowments of a specific country are oriented towards natural resource and other low technology industries, it will probably have a low R&D intensity. This intensity, however, may simply reflect its industrial structure. In addition, if the high technology sectors in a country account for a smaller than average share of the country's output, its R&D intensity is also likely to be relatively low. Nevertheless, this does not necessarily mean that the country is less technologically competitive. Its individual industries may have relatively high levels of R&D expenditure per unit of output compared to its competitors, but the industries themselves may not be characterized by heavy R&D expenditures. The use of R&D intensities to rank technological performance thus tends to be biased against

10 ibid.
11 ibid., p. 52

countries with favourable natural resource endowments and a comparative advantage in the low technology industries.'[12]

From the above quotation it is evident that the STIBERD indicator is meant to adjust manufacturing R&D intensity for industrial structure. However, it seems equally evident that STIBERD is meant to adjust for the allocation of R&D expenditures across industries as well. According to the OECD, STIBERD 'addresses the issue of whether the individual industries in a country, particularly those accounting for large value added shares compare favourably against the same industries in other countries.'[13]

To explain how this comparison works, the OECD sketches a hypothetical example: 'For example, assume that country A has R&D expenditures which are highly concentrated in a few high technology industries that do not account for a large share of value added while on the other hand country B has R&D expenditures which are concentrated in the low technology industries but allocated across these industries more in line with their shares of output. Country B will have a higher STIBERD than country A because it is performing relatively more R&D in the industries which are important in its economy, despite the fact that its R&D intensity when measured in the standard fashion would probably be lower than country A's.'[14]

Let us now look closer at the STIBERD indicator. As we saw, it is defined as

$$STIBERD = \sum_{i=1}^{n} \frac{I_i}{\bar{I}_i} \cdot w_i$$

Let us compare this to the *industry intensity component* defined above, which is given by

$$IIC = \sum_{i=1}^{n} \left(I_i - \bar{I}_i \right) \cdot w_i$$

---

12 ibid., p. 51.
13 ibid. p. 53.
14 ibid. p. 54.

where *industry intensity component* is abridged to *IIC* (this is the second part of the right hand side of equation (6), above).

In both STIBERD and the industry intensity component, the *typical* R&D intensity of each industry, denoted $\bar{I}_i$, is used as a standard of reference. In the STIBERD indicator, this is the *average* R&D intensity of each industry across all the countries included. The average would also have been used in the decomposition performed in the present paper, had it not been for the presence of a couple of extreme values making the average less suitable as a measure of the typical. Because of these outliers, the median was chosen instead. However, the point is to grasp the typical R&D intensity of each industry. Whether the mean or the median is chosen implies no fundamental difference.

Apart from this, the two expressions look quite similar. Both compare the actual R&D intensity of each industry in the country concerned to the *typical* R&D intensity of each industry, and then sums the results of these comparison, weighing each component of the sum by the share of total value added in the country concerned accounted for by each industry.

What distinguishes the two expressions, then, is that whereas the industry intensity component compares the actual R&D intensity of each industry to the typical R&D intensity of each industry by taking the *difference*, in percentage points, between the actual and the typical R&D intensity, STIBERD compares them by *dividing* the actual R&D intensity by the typical R&D intensity, which results in a *quotient*. Thus, if a country has an R&D intensity which is equal to the typical in every single industry, the industry intensity component will be 0, whereas STIBERD will be 1. For discrepancies between the actual and the typical, the industry intensity component will give weighted average differences while STIBERD will give weighted average quotients. Thus, if a country has an R&D intensity which is 2 percentage points above the typical in every single industry, the industry intensity component will be 2. By contrast, if a country has an R&D intensity which is 20 per cent higher than the typical in every single industry, STIBERD will be 1.2. These figures are not directly comparable, of course.

One implication of this is compared to the industry intensity component, STIBERD will give a larger weight to differences in R&D intensity in industries where the typical R&D intensity is low than in industries where the typical R&D intensity is high. For instance, let us compare two industries, one with a typical R&D intensity of 2 per cent and the other with a typical R&D intensity of 20 per cent. Now, if a country has an R&D intensity which is 1 percentage point higher than the typical in both the industries, that is 3 per cent in the first industry and 21 per cent in the other, the STIBERD contribution will be much higher in the industry with the low typical R&D intensity than in the industry with the high typical R&D intensity, namely 1.5 as against 1.05. The other way around, given that both industries in a given country have the same STIBERD component value, for instance 1.5, this will mean that the low R&D intensity industry has an R&D intensity which is 1 percentage point above the typical while in the high R&D intensity industry the difference from the typical will be 10 percentage points. Thus, the two indicators give quite different answers to the question of what shall count as equal distances from the typical when one adds up R&D performances in industries with very different typical R&D intensities.

The differences between the two indicators are more fundamental than the above suggests, however, so let us try to analyse more closely their properties. The perspective from which this analysis will be performed is what is stated as the explicit purpose of the STIBERD indicator, namely to get a general expression of how well different countries perform in terms of R&D effort given the industrial structure that they actually have, the point of departure being that just comparing R&D intensities of the manufacturing sector as a whole across countries does not capture this adequately because this measure does not take into account neither differences in industrial structure across countries nor how the R&D expenditures are allocated across industries in the different countries.

Given that the question here is conceptualized as one of adjusting the R&D intensity of the manufacturing sector as a whole for other relevant factors (industrial structure and the allocation of R&D expenditures across industries), our analysis will be performed from the perspective of what determines the magnitude of each of the two indicators *when the R&D intensity of the manufacturing sector as a whole is taken as given*.

Let us start with the *industry intensity component*, which is written

$$IIC = \sum_{i=1}^{n} \left(I_i - \bar{I}_i \right) \cdot w_i \qquad (9)$$

As we saw above, this expression is essentially a residual. Rearranging equation (6), above, we get

$$IIC = \sum_{i=1}^{n} \left(I_i - \bar{I}_i \right) \cdot w_i = I_m - \sum_{i=1}^{n} \bar{I}_i \cdot w_i \qquad (10)$$

Thus, the industry intensity component equals the actual R&D intensity of manufacturing as a whole *minus* the expected R&D intensity of the manufacturing sector given the industrial structure, i.e. the R&D intensity the manufacturing sector would have had if the R&D intensity in each industry was equal to the typical R&D intensity for the industry.

First we note that everything else equal, the industry intensity component will increase if R&D expenditures are increased, as this will everything else equal increase I_m. This is, of course, very reasonable.

Apart from this, *given* the R&D intensity of the manufacturing sector as a whole, i.e. given I_m, the industry intensity component will increase when the expression $\sum_{i=1}^{n} \bar{I}_i \cdot w_i$ *decreases*. This means that, for a given R&D intensity of the manufacturing sector as a whole, the industry intensity component will be the higher the lower R&D intensity we would have *expected* given the industrial structure. Thus, for example, if the industrial structure is unfavourable in relation to having a high R&D intensity of the manufacturing sector, and the R&D intensity of the manufacturing sector in spite of this in fact is relatively high, the industry component will be high. This seems very reasonable as an adjustment for *industrial structure*.

On the other hand, the industry intensity component does not adjust for anything else. Especially, it does not adjust for the allocation of R&D expenditures across industries. This is very easy to see from equation (10), above. The *actual* R&D expenditures only appear in the R&D intensity of the manufacturing sector as a whole, I_m, the first part of the right hand side. Where the industrial structure figures, in the expression $\sum_{i=1}^{n} \bar{I}_i \cdot w_i$, the second part of the right hand side, it is only in connection with the *typical* R&D intensities. This second part is thus a pure expression of the industrial structure. Thus, a given industry intensity component value may occur through any distribution of R&D expenditures across industries. It may be a highly balanced one or, on the other hand, a highly unbalanced one, where, for instance, a highly disproportionate part of R&D expenditures are concentrated in a few high intensity industries accounting for a very small share of total production.

Let us now examine the STIBERD indicator in the same way. Its expression is

$$STIBERD = \sum_{i=1}^{n} \frac{I_i}{\bar{I}_i} \cdot w_i \qquad (11)$$

Let us now transform this expression. Now, the R&D intensity of each industry i, I_i, is given by $\frac{R_i}{V_i}$, where R_i is the R&D expenditures and V_i is the value added of industry i. Furthermore, w_i is given by $\frac{V_i}{V_m}$, where V_m is the value added of the manufacturing sector as a whole. We may thus write

$$STIBERD = \sum_{i=1}^{n} \frac{\frac{R_i}{V_i}}{\bar{I}_i} \cdot \frac{V_i}{V_m}$$

Here V_i may be eliminated, giving

$$STIBERD = \sum_{i=1}^{n} \frac{R_i}{\bar{I}_i} \cdot \frac{1}{V_m}$$

Here the factor $\frac{1}{V_m}$ may be put outside of the summation. We thus end up with

$$STIBERD = \frac{1}{V_m} \cdot \sum_{i=1}^{n} \frac{R_i}{\bar{I}_i} \qquad (12)$$

Again, i.e. as in the case of the industry intensity component, we first note that everything else equal, STIBERD will increase if R&D expenditures are increased. This is, again, very reasonable.

Let us now turn to the question of what determines the magnitude of STIBERD *given* the R&D intensity of manufacturing as a whole, i.e. given the sum of R&D expenditures and total value added. The remarkable thing to notice about this last expression, equation (12), is that *it does not at all contain the industrial structure of the country in question*. Apart from the inverted of the total value added of the manufacturing sector, entering as a constant factor, the expression only contains the R&D expenditures of each industry and the *typical* R&D intensity of each industry. Thus, given the R&D intensity of the manufacturing sector as a whole, i.e. the total R&D expenditures divided by total value added, as well as the *typical* R&D intensities of each industry, STIBERD *only* depends on the *allocation* of (the given) total R&D expenditures across industries, but, to repeat, *not* on the industrial structure.

Now, given the R&D intensity of the manufacturing sector a whole, what kind of distribution of the R&D resources across industries will give a high STIBERD and what kind of distribution will give a low STIBERD. It is easy to show[15] that irrespective of what this distribution looks like one will *always* get a higher STIBERD if R&D resources are reallocated *from* an industry with a *higher typical* R&D intensity *to* an industry with a *lower typical* R&D intensity.[16] This means that

15 I have done this in an earlier paper, ‘Understanding R&D performance: A note on a new OECD indicator’, STEP Report 16/94, Oslo 1994, pp. 7-9. This paper discusses the STIBERD indicator in more detail.

16 Intuitively, it is not difficult to see why this is so. If the numerator is given, you get a larger result if the denominator is small than if it is large.

STIBERD reaches its *maximum* when *all* the R&D resources are concentrated in one single industry, namely the industry where the *typical* R&D intensity is the *lowest*. It follows from what was shown further above that this is true *irrespective of what the industrial structure looks like*. Thus, for a given R&D intensity of manufacturing as a whole, STIBERD will reach its maximum value when all R&D resources are allocated to the industry with the lowest *typical* R&D intensity, irrespective of whether this industry accounts for a high share of total production or an absolutely insignificant share of total production. More generally, it always pays in terms of getting a higher STIBERD value to reallocate R&D resources from an industry with a higher typical R&D intensity to an industry with a lower typical R&D intensity, even if the high R&D intensity industry in question accounts for a substantial share of total production in the country in question while the low R&D intensity industry accounts for an insignificant share, and the high R&D intensity industry in the country in question has a very low R&D intensity compared to other countries while the low R&D intensity industry already has a more than satisfactory R&D intensity by comparative standards.

Thus, the STIBERD indicator seems to have a number of very undesirable properties. It does not, contrary to what was claimed for it, adjust for industrial structure. Its value is unambiguously determined by the total R&D expenditures in manufacturing and the total value added in manufacturing together with the distribution of the R&D expenditures across industries. Second, it reaches its maximum value when this distribution is extremely unbalanced and irrational. Third, this is all the more unreasonable as STIBERD is increased when R&D resources are distributed across industries in a way which is the opposite of the pattern generally observed in reality. STIBERD implies that R&D resources should always be concentrated in industries where the R&D intensity in general is low. However, there are obviously reasons for the kind of distribution of R&D intensity across industries which we in fact typically observe, and presumably they are at least in part good reasons. Thus, it would seem much more reasonable to give the *typical* distribution some kind of *positive* normative status, not a negative normative status.

In practice, the results we get if we use STIBERD as an indicator will probably not be so disastrous as the above might suggest. In practice, R&D resources are not that unequally distributed. In fact, even if one does accept that the pattern of *typical* R&D intensities across industries in an approximate way corresponds to differences in ‘real needs’ for R&D resources across industries, there is a widespread belief that in general R&D resources are too heavily concentrated in the high R&D intensity industries and that the low R&D intensity industries get too little. Thus, it may be the case that we practically never come across instances where the distribution is too unequal in favour of low intensity industries. We may thus happen to be in a situation where a redistribution of R&D resources from a high R&D intensity industry to a low R&D intensity industry practically always means a better distribution. Also, if a country has distributed relatively more R&D resources to low R&D industries and relatively less to high R&D industries than what is typical, the R&D resources will also be likely to be less concentrated in industries which account for relatively low shares of total production, as the really high R&D intensity industries generally account for a relatively low share of total production. Thus, STIBERD values may in practice not be so unreasonable.

Nevertheless, even if STIBERD values in practice should turn out to look quite reasonable, this result would seem to be somewhat more accidental that we should accept. At a more fundamental level, in its logic, STIBERD does seem to be quite defective.

To sum up the comparison of the two indicators, the *industry intensity component* adjusts for industrial structure in a straightforward and reasonable way, but does not take the allocation of R&D resources across industries into account at all. This is quite all right as far as it goes. The STIBERD indicator, on the other hand, does not at all take account of industrial structure, but only of the allocation of R&D resources across industries. This allocation, however, it takes account of in a highly distorted and unreasonable way. This is *not* all right, although the results may not be so disastrous in practice, given that some additional assumptions concerning how R&D resources generally are distributed happen to be true.

The problem of statistical interaction

The idea of adjusting for industrial structure to find an expression for how well each country performs in terms of R&D intensity given the industrial structure that it has would have been quite straightforward if for every country its performance compared to the typical was the same in all industries. If in one country the R&D intensity was 1 percentage point above the typical in every single industry and in another country it was 0.5 percentage points *below* the typical in every single industry, the *industry intensity component* would have been 1 and - 0.5, respectively, and these values would be very easy to interpret. Similarly, if instead for each country the *ratio* of its R&D intensity to the typical R&D intensity was the same in all industries, STIBERD would have given a very straightforward expression of the performance inside individual industries. If, for instance, one country had an R&D intensity which was 10 per cent higher than the typical in every single industry and another country had an R&D intensity which was 10 per cent *lower* than the typical in every single industry, STIBERD would have been 1.1 and 0.9, respectively, and again these values would be very easy to interpret.

However, in reality this is not the case. Instead, any country performs differently in different industries. In some industries a given country may perform substantially above normal, in others substantially below normal, in still others its performance may be approximately normal. Thus, when controlling for industry we find that the effect on R&D intensities of being in one country as opposed to being in another country varies across industries. This is what is called statistical interaction.

It is if this interaction is substantial and does not just represent relatively minor deviations from a pattern of approximately equal effects across industries that we get into trouble with the indicators considered above. For instance, if a high *industry intensity component* value comes about through a combination of R&D intensities which are very much higher than the typical in a few high R&D intensity industries which account for a very minor share of total production and R&D intensities which are well below the typical in a majority of the industries which together account for the bulk of total production, it would be problematic to interpret this indicator value as a measure of how well the country in general performs inside each industry. And we have seen that even more unreasonable distortions may result from application of the STIBERD indicator.

Let us now try to give an idea of how one might take interaction into consideration. We start with the decomposition of R&D intensity in manufacturing into a structure component and an industry intensity component (expression 6 above):

$$I_m = \sum_{i=1}^{n} \bar{I}_i \cdot w_i + \sum_{i=1}^{n} \left(I_i - \bar{I}_i\right) \cdot w_i \qquad (6)$$

Here the first term on the right hand side, $\sum_{i=1}^{n} \bar{I}_i \cdot w_i$, expresses the industrial structure, giving the R&D intensity in manufacturing which we would expect from knowledge only of the industrial structure of a given country. The second term is the industry intensity component, $IIC = \sum_{i=1}^{n} \left(I_i - \bar{I}_i\right) \cdot w_i$, which is in essence a residual. Let us now introduce the idea of the typical *weight* of each industry, corresponding to the idea of the typical R&D intensity of each industry applied above. This we will define simply as the *mean* or average weight of each industry across all countries, to be denoted $\overline{w}_i$. Together these $\overline{w}_i$s define the *typical* industrial structure. They sum up to 1, of course.

Let us now transform the above expression (equation 6) by adding $\overline{w}_i - \overline{w}_i$, i.e. by adding 0, to the factor w_i in the second component (the industry intensity component). This gives

$$I_m = \sum_{i=1}^{n} \bar{I}_i \cdot w_i + \sum_{i=1}^{n} \left(I_i - \bar{I}_i\right) \cdot (w_i + \overline{w}_i - \overline{w}_i)$$

Rearranging, we get

$$I_m = \sum_{i=1}^{n} \bar{I}_i \cdot w_i + \sum_{i=1}^{n} \left(I_i - \bar{I}_i\right) \cdot \overline{w}_i + \sum_{i=1}^{n} \left(I_i - \bar{I}_i\right) \cdot (w_i - \overline{w}_i) \qquad (13)$$

We now have expressed R&D intensity in manufacturing as a sum of *three* components. The first component is the structure component from above, which expresses the effect of industrial structure. Then comes a component which may in a sense be thought of as a ‘pure’ expression of the effect of the intensity inside each industry and lastly comes an interaction effect.

Since the first component is the familiar structure component from above, let us just look closer into how the industry intensity component IIC is divided into the sum of two components:

$$IIC = \sum_{i=1}^{n}\left(I_i - \bar{I}_i\right)\cdot \overline{w}_i + \sum_{i=1}^{n}\left(I_i - \bar{I}_i\right)\cdot (w_i - \overline{w}_i) \qquad (14)$$

The first part of this expression, $\sum_{i=1}^{n}\left(I_i - \bar{I}_i\right)\cdot \overline{w}_i$, in essence says what the R&D intensity in manufacturing in the country concerned would have been if it had had an industrial structure equal to the typical industrial structure, given the R&D intensity in each industry which it actually has. Then is subtracted a point of reference defined by the R&D intensity in manufacturing of a hypothetical country which has both the typical industrial structure and an R&D intensity equal to the typical in each industry. Thus, this component says what the R&D intensity in manufacturing in the different countries would have been if they all had the *same* industrial structure. In this sense one may think of it as an expression of the pure effect of R&D intensities inside each industry, controlling for industrial structure, i.e. holding industrial structure constant.

The second part of the expression, $\sum_{i=1}^{n}\left(I_i - \bar{I}_i\right)\cdot (w_i - \overline{w}_i)$, is an interaction term. Let us see how it works. First, let us suppose that in a given country the performance in terms of R&D intensity compared to the typical is the same in all industries, say, 2 percentage points higher. Since both the w_is and the $\overline{w}_i$s sum to 1 over all industries, the sum of differences between these magnitudes over all industries must be 0. Since each of these differences by assumption are to be multiplied by the same factor, namely 2, the interaction term itself in this case also would have to be 0. Thus, when the relative performance is the same in all industries, there is no interaction. This is in accordance with what we said above.

If the relative performance is *not* the same in all industries, we have the following, as we see by examining the interaction term. Industries in a given country with a *higher* share of total value added than the typical and a *higher* R&D intensity than the typical, contribute *positively* to the interaction term. Industries with a *higher* share of total value added than the typical and a *lower* R&D intensity than the typical, contribute *negatively* to the interaction term. Industries with a *lower* share of total value added than the typical and a *higher* R&D intensity than the typical, also contribute *negatively* to the interaction term. Finally, industries with a *lower* share of total value added than the typical and a *lower* R&D intensity than the typical, also contribute *positively* to the interaction term.

The interaction term will be the sum of these different contributions. The end result is that if a country tends to perform comparatively *better* in industries which account for *higher* than typical shares of total value added than in industries which account for lower than typical shares of total value, then the interaction term will be *positive*. If, on the other hand, it tends to do comparatively better in industries which account for less of total value added in this country than on average in other countries, the interaction term will be *negative*. If there is no clear tendency in either direction, the interaction term will be 0.

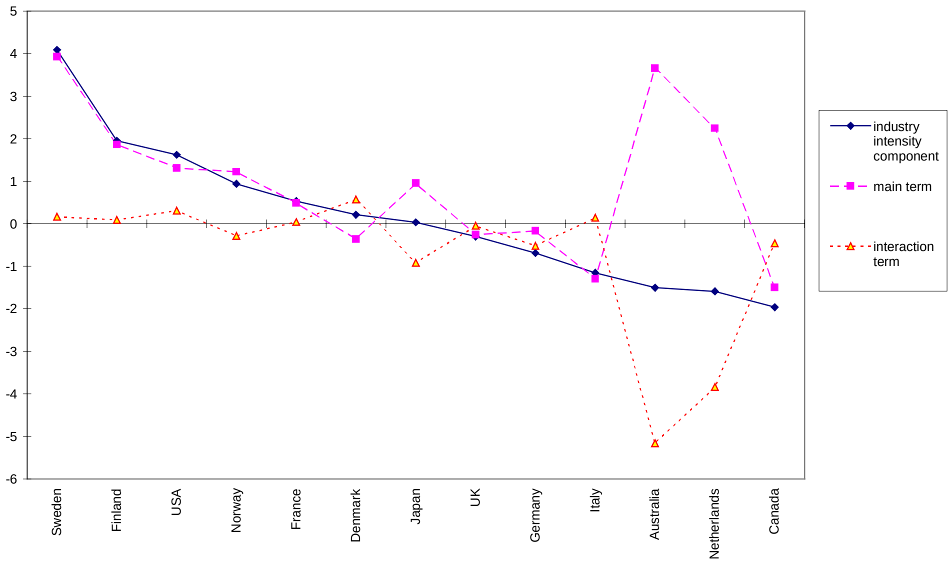
It was mentioned above that the industry intensity component is not standardized in such a way as to make the mean across all countries equal to 0. We may now see why this is so. In our case there are two reasons for this.

The most fundamental is the following. Let us assume that the *mean* of the R&D values across countries had been used as the typical R&D intensity in each industry, instead of the median. In that case, the mean across all countries of first term of the industry intensity component, $\sum_{i=1}^{n}\left(I_i - \bar{I}_i\right)\cdot\overline{w}_i$, which we may call the *main term*, would have to be 0. For each single industry, the sum of positive deviations across countries must then equal the sum of negative deviations, since by assumption given the industry the deviations are to be multiplied by the same weights in all countries. Summed over all industries, we then still get 0. Thus the *main term* would have to be 0. However, the industry intensity component also contains the interaction term, and there is no reason why the mean of this across all countries should necessarily be 0. More specifically, if there is a tendency for countries to perform comparatively better in terms of R&D in industries where they themselves have a higher than average share of their value added than in industries where they have a lower than average share of their value added, the mean of the interaction term across countries, and thereby also the mean of the industry intensity component, will be higher than 0. If the tendency is in the opposite direction, the mean of the interaction term and thereby also the mean of the industry intensity component will be lower than 0.

The second reason that the industry intensity component is not standardized in such a way as to make the mean across all countries equal to 0 is that in our case the typical R&D intensities of each industry are represented not by the mean but by the *median*. In that case, the mean of the *main term* is not necessarily 0 either. In fact, for almost all industries the distribution of R&D intensity across countries is skewed in such a way that the median is lower than the mean. This means that in our case the mean of the main term will be positive. We saw above (Table 13) that the mean of the industry intensity component across countries is 0.2. This can be decomposed into a mean of the main term of 0.9 and a mean of the interaction term of - 0.8 (the reason for the discrepancy when we add these components up is rounding error).

Let us now see how the industry intensity component is decomposed into the main term and the interaction term for our 1991 data. This is shown graphically in Figure 26, below.

Figure 26: Decomposition of industry intensity component into main term and interaction term, 1991.

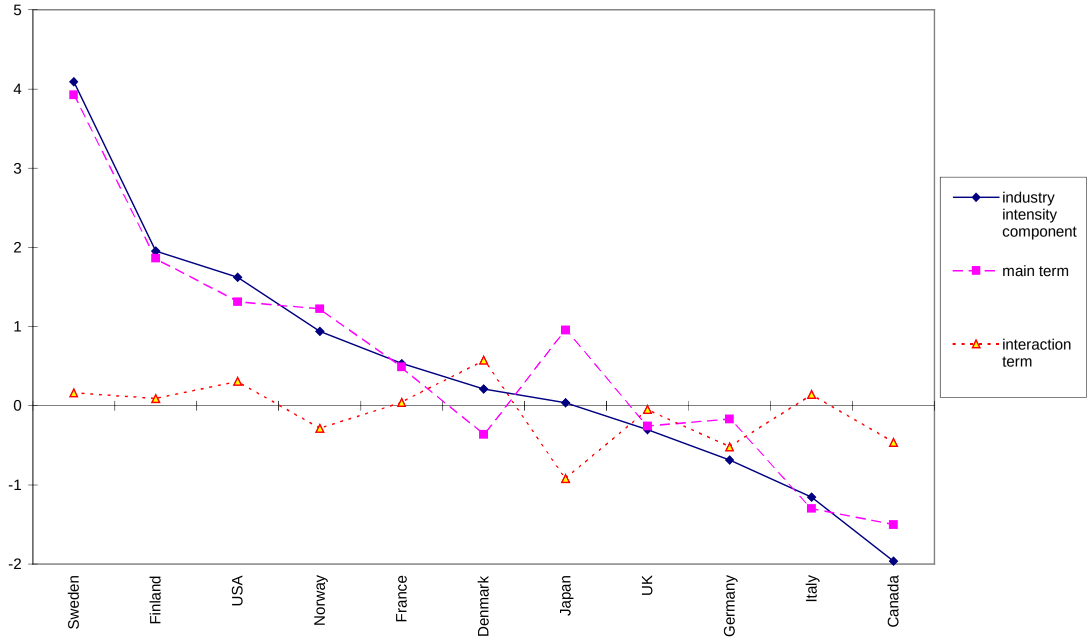


Let us treat the cases of Australia and Netherlands separately from the others, as it is very clear that they are special. In both cases, the industry intensity component is low. What is special is that this value comes about through the combination of a very high main term value and a very low interaction term value. The reason is that both these countries compared to the other countries have an extremely high R&D intensity in one industry each (*Radio, TV and Communication Equipment* in Australia and *Electrical Machinery excluding. Communication Equipment* in Netherlands). At the same time, in both countries the industry in question accounts for a much lower share of total value added than in all countries on average. Thus, if each country had had an industrial structure equal to the typical, as is the assumption behind the main term expression, the extremely high intensity in each of the two countries would have achieved a much higher weight. Consequently, on the assumption that all countries had this same industrial structure, both these countries would have come out with a high R&D intensity in manufacturing. Thus, the main term value is high in these two countries. The complement to this high main term value is that in the interaction term for these two countries these very large positive deviations from the typical in each of the extreme value industries get a sizeable negative weight. The result is that the interaction term gets a very high *negative* value.

Thus, the deviant pattern of main term and interaction term values in these two countries expresses the presence of extreme values. In fact, the values in question are so high that one may doubt if they are correct. For instance, one may suspect that total R&D expenditures have not been correctly ascribed to the different industries or that there are differences in the classification of industries between the R&D expenditures statistics and the value added statistics.

Thus, let us exclude Australia and Netherlands when we look at the other countries. The decomposition of the industry intensity component into a main term and an interaction term for the remaining countries is shown graphically in Figure 27, below.

Figure 27: Decomposition of industry intensity component into main term and interaction term, excluding Australia and Netherlands, 1991.



We see that when we exclude the other countries, the main term traces the industry intensity component quite closely, while the interaction term lies quite close to 0. (In fact, the correlation between the industry intensity component and the main term among these countries is as high as 0.97, while the correlation between the industry intensity component and the interaction term is only 0.39.) We see that Denmark has the highest interaction term value (0.6), while Japan has the lowest (- 0.9). Now, since Japan has a high share of its production accounted for by high R&D intensity industries (as we saw, Japan has the highest structure component value of all the countries in 1991), this may suggest that Japan performs relatively better in low R&D intensity industries than in high R&D intensity industries.

However, an interaction term value close to 0 does not necessarily mean that R&D resources are distributed across industries in such a way that the R&D performance compared to the typical is the same in all industries. To see this, note that if a country has an industrial structure which is identical to the typical, the interaction term has to be 0, no matter how bizarre the distribution of R&D resources across industries. Thus, for a country which has an industrial structure which is close to the typical,

Furthermore, also for countries with an industrial structure quite different from the typical an interaction term value close to 0 is compatible with quite anomalous distributions of R&D intensities across industries, if the positive contributions and the negative contributions to the interaction term roughly balance each other.

If one wanted to use the typical distribution of R&D intensities across industries as a norm for what is a rational distribution, and wanted to incorporate this in an indicator which measured R&D intensity performance in a way which adjusted for both industrial structure and the allocation of R&D expenditures across industries, one could perhaps proceed roughly along the following lines: First one could, on the basis of the industry intensity component value, find out what the R&D intensity in each industry would have been if relative R&D performance was the same in all industries, as measured by the difference from the typical R&D intensity. Then one could construct a measure of the sum of absolute deviations of the actual R&D intensities from the hypothetical same performance intensities across all industries. Finally, one could take the industry intensity component as a point of departure and the *subtract* this deviation measure to arrive at an indicator which adjusted for both industrial structure and the allocation of R&D resources across industries.

However, we will not go further along these lines here. Instead we will propose another indicator with the same objective as the STIBERD indicator, but which hopefully has more desirable properties.

An alternative indicator: weighted mean of ranks

The objective of this indicator, as of the OECD STIBERD indicator, is thus to give an expression of how well a given country on average performs in terms of R&D intensity inside each industry, given the industrial structure that it actually has.

The industry intensity component defined above in a straightforward and sensible way adjusts R&D intensity in manufacturing for industrial structure. It is quite useful if we just want to do that. However, however, of an indicator which is to measure how well a country typically performs inside its individual industries we would want that it should also take into account the distribution of R&D expenditures across industries. We have seen that measures which just add together measures of performance inside different industries in such a way that extreme performances in a few industries may more than weigh up for weak performances in the majority of industries are highly problematic.

A way to solve this problem might thus be to define a measure where once a certain level of R&D intensity in a given industry is reached one does not get any more 'points' by allocating more resources to this industry. To construct an indicator along these lines, we propose to base the indicator on simple rank numbers among the countries inside each industry.

In each industry the country with the highest R&D intensity gets the rank number 1, the country with the second highest R&D intensity gets the rank number 2, and so on down to rank number 13 for the country with the lowest R&D intensity in the industry. To each country there is thus assigned 22 rank numbers, one for its rank in terms of R&D intensity in each industry.

One possibility here would be to define the indicator simply as the *average* of these 22 rank numbers for each country. A country which was ranked among the 5 highest performers in every single industry would then get an average rank of perhaps around 3 or 4, while a country which was ranked among the 5 *lowest* performers in every single industry would get an average rank of perhaps around 9 or 10. If we had done this, the average of these indicator values across all countries would necessarily be 7, since this is the average for each single industry.[17] If some countries are on average better than 7th across all industries, this must be balanced by other countries being on average worse than 7th.

However, in agreement with the logic of the argument so far in this paper, we will not use the *simple* average of rank numbers. Instead, we will use the *weighted* average, where the weight of each industry, as above, is defined by the share of the total value added in manufacturing which it accounts for in the country in question.

When we do this, however, there is no longer any reason that the average of the indicator values across all countries should be 7. We can easily see this if we consider a hypothetical extreme example. Assume that in each country one single industry accounts for 95 per cent of value added in manufacturing and that this industry is a different industry in all countries, i.e. country A has 95 per cent of its production in industry 1, country B has 95 per cent of its production in industry 2, and so on. Assume further that each country is ranked in first place, i.e. has the highest R&D intensity among all the countries, in precisely the industry which dominates its manufacturing sector. In that case, it is easy to see that all countries would have an indicator value close to 1, no matter how each of them ranked in the remaining industries accounting for 5 per cent of their production. By contrast, if each country ranked in last place in precisely the industry which dominates its production, al countries would have an indicator value close to 13.

More generally, if there is a tendency across the countries for performing better in industries which account for a higher share of total manufacturing production in the country in question than in all countries on average, then the average of the indicator value across all countries will be better than 7, i.e. be lower than 7 in value. If the

---

17 In our case, this is not strictly true, for a specific reason. In the Danish case there are no figures neither for R&D expenditures nor for value added in the Aircraft industry. I have treated this as if Denmark does not have production in this industry at all, which is probably not far from the truth anyway. Thus, in the Aircraft industry the ranks go from 1 to 12 only. I have done nothing to adjust for this anomaly. Thus, in our case the average of the indicator across all countries would have been slightly better than 7 (to be precise, 6.955).

opposite tendency is operative, the average of the indicator value across all countries will be higher than 7 in value. If there is no tendency in either direction, the average will be around 7.

This weighted mean of ranks indicator does not have the same undesirable properties as STIBERD. It is not possible here that an extreme concentration of R&D resources in one industry gives a maximum indicator value. Rather, if country A has an R&D intensity in a given industry which is much higher than that of the second ranked country, country A may reallocate resources from this industry to climb in the classification in other industries without the rank in the first industry being affected at all. When the first position in a given industry is reached, there is nothing more in terms of the indicator value to allocate more resources into this industry, whereas an allocation to other industries where the country is not in first place may improve the rank in these industries.

Also, the distribution of R&D intensities across countries inside each industry will generally be such that the differences between the countries are larger towards the top of the distribution than towards the middle. This means that, for instance, if a country has reached second place in the ranking in one industry, it will generally, everything else equal, require more R&D resources to rise from second to first place in this industry than it will require to rise from, say, seventh to sixth place in another industry. Thus, everything else equal, one will gain more in terms of the indicator value by allocating resources to industries where the country is ranked towards the middle of the distribution than towards the upper end of the distribution. There might also be a tendency for distances between the rank numbers to be larger towards the bottom of the distribution than towards the middle. However, this tendency is not likely to be substantial, as the distributions are limited downwards by 0 being the absolute minimum value.

We will come back in more detail to how this indicator works further below. Let us first, however, look at the results we get when we apply the indicator to our 1991 data. In Table 27, below, the weighted mean of ranks indicator values for all the countries are shown, the countries being ranked from top to bottom.

Table 27: Countries ranked according to the weighted mean of ranks indicator.

Country	*Mean*
Finland	2.9
Japan	4.2
Sweden	5.0
USA	5.6
Norway	6.7
Germany	6.8
Denmark	7.1
France	7.6
Netherlands	7.9
Australia	8.1
UK	9.0
Canada	9.7
Italy	11.2

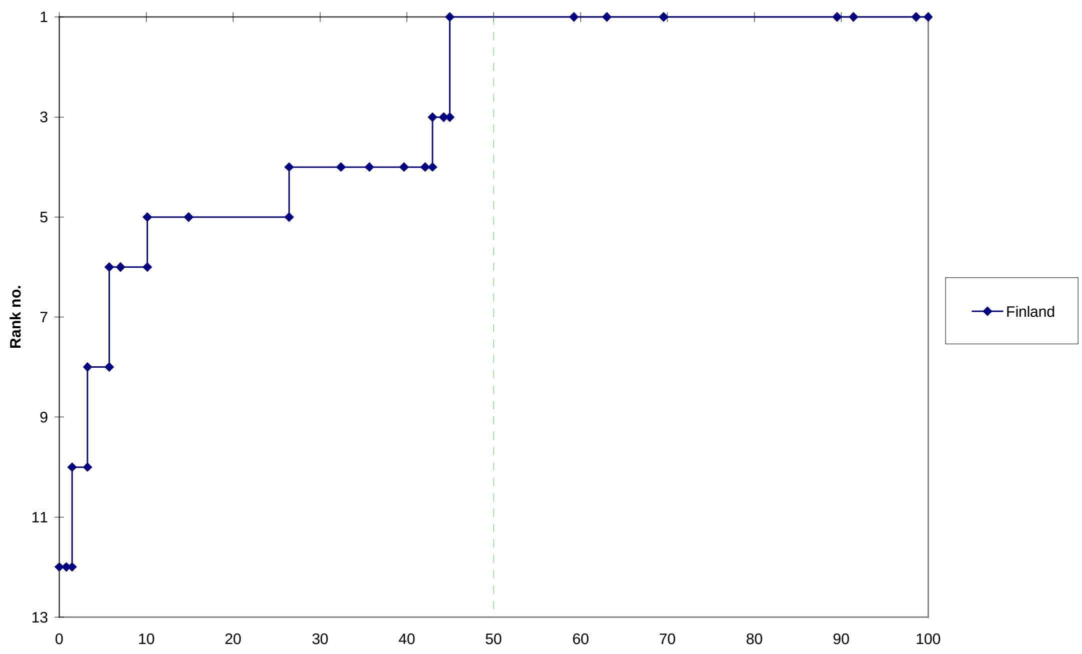
We see that Finland now comes up on top with a very high mean rank (low figure) of only 2.9, i.e. on average Finland ranks third inside the individual industries, when the industrial structure has been taken into account. Japan is in second place here with a quite high mean rank of 4.2. Then follow Sweden and USA.

In the bottom Italy is as unchallengeable as Finland is in the top, with a weighted mean rank of 11.2 (of maximum possible of 13!). Then follow Canada, UK and Australia.

The average of this distribution is 7.1, which means that there is no tendency neither for countries to perform better in the industries which are comparatively more important for them relative to other countries, nor for the opposite.

Let us now see how these indicator values come about for some of the countries. We start with Finland, who apparently is very highly ranked in all the industries which matter quantitatively speaking in its manufacturing sector. A graphical presentation of Finland’s performance is given in Figure 28, below.

Figure 28: Finland's R&D intensity ranking in different industries (y-axis). Industries weighted by share of value added and ranked according to performance (x-axis). 1991.

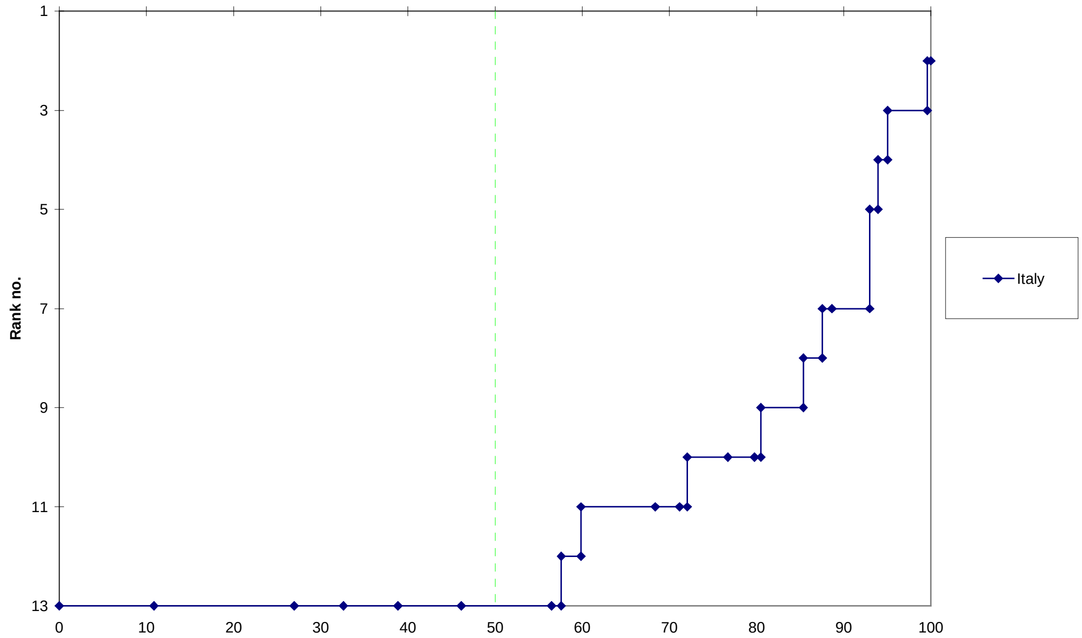


In this figure, the rank of Finland in the different industries among the 13 countries is shown along the y-axis. The industries are lined up along the x-axis. They each occupy a fraction of this axis which is proportional to their share of manufacturing value added. They are ranked according to performance as measured by Finland's rank among the 13 countries in each industry.

We see that Finland ranks highest in seven of the 22 industries, and these seven industries account for more than half of total manufacturing value added in Finland. At the lower end, only four industries rank lower than seventh place, and these four industries account for less than six per cent of value added. This is quite extraordinary. (Can it really be true?)

Let us now look at the country placed at the opposite end of the list, namely Italy. Italy's ranking in the different industries is shown in Figure 29, below.

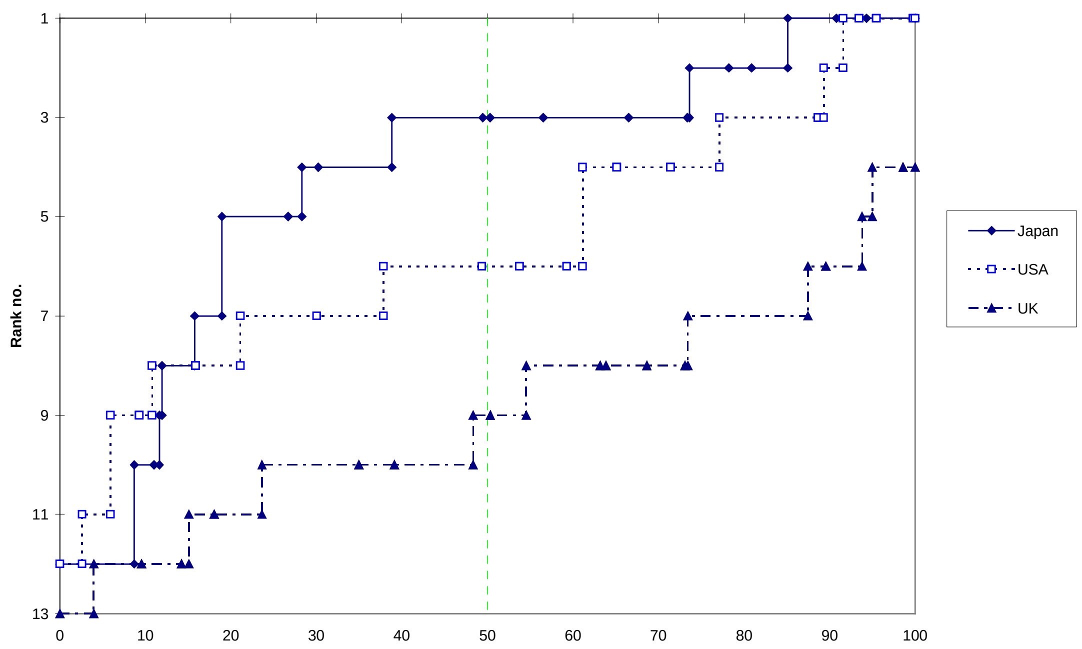
Figure 29: Italy's R&D intensity ranking in different industries (y-axis). Industries weighted by share of value added and ranked according to performance (x-axis). 1991.



This figure is almost the exact opposite of the one which represented Finland. Italy ranks in last place in seven industries, which together account for well over 50 per cent of manufacturing value added in Italy. Only in four industries is Italy placed in the upper half of the ranking, i.e. better than seventh. These four industries account for no more than seven per cent of value added.

Let us also give examples of countries between these two extremes. In Figure 30, below, Japan, USA and UK are chosen.

Figure 30: Japan's, USA's and UK's R&D intensity ranking in different industries (y-axis). Industries weighted by share of value added and ranked according to performance (x-axis). 1991.



Here it emerges quite clearly that by and large Japan performs better than USA and USA better than UK. But note that the three countries do not have the same ranking of industries along the x-axis. Note also that UK generally comes out quite low in most industries. In no industry UK ranks better than 4th.

Let us now look at the relationship between the industry intensity component values and the weighted mean of ranks indicator values. Table 28, below, shows how the different countries rank on each of these indicators for the 1991 data.

Table 28: Ranks of the different countries on the industry intensity component (IIC) and the weighted mean of ranks (WMR) indicators, 1991.

Country	*IIC*	*WMR*
Finland	2	1
Japan	7	2
Sweden	1	3
USA	3	4
Norway	4	5
Germany	9	6
Denmark	6	7
France	5	8
Netherlands	12	9
Australia	11	10
UK	8	11
Canada	13	12
Italy	10	13

The most spectacular change in the rankings when we go from the industry intensity component to the weighted mean of ranks indicator is that Japan rises from 7th to second place. Finland also does better, rising from second to first place. Further down in the distribution both Netherlands and Germany rise three places, to no. 6 and no. 9, respectively.

In the opposite direction, Sweden drops from a very clear first position to third. Further down both France, UK and Italy drop three places, to 8th, 11th and 13th, respectively.

To look further into the relationship between the two indicators, let us correlate the variables with all 13 countries as units of observation. First, however, we propose to transform the weighted mean of ranks indicator so that it too gives the best performers high values and the worst performers low values. we also propose to make it vary such that 10 is the maximum value and 0 is the minimum value. This is accomplished by the following formula:

$$WMR^* = \frac{(13 - WMR) \cdot 10}{12}$$

where *WMR* is the original indicator and *WMR** is its transformed expression. As one can easily verify, a *WMR* of 13 gives a value of 0 in the transformed expression, while a *WMR* of 1 gives a value of 10 in the transformed expression.

Table 29, below, gives the weighted mean of ranks indicator values in both their original and transformed expression for all the countries.

Table 29: The weighted mean of ranks indicator values, original and transformed expression, 1991.

Country	*WMR*	*WMR**
Finland	2.9	8.4
Japan	4.2	7.3
Sweden	5.0	6.7
USA	5.6	6.1
Norway	6.7	5.2
Germany	6.8	5.2
Denmark	7.1	4.9
France	7.6	4.5
Netherlands	7.9	4.2
Australia	8.1	4.0
UK	9.0	3.4
Canada	9.7	2.8
Italy	11.2	1.5

When we now correlate the industry intensity component with the transformed expression for the weighted mean of ranks indicator, we get $r = 0.70$. (Of course, if we correlated *IIC* with the untransformed version of *WMR*, the correlation coefficient would simply changes sign, becoming - 0.70 instead of 0.70.)

Thus, there is a very clear positive relationship between the two indicators. In general, a country with a high *IIC* will also be high on *WMR*, and vice versa. This, of course, seems very reasonable.

At the same time, the relationship is far from perfect. With $r = 0.70$, we have an R^2 of 0.49, which means that only half of the variation in the weighted mean of ranks indicator is accounted for by the variation in the industry intensity component. What else influences the weighted mean of ranks indicator, then?

Let us go back to the discussion above about what kind of distribution of R&D expenditures across industries which will give a high weighted mean of ranks indicator value. There it was said that everything else equal one would gain more in terms of the indicator value by allocating resources to industries where one at the outset rank more towards the bottom and especially towards the middle end of the distributions than towards the top. The reason is that the distances between the countries tend to be larger towards the top. However, it will also in general be the case that in industries with high R&D intensity the distances between the countries in the classification will be greater than in industries with low R&D intensity. Thus it will in general require less R&D resources to climb in the classification in industries with low R&D intensity than in industries with high R&D intensity, provided that one is not already at the top in the low R&D intensity industries and thus have little more to gain there. This means that we should expect that everything else equal, notably given the industry intensity component value, countries who tend to perform relatively better in industries with low R&D intensity than in industries with high R&D intensity will also tend to get a higher weighted mean of ranks indicator value than countries where the opposite is the case.

How can we measure this? we propose for each country to correlate two variables, with the 22 industries as observation units. The first variable is simply the *typical* R&D intensity of each industry. This variable will thus be the same for all countries. The second variable will be an expression of how well the country performs relative to other countries in each industry. We here propose to use a modified version of standard scores or *z scores*. Z scores are scores which are standardized by means of the mean and the standard deviation of the original scores in a distribution. One takes the original score, subtracts the mean and then divides the difference by the standard deviation. Thus an observation which has a score equal to the mean gets a z score of 0, an observation with a score which is one standard deviation above the mean gets a z score of 1, and so on. The *modification* we propose here is motivated by the presence of some extreme values as explained before. Thus we will substitute the *median* for the mean, and instead of the standard deviation for the full distribution we will use the standard deviation for the distribution excluding its maximum and minimum score.

The results are shown in Table 30, below.

Table 30: Correlation, for each country, between the typical R&D intensity of each industry and the country's relative R&D performance in each industry as measured by modified z scores, 1991.

Country	*Correlation*
UK	0.47
Italy	0.38
France	0.36
Sweden	0.29
Norway	0.11
Australia	0.10
Canada	0.08
USA	0.02
Germany	-0.01
Netherlands	-0.02
Denmark	-0.22
Japan	-0.47
Finland	-0.57

Let us refer to the correlation coefficients reported in Table 30 as "relative performance values," meaning by that that they indicate whether the country in question performs relatively better in high R&D intensity industries or in low R&D intensity industries.

We see that UK, Italy, France and Sweden have a tendency for performing relatively better in industries with high R&D intensity than in industries with low R&D intensity, while the opposite is the case for Finland and Japan. Now, UK, Italy, France and Sweden were precisely the countries which most clearly performed better on the industry intensity component indicator than on the weighted mean of ranks indicator, while the opposite was the case for Finland and Japan.

Let us correlate this relative performance variable with the weighted mean of ranks indicator in the standardized version (*WLR**). The result is a correlation coefficient (r) of - 0.70. This means that the more a country tends to perform relatively better in industries with high typical R&D intensity than in industries with low typical R&D intensities, the *lower* it tends to perform on the weighted mean of ranks indicator.

Thus, we have seen that both the industry intensity component and the relative performance variable correlates quite strongly with the weighted mean of ranks indicator for our 1991 data, both coefficients being 0.70 in absolute value. Let us now set up a regression equation to see how these two variables combine to influence the weighted mean of ranks indicator. In order that the regression coefficients be comparable to the correlation coefficients already reported, we choose to report standardized regression coefficients. These are obtained by transforming all three variables, the dependent and the two independent variables, into z scores, where the raw scores are transformed by subtracting the mean and dividing by the standard deviation, as explained above. The regression is then run on these transformed scores. The coefficients will then measure the change in the dependent variable in terms of standard deviations associated with a one standard deviation increase in each of the independent variables, holding the other independent variable constant. This is exactly what the correlation coefficient measures in the bivariate case.

Let $\hat{Y}$ be the predicted value of the weighted mean of ranks indicator in the standardized version, X_1 the industry intensity component and X_2 the relative performance variable, all values expressed in standardized or z scores. We then get the following results:

$$\hat{Y} = \underset{(0.10)}{0.64\,X_1} - \underset{(0.10)}{0.65\,X_2}$$

$R^2 = 0.90$ N = 13

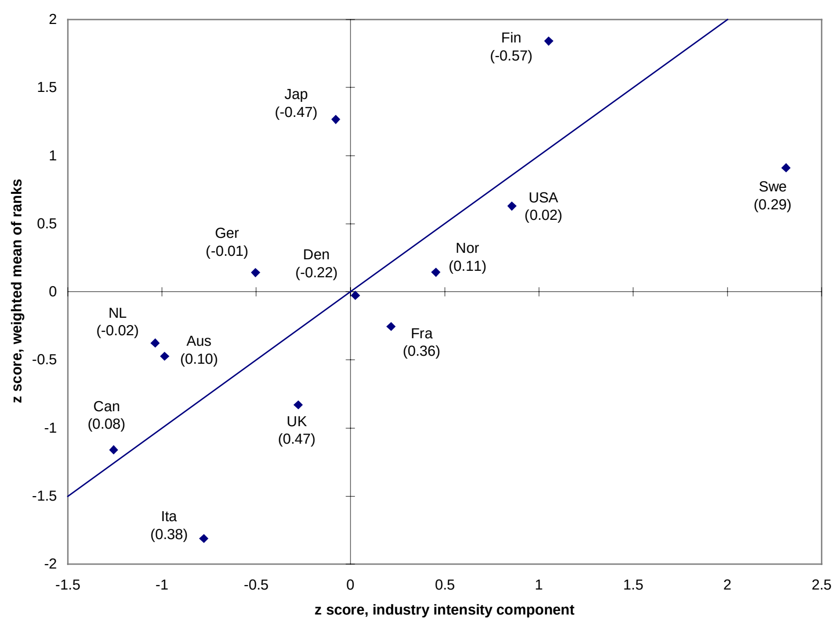
The figures in parentheses are the standard errors of the coefficients. Since we use z scores and thereby standardized regression coefficients here, the constant term is necessarily 0.

We see that the standardized regression coefficients are 0.64 for the industry intensity component and - 0.65 for the relative performance variable. Thus, the coefficients are almost the same as in the bivariate cases, where they were 0.70 and - 0.70, respectively (the correlation coefficients). This means that the correlation between the two independent variables of the above model is low; in fact, the correlation coefficient is - 0.09. With an R^2 of 0.49 for both the bivariate models, this means that R^2 for the model with both independent variables is very high, namely 0.90.

Substantively, this means that holding the relative performance variable constant, the value of the standardized weighted mean of ranks indicator increases as the industry intensity component increases. More precisely, it increases by 0.64 standard deviations for every standard deviation increase in the industry intensity component. Likewise, holding the industry intensity component constant, the value of the standardized weighted mean of ranks indicator *decreases* as the relative performance variable increases. More precisely, it decreases by 0.65 standard deviations for every standard deviation increase in the relative performance variable. Together these two variables account for almost all of the variation in the weighted mean of ranks indicator, or, to be precise, 90 per cent of the variation.

The relationship between these variables for our 1991 data is shown in Figure 31, below.

Figure 31: Z scores weighted mean of ranks indicator, standardized version (y-axis), z scores industry intensity component (x-axis), relative performance correlation coefficient (in parentheses), 1991.



Along the x-axis we have the z scores of the industry intensity component, along the y-axis the z scores of the weighted mean of ranks indicator (in the standardized version), and in parentheses are the relative performance correlation coefficients. Notice that these figures are the actual correlation coefficients, not their z scores. The line drawn through the diagram is a 45 degree line. Countries on this line have the same value on the industry intensity component and the weighted mean of ranks indicator, as measured by the z scores, i.e. by the distance from the mean in terms of standard deviations. Countries to the left of and above this line have a higher weighted mean of ranks z score than industry intensity component z score, while the reverse is true for countries to the right of and below the line. Thus, this is *not* a regression line.

We clearly see that there is a positive association between the scores on the industry intensity component and the scores on the weighted mean of ranks indicator. Since the correlation coefficient between these two variables was 0.70, the gradient of the regression line would have been 0.70 (the 45 degree line, of course, has a gradient of 1). We also see that for approximately the same industry intensity component value, the weighted mean of ranks indicator value tends to be higher the *lower* the relative performance correlation coefficient. To state this differently, we see that countries with a negative relative performance correlation coefficient tend to have a higher weighted mean of ranks z score than industry intensity component z score, while the reverse is true for countries with a relative performance correlation coefficient.

Does this seem reasonable? Clearly, it seems reasonable that the indicator should increase with increasing value of the industry intensity component, everything else equal. It also seems reasonable that this correlation is fairly high, but not so high that the two measures in practice become the same.

Next, we turn to the correlation between the weighted mean of ranks indicator and the relative performance coefficient. Is it reasonable that this correlation should have a fairly high negative value, everything else equal? Stated differently, is it reasonable that, for any given value of the industry intensity component, the indicator we are trying to construct should increase when R&D resources are allocated away from high R&D intensity industries and towards low R&D intensity industries? If one believes that R&D resources in general are too heavily concentrated in high R&D intensity industries with too few resources going to low R&D intensity industries, this result would indeed seem reasonable. However, if one believes that the distribution which one typically finds is by and large a rational distribution, this correlation between the weighted mean of ranks indicator and the relative performance coefficient would seem to imply that the weighted mean of ranks indicator is not a reasonable indicator. If one believes that the typical here should have a normative status, a relevant test of proposed indicators might be the following. One constructs a variable which has a maximum value when the distribution of R&D intensities across industries matches the typical distribution, and which decreases as the distribution of R&D intensities across industries deviates more and more from the typical distribution. A good indicator should then correlate positively with this distribution variable when the industry intensity component is controlled for. One possibility for an indicator might here simply be to use this distribution variable in its construction, for instance by constructing the indicator as a weighted sum of this distribution variable and the industry intensity component.

Let us now briefly compare the results we get when we use the weighted mean of ranks indicator with the results we get if we instead use STIBERD. In fact, it turns out that for our 1991 data the results are almost identical. The correlation of the STIBERD values with the weighted mean of ranks values is as high as 0.95. The correlation of STIBERD with the industry intensity component is 0.70, with the relative performance coefficient - 0.69. Regressing STIBERD (Y) on the industry intensity component (X_1) and the relative performance coefficient (X_2), using standardized regression coefficients, we get

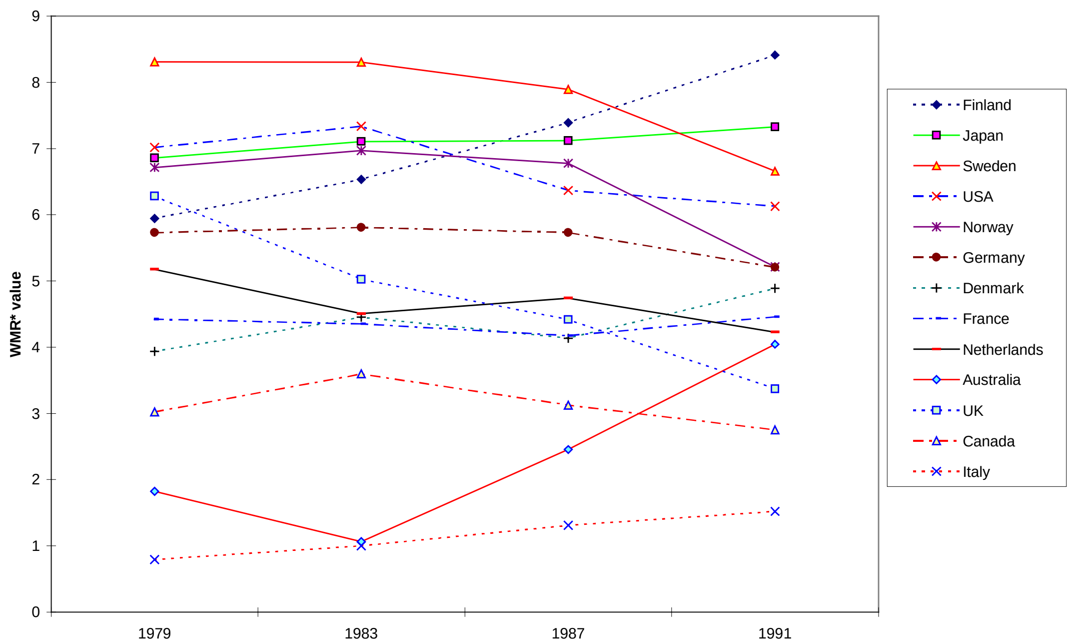
$$\hat{Y} = \underset{(0.10)}{0.65\, X_1} - \underset{(0.10)}{0.63\, X_2}$$

$R^2 = 0.89$ N = 13

Obviously, we are in actual fact far from situations where the use of STIBERD gives 'wild' results. With the type of distribution of R&D resources across industries which we actually have in the different countries in 1991, STIBERD seems to function reasonably well. However, the logical problems exposed further above still makes it unattractive as an indicator, in the present author's opinion.

We now turn to a comparison of the weighted mean of ranks indicator values of 1991 with previous years. In Figure 32, below, the distributions in 1979, 1983, 1987 and 1991 are compared.

Figure 32: Weighted mean of ranks indicator values (standardized version), 1979, 1983, 1987 and 1991.



Note that this is the standardized version of the indicator, where a first place in all industries will give the value 10 and a bottom place the value 0.

We see that there are some changes in the relative position of the countries over these 12 years. For instance, Finland steadily and substantially improves its position (moving from 6th to 5th and then to 2nd and finally 1st place), while the UK equally steadily and substantially declines (dropping from 5th to 7th and then to 8th and 11th place).

To get a further idea of how these indicator values change, let us look at the correlation of the distribution of these values between each pair of these year. The correlation coefficients are shown in Table 31, below.

Table 31: Correlation matrix of distribution of weighted mean of ranks indicator values across countries between 1979, 1983, 1987 and 1991.

Year	*1979*	*1983*	*1987*	*1991*
1979	1			
1983	0.97	1		
1987	0.94	0.96	1	
1991	0.73	0.79	0.89	1

We see that the correlation between the first and the last of these years, 1979 and 1991, is 0.73. This is not nearly as high as the corresponding correlation for the structure component, of course, which was 0.94. However, it is noticeably higher than the corresponding correlation for the industry intensity component, which was only 0.60.

Again, there seem to occur more substantial changes between 1987 and 1991 than earlier in the period. Thus, at a four year distance, the correlation between 1979 and 1983 is 0.97 and between 1983 and 1987 it is 0.96, while between 1987 and 1991 it is only 0.89. Similarly, at an eight year distance, the correlation between 1979 and 1987 is 0.94, while it is only 0.79 between 1983 and 1991.

Lastly, let us see if there is any association between score on the weighted mean of ranks indicator and size of economy. In Table 32, below, the correlation between, on the one hand, the weighted mean of ranks indicator, in its standardized version, and, on the other hand, GDP and Ln GDP, respectively, is shown for the years 1979, 1983, 1987 and 1991.

Table 32: Correlation coefficient (r) *between, on the one hand, the weighted mean of ranks indicator, standardized version, and, on the other hand, GDP and Ln GDP for the years 1979, 1983, 1987 and 1991.*

Year	*GDP*	*Ln GDP*
1979	0.26	0.01
1983	0.29	- 0.02
1987	0.18	- 0.14
1991	0.20	- 0.15

As we see, we find no clear association between size of economy and score on the weighted mean of ranks indicator. None of the correlation coefficients are statistically significant at the 5 per cent level. The correlation coefficients involving GDP have very small positive values. The coefficients involving Ln GDP are virtually zero or slightly negative.

CONCLUSION

In this paper we have looked at variation across countries in R&D intensity in manufacturing, and we have looked at how R&D intensity in manufacturing is related to size of the economy. Into this picture we have then brought variation in industrial structure across countries. We have also discussed R&D intensity in manufacturing as an indicator of innovative performance of countries.

In our data we find at best a weak positive relationship between economy size and R&D intensity in manufacturing. We find no clear evidence of a positive association between these variables. On the other hand, neither do our findings constitute any clear evidence against an hypothesis that there is a substantial association between size of economy and R&D intensity in manufacturing.

However, when we decompose R&D intensity in manufacturing into, on the one hand, a component expressing the industrial structure of the country in question, and, on the other hand, a component expressing how the country in general compares with the other countries in terms of R&D intensity inside each industry, we find a clear and strong positive association between economy size and the *structure component*. The larger the size of the economy, as measured by GDP, the higher the R&D

intensity in manufacturing we would predict from knowledge only of the industrial structure, or, in other words, the more the industrial structure is favourable to a high R&D intensity in manufacturing.

Seeing that some countries have an industrial structure which would make us expect a *high* R&D intensity in manufacturing as a whole while other countries have an industrial structure which would make us expect a *low* R&D intensity in manufacturing, the idea of controlling or adjusting for industrial structure when comparing R&D intensities in manufacturing was introduced. One would thus try to measure how a given country in general compares to other countries in terms of R&D intensity given the industrial structure that it really has. It was pointed out that, even disregarding the more general problems of using R&D intensity as a performance measure, it is not unproblematic to use an indicator of R&D intensity which is adjusted for industrial structure as a performance measure. Industrial structure cannot simply be treated as a given, as an ‘endowment’, but will itself precisely to some extent express innovative capacity, i.e. performance. Using these kinds of indicators thus presupposes careful interpretation: to what extent is the industrial structure in question a rational one, and to what extent does it precisely reflect low innovative capability, etc.?

One indicator proposed was simply to take the difference between actual R&D intensity in manufacturing and the R&D intensity predicted from knowledge of the industrial structure. This indicator says how the country in question on average compares in term of R&D intensity to what is typical among all countries. The average here is a weighted one, where the weights are defined by each industry’s share of total manufacturing value added in the country in question. Thus the indicator compares the R&D intensity of country in question with that of other countries on the basis of the industrial structure which it has. This indicator in a quite straightforward sense adjusts for differences in industrial structure.

However, this indicator only takes account of industrial structure, but not of the distribution of R&D expenditures across industries. Since an average, including a weighted average, of performances inside the individual industries may be an average of very different performances, it seems reasonable that an indicator which intends to capture how a country in general performs inside the individual industries would have to take account also of the distribution of R&D expenditures across industries in addition to industrial structure.

A point made in the paper is that an indicator which takes account also of the distribution of R&D expenditures across industries is bound to presuppose what a rational distribution of expenditures would be which the actual distribution is measured against. If this is not done explicitly, it will nevertheless be contained implicitly in the indicator. It was shown that the indicator presented in the paper, the weighted mean of ranks indicator, increases in value as more R&D expenditures are allocated in the direction of low R&D intensity industries and less in the direction of high R&D intensity than what today is typical or average (median) among the countries included in the study. Some would say that this is reasonable because R&D resources today in general tend to be too highly concentrated in a few high R&D intensity industries to the detriment of the large number of low R&D intensity industries which account for the bulk of manufacturing production. Others would say that this is not reasonable because the distribution we typically find is by and large reasonable and should be taken as the norm. In conclusion, to take account of the distribution of R&D expenditures across industries in a meaningful way, one should have a well grounded idea of what a rational distribution looks like.