Improving the understanding of innovation by using test techniques

Vincent Fructuoso van der Veen

Discussion paper (09021)

Statistics Netherlands

The Hague/Heerlen, 2009

Explanation of symbols

.	= data not available
*	= provisional figure
x	= publication prohibited (confidential figure)
–	= nil or less than half of unit concerned
–	= (between two figures) inclusive
0 (0,0)	= less than half of unit concerned
blank	= not applicable
2007-2008	= 2007 to 2008 inclusive
2007/2008	= average of 2007 up to and including 2008
2007/’08	= crop year, financial year, school year etc. beginning in 2007 and ending in 2008
2005/’06–2007/’08	= crop year, financial year, etc. 2005/’06 to 2007/’08 inclusive

Due to rounding, some totals may not correspond with the sum of the separate figures.

Publisher
Statistics Netherlands
Henri Faasdreef 312
2492 JP The Hague

Prepress
Statistics Netherlands - Facility Services

Cover
TelDesign, Rotterdam

Information
Telephone .. +31 88 570 70 70
Telefax .. +31 70 337 59 94
Via contact form: www.cbs.nl/information

Where to order
E-mail: verkoop@cbs.nl
Telefax .. +31 45 570 62 68

Internet
www.cbs.nl

ISSN: 1572-0314

6008309021 X-10

Improving the understanding of innovation by using test techniques

Vincent Fructuoso van der Veen[1]

Abstract

Scaling or test techniques may enlarge our knowledge of innovation dimensions. Such techniques include Reliability Analysis, Factor Analysis and Optimal Scaling, such as HOMALS, all which can be used to combine existing innovation indicators into new, reliable and valid innovation constructs. Scaling techniques offer an alternative methodology as well to make such new combinations with CIS microdata. The traditional way of calculating 'innovators' makes it impossible to calculate directly the correlation of innovators and 'R&D expenditure', since only innovators report 'R&D' and all innovators are treated in the same way. Scaling innovators does enable calculating this correlation: r = 0.34 in the Netherlands.

Keywords: innovation indicators, CIS, microdata, scaling, innovation modes, factor analyses, HOMALS, optimal scaling, OECD, NESTI.

1. Introduction

At the meeting of National Experts on Science and Technology Indicators, in Paris in June 2007 (NESTI 2007), M. Frenz and R. Lambert (UK, Department of Trade and Industry, now: DIUS) presented a paper (DSTI/EAS/STP/NESTI(2007)17) on the subject *Indicators of Non-Technological Innovation*. The OECD set up a microdata project in which Statistics Netherlands is participating. The objective of this specific microdata project (*Modes of innovation, technological and non-technological innovation*), which is led by the UK, is to improve our understanding of innovation. The present report builds further on the UK work, exploring different modes of innovation and presenting Dutch results. Including new combinations of existing innovation indicators in CIS (Community Innovation Surveys) may help to achieve this goal.

2. Considerations and approach

1. The definition of innovation in the *Oslo Manual* changed several times between CIS1 to CIS4.[i] As a result the core questionnaire also changed. There used to be a strong emphasis on technology, representing 'product and process innovation', but later the word 'technology' disappeared from the CIS. After that, the definition of innovation was extended to include 'organisational and marketing innovation'. These constructs reflect 'non-technological aspects of

[1] Statistics Netherlands, e-mail: va.fructuosovanderveen@cbs.nl

innovation'. As a result, statistics and our understanding of innovation have changed;

2. Innovation is difficult to measure, because no direct measurement of it exists. Using a survey in which indicators refer to definitions of innovation is used as an alternative;

3. New combinations of existing indicators could improve our current understanding of innovation, although it is unclear exactly which combinations are advisable. Also, it is unclear whether proposed new combinations would perform better than existing CIS indicators. In short, how could a search for new combinations be managed?

4. Scaling and test techniques may clarify these two problems. Test and scaling techniques include Reliability Analysis (RA), Factor Analysis (FA) and Optimal Scaling such as HOMALS; of course, validation is needed too;

5. So, analysing the properties of the CIS microdata with the aid of these techniques will indicate the direction in which a search for new combinations of existing indicators would be advisable. However, it is unclear which existing indicators are the best candidates for analysis for the purpose of this project. As 'product', 'process', 'organisational' and 'marketing innovation' are the core of innovation, these constructs will be analysed;

6. If any of the four types of innovation definitions is an empirically deviant construct, this will lead to finding the first part of a new combination. 'Deviant' means that the empirical results of the analysis contradict expectations based on existing definitions;

7. Reliability refers to the consistency of measurement, validity to the meaning of the construct. HOMALS is a FA for nominal data. FA is a structural model based technique. It reduces data in a meaningful way and is often used in constructing surveys. This may all improve the understanding of innovation;

8. The four types of innovation are measured by ten indicators. So, FA on these ten indicators should result in four factors, which in turn represent the four constructs. If not, this knowledge will show the direction in which improvements should be sought; it will provide the first part of a new combination of indicators to be recommended. If a theoretical construct, such as process innovation, does not consist of one dimension, in empirical terms, this means that corresponding indicators are invalid or the definition is invalid. In this case, for example, all indicators should measure process innovation. Finding a factor which explains one construct means that the construct consists of one dimension.

9. HOMALS will lead to a visible presentation of how well the four types of innovation fit in the technology dimension. First, this may answer an important question: Is the typology 'technological – non-technological innovation' oversimplified? Second, if a construct, or a specific indicator, does not fit in this typology as it is expected to, this knowledge provides the first part of an advisable new combination as well. It would not be a desirable sign if, for example, the process innovation indicators are in the non-technological part of a technology dimension, if such a dimension could be found at all, since it is

considered to be a technological kind of innovation. Such a distinction is oversimplified on the one hand, but still indicative the other.

10. Lastly, RA will indicate how reliable the measurement of innovation is, and this may also lead to advisable combinations. Sufficient reliability of a construct means that the supposed dimension is measured consistently. RA treats the constructs, for the moment, as ordinal scales. This means that firms which reported, for example, all three indicators of 'process innovation', are innovating their processes to a greater extent than firms which reported just one. In an ordinal scale it is irrelevant how much greater this extent is (otherwise it would be an interval measurement). So, the least kind of information is used: three is more than one, and how much more does not matter, at least in data technical terms.

11. However, after finding the first part of an advisable new combination of existing indicators, if at all, it is still unclear what the second part should be. Combining the first part with other existing indicators is advisable if this leads to a combination that could be explained by one factor. So, in that case it is certain that adding new indicators to the first part will lead to a new one-dimensional construct. This method prevents indicators being added randomly or based on a whim. Instead, it uses empirical characteristics. Secondly, the emerging new factor should not explain other types of innovation as well. This requirement prevents the addition of indicators which may all be part of different dimensions. Thirdly, the proposed combination must be reliable. This means that indicators that make up the new constructed combination should correlate strongly. If they do not, they measure different aspects. Fourthly, the new combination must be valid. This means that the new construct must show relations with other indicators or constructs that can be expected theoretically. If not, it does not fit in an innovation theory.

12. Validity will be analysed by correlating the proposals with 'R&D expenditure'. The proposals are considered to have an added value if they correlate more strongly with 'R&D expenditure' than existing indicators, thus leading to an improved understanding of innovation.

13. Also, it is unclear whether innovation dimensions differ between sectors of economy (NACE) or size classes (firm size). If true, new combinations of existing indicators may differ for both;

14. The results of these analyses will be used as input for a proposal for new combinations of existing indicators. This report describes the issues raised above.

3. Summary of results

In the Netherlands the four types of innovation perform very well, i.e. as expected, except for the construct 'marketing innovation'. Firstly, analyses of CIS data reveal that the measurement of 'marketing innovation' is inconsistent. Secondly, HOMALS reveals that neither indicator of marketing innovation is a measurement of ‘non-technological innovation’. This is caused by the first indicator: ‘Changes to the design or packaging of a good or service’, which even has technological relevance for ‘Manufacturing’ and big firms (250+). Thirdly, no factor represents the construct. These observations make the measurement or the definition of ‘marketing innovation’ a candidate for improvement in the Netherlands. Therefore, the indicator ‘Changes to the design or packaging’ is the first part of an advisable new combination. After all, it is possible that a combination with other indicators would work very well.

As all other indicators of the four types of innovation fit in the technology dimension very well, this typology is still very useful in the Netherlands. Also, factors representing the other constructs do exist.

In the Netherlands, the properties of data show that it is advisable to combine ‘New design or packaging’ - the first part - with the indicator ‘New goods’ (product innovation of goods). This new combination refers to ‘classical innovation’. It makes empirical sense to combine this first indicator of ‘marketing innovation’ with the indicator ‘New goods’. ‘Classical innovation’ views ‘changing design and packaging’ – of goods, in most cases – and ‘innovation in goods’ as part of one technology-based dimension.

This combination is empirically quite satisfactory. However, ‘classical innovation’ is improved when other indicators - closely related to new goods - are added. This leads to the final proposal, which satisfies all the demands stated in this report. It recommends combining the indicators ‘New goods’, ‘New to the market’, ‘Market introduction of innovations’, ‘Intramural R&D’ and ‘Patent apply’. This combination consists of one factor which does not disturb the other found factors of innovation. So, this extra demand assures that the new combination does not interfere with existing definitions of innovation. ‘Classical innovation’ is furthermore a consistent measurement, which seems valid, since its correlation with ‘R&D expenditure’ is quite strong. ‘Classical innovation’ distinguishes innovators in goods, performing intramural R&D – which has resulted in first entries on new markets and patenting the innovation – from firms which do not perform these activities. The new combination can be used as an ordinal scale measurement as well.

The second proposal is scaling. Scaling of innovation indicators leads to ordinal scales. These scales use existing information in a better way than the ‘traditional method of combining indicators’. The traditional method combines indicators, using or/or statements. Scaling on the other hand adds indicators, after which the sum is divided by the number of items. The difference is that the latter uses the information of the number of reported innovation types. The traditional method uses only the information in terms of conditions. The proposed scales are the ‘innovator scale’, the ‘innovation scale’, the ‘process innovation scale’, the ‘organisational innovation scale’ and the ‘classical innovation scale’. These scales are allowed. A ‘marketing scale’ and a ‘product innovation scale’ are not allowed. The fact that the latter is not reliable is acceptable, since goods and services are two different dimensions. Whether scales are allowed depends on their reliability and whether a separate factor for them (one way of validity) can be found. As the scales are ordinal, the different indicators do not need to be weighed: this is only required if the scales are used as interval scales. Although this is possible, it would be complex. With ordinal scales only the number of reported items is used, not their supposed relevance. Of course, such weights would be possible, but proposals on which weights to use for which indicators should be well founded.

The ‘innovator scale’, for example, improves our understanding of innovation. The traditional way of calculating ‘innovators’ makes it impossible to calculate directly the correlation of innovators and ‘R&D expenditure’, since only innovators report ‘R&D’ and all innovators are treated in the same way. Scaling innovators does enable calculating this correlation: r = 0.34 in the Netherlands.

So, innovators are firms which report, at least, one of the six indicators of ‘product and process innovation’ (5) and ‘ongoing or abandoned innovation activities’ (1). The traditional calculation of innovators leads to two groups: “0: no innovator” and “1: innovator”. However, the ‘innovator scale’ sub-divides the group of innovators by the number of reported innovation types. It distinguishes innovators which have realised, for example, just one type of innovation from innovators which have realised, for example, four. In this way the scale is a measurement of renewal, of how innovative a firm really is. Obviously, the latter firm in the example innovates to a greater extent. This way of treating innovators firstly enables the calculation of the correlation with ‘R&D’ directly, making it possible to answer an important policy and science related question: “What is the relation between innovators and R&D?”: the more innovators spend on ‘R&D’, the more innovative they are’. Secondly, it does not treat all innovators in the same way by putting them all in one box, ignoring the reported number of innovation types.

The next chapters will go into more detail on the topics: definition of the problem, basic correlations, reliability, factors of innovation, role of NACE and firm size, classical innovation, validity issues, application of scaling innovation indicators. This paper ends with a discussion.

4. Definition of problem

Definitions of innovation have changed several times, as have indicators in the several innovation surveys (CIS). Indicators measuring ‘organisational and marketing innovation’ were added to indicators of ‘product and process innovation’. The latter can be regarded as ‘technological innovation’ indicators, although the word ‘technological’ is no longer used in the innovation survey. Indicators of organisational and marketing innovation reflect ‘non-technological’ innovation.
The distinction between the two types may be oversimplified. They are probably related, and both ‘technological’ and ‘non-technological’ knowledge or activity, may be part of innovation of any kind. Furthermore, the way in

which innovations are realised may differ between sectors of economy and size classes.
To fill this gap in our understanding of innovation, the OECD set up its Microdata Project in which the Netherlands is participating; this report is limited to results of the Netherlands alone.

The objective of this report – as of the overall project – is to improve our understanding of innovation by making new combinations of existing innovation indicators. This objective, however, raises two questions:

1. *Which combinations of existing indicators are advisable?*
2. *What criteria should be used to determine whether the new combinations have an added value?*

The first question could be answered in many ways. Because the project is based on microdata, empirical features of existing indicators are used to determine which indicators lead to deviant results, i.e. results that do not empirically fit in existing innovation definitions. This will direct the search for new indicator combinations. This question (1) in turn raises two new questions:

A) What empirical features should be used?
B) Which existing indicators are the best candidates to be examined?

Obviously, the understanding of innovation will be improved most by examining indicators that measure 'product', 'process', 'organisational' and 'marketing' innovation. Standard empirical features are reliability and validity. So, if any construct is less reliable or valid, then a specific direction will emerge. As validation is a theoretical issue, analysing reliability is more important, at least for our purpose.
Analysing reliability is possible if the constructs are treated as scales for a moment. This is justified as innovation types may be ranked. For example, a firm reporting three types of 'organisational innovation' actually innovates the organisation more than a firm which reports just one type. Just how much more is irrelevant in such a scale.
So, reliability says the correlation between the indicators of a specific construct must be strong. Not as strong as possible, as the indicators would then measure exactly the same thing; but not too weak either, as it would then be unreliable. Furthermore, the indicators of one construct should correlate more strongly with each other than with indicators of other innovation constructs. These requirements follow from test theoretical principles, to be used when creating surveys.

The second question (2) raises the issue of which criteria should be used to decide whether new combinations of existing indicators are better than the existing indicators alone. First of all, again if the reliability and validity of proposed new combinations of existing indicators are better, they are considered to have an added value. Secondly, the new combination should be explained by one factor. Thirdly, this factor should not explain other types of innovation as well. If these conditions are met, the proposals are empirically justified.

Method
The questions will be answered with the aid of the following test and scaling techniques:

- Correlations;
- Reliability Analysis (RA);

- Factor Analysis (FA) and HOMALS;
- Factor-Analytic Scale Construction.

5. Basic indicators of innovation and how they interrelate

CIS4 (reference period 2002-2004) comprises ten indicators of innovation type (table 1). ‘Product’ and ‘process’ innovation are defined as ‘technological innovation’, ‘organisational’ and ‘marketing’ innovation as ‘non-technological innovation’.

Table 1. Types of innovation

Indicator	Code	Construct
		Product innovation
Prod1	INPDGD	New or significantly improved goods
Prod2	INPDSV	New or significantly improved services
		Process innovation
Proc1	INPSP	New or significantly improved methods of manufacturing or producing
Proc2	INPSLG	New or significantly improved logistics, delivery or distribution
Proc3	INPSSU	New or significantly improved supporting activities for processes
		Organisational innovation
Org1	ORGSYS	New or significantly improved knowledge management systems
Org2	ORGSTR	A major change to the organisation of work
Org3	ORGREL	New or significant changes in relations with others
		Marketing innovation
Mar1	MKTDES	Significant changes to the design or packaging
Mar2	MKTMET	New or significantly changed sales or distribution methods

Correlations

The correlation matrix of these innovation indicators is presented in table 2. The correlations are a first check of the question whether relations between types of innovation exist. The data are weighted. [ii]

Table 2. Correlations of types of innovation in the Netherlands (CIS4)

	Prod1	Prod2	Proc1	Proc2	Proc3	Org1	Org2	Org3	Mar1	Mar2
Prod1	1.00									
Prod2	0.27	1.00								
Proc1	0.38	0.26	1.00							
Proc2	0.20	0.30	0.36	1.00						
Proc3	0.29	0.30	0.38	0.43	1.00					
Org1	0.16	0.22	0.23	0.23	0.33	1.00				
Org2	0.17	0.20	0.19	0.18	0.26	0.47	1.00			
Org3	0.14	0.24	0.18	0.18	0.19	0.34	0.43	1.00		
Mar1	0.32	0.18	0.25	0.18	0.20	0.21	0.19	0.22	1.00	
Mar2	0.11	0.19	0.11	0.19	0.19	0.27	0.26	0.26	0.28	1.00

All correlations are significant at the 0.01 level (two tailed), N = 5 927.

What are the test and scaling requirements? First, all indicators should correlate and all should have a positive sign, as they all measure innovation. [iii] If they do not, they should not be part of a measurement of innovation definitions. Secondly, the indicators of the four different constructs should not correlate in the same way: the correlations of indicators within the same construct should be stronger than those with indicators of other constructs. Otherwise it is unclear why an indicator is part of a specific construct. The correlations within one construct are boxed in table 2.

Observations and conclusions:

1. All the ‘process and organisational innovation’ correlations within one construct are stronger than their correlations with indicators of the other constructs. They are marked green in table 2.
2. $r_{\text{proc1(New manufacturing or producing)-prod1 (New goods)}}$ is larger than expected. On the other hand, the strong correlation between ‘process and product innovation’ is still consistent with the idea of ‘technological innovation’ (both are part of ‘technological innovation indicators’). Secondly, this correlation is not stronger than all correlations of ‘process innovation’ within the construct. This pattern is marked yellow;
3. $r_{\text{prod1 (New goods)-mar1 (New design or packaging)}}$ does not meet test requirements, as this correlation is stronger than $r_{\text{mar1-mar2}}$ is. This observation is marked red in table 2. This is a first indication that it not clear why indicator ‘Mar1’ is part of the construct ‘marketing innovation’. Secondly, the construct ‘marketing innovation’ has no other (third) indicator to measure it, which may pose a problem if the reliability of such a scale and the validity of the construct are insufficient as well, and if no factor can be found for ‘marketing innovation’;
4. Indicator ‘Prod2’ (New services) ‘behaves’ unusually as well: new services correspond better with ‘Proc2’ (New logistics, delivery or distribution) and Proc3 (New supporting activities). This type is marked yellow as well;
5. The correlations are not corrected (yet) – for example by computing (semi) partial correlations – so some caution in interpretation is needed. [iv] However, the picture that emerges will be confirmed, later on in chapters 3 and 4.

6. Reliability of innovation constructs

A construct should preferably consist of one single dimension. All indicators that are part of the measurement of a construct should measure that dimension consistently. Analysing reliability (RA) checks this. If a construct (or definition) consists of two dimensions, more than one indicator should reflect each dimension. Both dimensions should consist of different indicators. If they do not, clustering indicators is useless.
Indicators should therefore measure a construct consistently. If they do not, the construct is unreliable. This could mean that the construct consists of more than one dimension, in which case it should be reconsidered. This may require breaking it down into several dimensions, which will in fact change the definition. However, it could also mean that only one of the indicators does not fit in the construct. In that case, the solution is replacing the indicator concerned by another, at least if this leads to an improved measurement.

So, the four constructs of innovation should lead to a one-dimensional measurement, which is reliable and valid. This chapter deals with reliability, treating the constructs, for the moment, as ordinal scales.
When conducting RA, or any other scaling technique, the term ‘item’ is preferred to the term ‘indicator’. Table 3 displays the supposed scales, the number of items that are part of the scales and Cronbach’s Alpha (α), which is the predicted proportion of variance between ‘observed’ and ‘true score’, thus providing information on the consistency of measurement (squared correlation). The value of α decreases if items measure a construct inconsistently. [v]

The higher the value of α, the lower random error is. However, a perfect reliability (α = 1) does not exclude bias. The higher the value of α, the higher the probability is of obtaining a valid measurement. [vi]

Table 3. Scale properties of the four innovation constructs in the Netherlands (CIS4)

Construct	# items	α [vii]
Organisational innovation	3	.67
Process innovation	3	.65
Marketing innovation	2	.44
Product innovation	2	.42
All	10	.76

Observations and conclusions:

1. All separate RA outcomes show that the deletion of no item increases α; of course, this can only be tested for 'process and organisational innovation', as the other two supposed scales consist of only two items;
2. Table 3 reveals that the items of 'process and organisational innovation' are consistent measurements. An α between .60 and .80 is reasonable, certainly taking into account the small number of items. For example, an α of .67 of the 'organisational innovation scale' means that the correlation between the 'observed' and the 'true organisational innovation score' is √ .67 = .82. So, both would be reliable scales, which could be used, if desired. [viii]
3. Unfortunately, the items of 'marketing innovation' do not measure the supposed dimensions consistently, at least not in the Netherlands. This means that these items cannot be scaled;
4. The low RA of the construct 'marketing innovation' does not necessarily mean that the items do not measure one dimension. Both could be a measurement of one single dimension in which one item performs badly. The following chapter will answer the question of whether the construct is one-dimensional or not. For example, RA does not clarify which of the two 'marketing innovation indicators' is problematic, but HOMALS will;
5. Only two items are considered to indicate 'marketing innovation', which is obviously a very small number to test scale properties. As a result, it is known beforehand that the correlation between the two indicators needs to be very strong to obtain a reliable measurement. Table 1 already revealed that the inter-item correlation was r = 0.28. Although significant, this is insufficient. However, three items is a small number as well and the first two scales did lead to sufficient reliability;
6. The RA shows that a search for new combinations of indicators in the direction of 'marketing innovation' would be useful in the Netherlands;
7. Scaled 'Product innovation items' seem to 'behave' in the same way as scaled 'marketing innovation items'. However, this is logical, because it is easy to understand that 'goods' and 'services' are two dimensions. The idea of scaling them is based on the fact that a firm reporting 'New goods' is less 'product innovative' than a firm reporting both 'New goods' and 'New services'. However, this does not lead to a reliable scale. In CIS, both indicators are labelled as 'product innovation', but it is unclear whether this label adds any value. The next chapter will answer this question empirically; [ix]
8. The last RA result of table 3 (bottom row) is a test to see whether all items add up to one reliable measurement. The resulting scale is indeed more reliable, but what does it measure? The scale could called an 'innovation scale'. It measures the number of reported types of innovation. This is a new combination of indicators with improved reliability, compared with the existing

indicators. See the following chapter and chapter 6 for further elaboration of this scale. [x]

The conclusion is that 'process' and 'organisational' innovation measure one dimension consistently, while 'marketing' and 'product' innovation do not. As table 2 demonstrated that indicator 'Mar1' (New design or packaging') correlates more strongly with indicator 'Prod1' ('New goods') than with indicator 'Mar2', the problem is probably in indicator 'Mar1'. It would be wrong to argue that 'marketing innovation' consists of two dimensions instead of one, and that both indicators measure precisely these dimensions. [xi]
So, this chapter suggests that indicators 'Mar1' and 'Prod1' are the first two parts of recommended new combinations of existing indicators. It is less imperative to search in other directions for new combinations. However, we still need to conduct HOMALS and the FA as well, as these techniques will demonstrate why 'Mar1' is the problem.

7. Factors of innovation

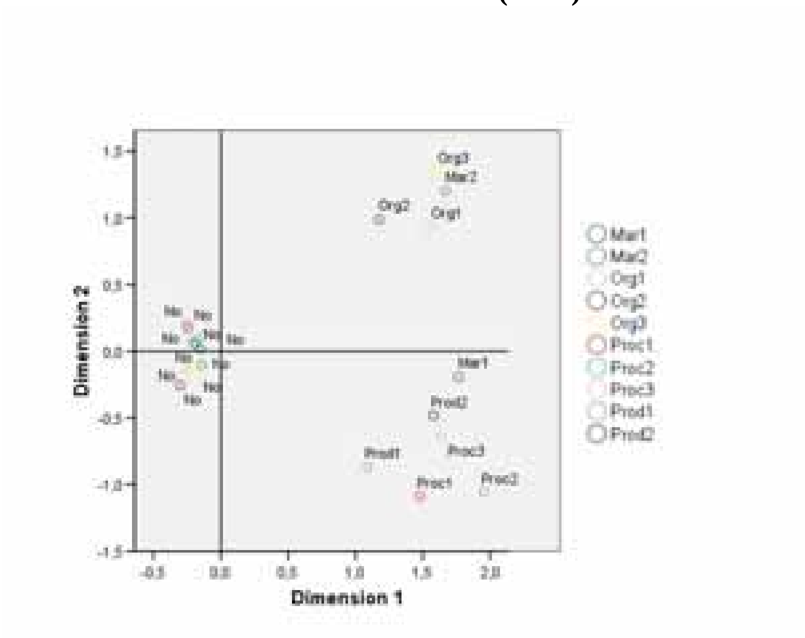
What are factors of innovation? In this chapter we shall answer this question, using factor analysis on the ten indicators of types of innovation in the Netherlands. First, Optimal Scaling will be used to learn more about the innovation definitions. This will improve the understanding of the FA conducted later on. The purpose of these analyses is to find out more about the direction in which a search for new combinations of existing indicators is advisable.

HOMALS of the ten innovation items
HOMALS (HOMogeneity through Alternating Least Squares) is based on an idea very similar to FA. It is, however, used to analyse nominal data, while FA is used to analyse interval variables. First HOMALS transforms the data to obtain the best category quantifications; then it conducts a classical PCA (Principal Components Analysis). [xii]
HOMALS allows a simultaneous clustering of both firms and variables (categories). The responses are grouped into similar categories and into similar firms at the same time, thus making a separate cluster analysis superfluous. In the following chapter NACE and Firm Size are included in the HOMALS, which will demonstrate this benefit even more.

Homogeneity can be visualised as relative distances in a Euclidean space. HOMALS is best at plotting two-dimensional solutions. However, this does not mean that only two dimensions exist; more may exist. HOMALS plots category-points, in which all responses of firms, on the indicators, are summarised as distances. The larger the distance between two categories, the less homogeneous they are. So, homogeneous indicators (categories) should be plotted close to each other.
What then are the requirements for HOMALS? Firstly, 'technological innovation indicators' (categories) should be plotted closer to each other than to 'non-technological indicators'. Secondly, indicators which are part of one construct should be plotted closer to each other than to indicators that are part of other constructs.
In the category plot on the next page, the categories are labelled with the indicator names, making them easier to recognise. However, it should be kept

in mind that these labels actually represent the categories of ‘Yes’ answers to the ten types of innovation.
In addition to the category plot, HOMALS plots variables and object scores. Such a variable plot is shown when ‘NACE and Firm Size’ are included. [xiii]
In figure 1, the two categories of each indicator are displayed. The indicator names are given on the right-hand side of the figure. For example, the ‘purple marks’ belong to indicator ‘Org2’ (New organisation of work). The purple mark indicating ‘No’ reflects the vector of the ‘No’ answers to indicator ‘Org2’. Similarly, the purple mark, indicated by ‘Org2’, reflects the ‘Yes’ answers to the indicator ‘Org2’. It is important to remember that HOMALS treats all indicators separately, not as part of a scale as in the previous chapter.
The first dimension in the HOMALS solution describes 33 percent of variance, and the second 13 percent, which gives a total of 23 percent. Note that later on the FA will explain more variance, as both Eigenvalues are added. [xiv]

Figure 1: Plot of categories of types of innovation in the Netherlands (CIS4)



How should figure 1 be interpreted? Observations and conclusions:
1. The first dimension distinguishes ‘Yes’ and ‘No’ answers. All ‘No’ categories of innovation types are located close to the origin. The first dimension can thus be interpreted as a general ‘innovative or not’ status. Note that this is the ‘innovation scale’, which proves its relevance in passing (see chapter 3, observation 8). In retrospect, it is obvious that the indicators had to discriminate between firms along this axis first. First we have to know whether firms are innovative or not – using these indicators – only after that does it become relevant whether the innovations are ‘technological’ or ‘non-technological’ – and only after that again is it relevant which of the four types of innovation is reported;
2. The group in the top right-hand corner of figure 1 corresponds to all positive categories of ‘organisational innovation indicators’, together with the second indicator of ‘marketing innovation’. The group in the bottom right-hand corner corresponds to all positive categories of ‘process and product innovation indicators’ and to the first indicator of ‘marketing innovation’;

3. The categories in the top right-hand corner all belong to 'non-technological innovation indicators' except for the positive category of indicator 'Mar1' (New design or packaging). The RA (chapter 3) led to the conclusion that 'marketing innovation' is not a consistent measurement. This could explain why the positive category of indicator 'Mar1' is in the 'wrong' group;
4. This second group includes all positive categories of 'technological innovation indicators', and of indicator 'Mar1'. So, the second dimension discriminates between 'technological' and 'non-technological' innovation very well: it can be interpreted as the technology dimension. Figure 1 proves that the typology is applicable in the Netherlands. All constructs and indicators fit in it, except for indicator 'Mar1'. This indicator seems to measure some technological content;
5. So the Dutch CIS4 data reveal that the construct 'marketing innovation' is not fully 'non-technological'. In empirical terms, it would be a 'non-technological' construct if indicator 'Mar1' was in the top right-hand corner of figure 1;
6. An improved reliability of a supposed scale 'marketing innovation' (chapter 3) would lead to 'migration' of indicator 'Mar1' to this desired position. This indicator should also migrate to a position closer to indicator 'Mar2' than to the three 'organisational innovation' indicators. Such a migration would lead to a perfect fit of the technology dimension with respect to these indicators.

Now we know why 'marketing innovation' is an inconsistent measurement or construct (see chapter 3). The first indicator of 'marketing innovation' is problematic, i.e. the definition of the construct is, as the Oslo Manual describes it as a 'non-technological' construct, or at least it is still treated thus by all EU countries.[xv].
It also makes sense of indicator 'Mar1' being homogeneous with the 'technological innovation indicators': 'changing design or packaging' means changing aspects of products and 'product innovation' is on the technological polar.
Figure 1 could help other countries answer the question of whether the technology typology is applicable in their country (see Appendix for the SPSS syntax). Furthermore, it helps detect deviating constructs or indicators in other ways as well. The second dimension should always be the technology dimension, at least if this typology makes sense.

FA of the ten innovation items
Do we still need to conduct an FA? [xvi] Yes: a factor solution always explains more variance than a HOMALS solution, since in an FA the Eigenvalues are added. The HOMALS plot served only one purpose: a scan of how the indicators perform on the important technology dimension, in a visual to ease interpretation. [xvii]
However, a third dimension may still exist, which is another reason for conducting the FA. This third factor cannot be a factor which explains 'marketing innovation', since indicator 'Mar1' is in the wrong HOMALS group. So, it can already be concluded that the variance of both 'marketing innovation' indicators cannot be explained by one, unique factor.
The FA should therefore find three factors: one explaining the 'process innovation' indicators, one the 'organisational innovation' indicators and one the 'product innovation' indicators. However, a factor discriminates firms on one dimension and in chapter 3 it was argued that goods and services could be considered as being different dimensions. In CIS, goods and services are

clustered under the label 'product innovation'. However, using the same term for goods and services does not make the construct one-dimensional. If no factor is found for this construct, no empirical reasons exist to cluster the two as 'product innovations'. However, HOMALS showed that clustering both 'product innovation indicators' as 'technological indicators' is reasonable, so it is not a real problem if we do not find a separate factor for both indicators: it would just mean that the label 'product innovation' does not imply added value, at least in empirical terms.

This means that at least two factors should be found: one unique factor for 'process innovation' and one for 'organisational innovation'. If this is the case, this could be considered as a (first) validation that both constructs are one-dimensional. [xviii] This would mean that this is not an advisable direction to search for new combinations of existing indicators.

The FA can now be generated and interpreted with more knowledge about the separate indicators and constructs. However, HOMALS is not a 'structural model-based analysis', like FA. This means that factors found in an FA cause the variance of the ten innovation definitions. HOMALS uses the term 'dimension', thus not assuming theoretical cause. Likewise, PCA uses the term 'component'. The term 'factor' in FA refers to this difference (in a plot of the corresponding models the arrows would have other directions).

The two-dimensional HOMALS solution (see figure 1) appears to be the second best way to describe the four constructs defining innovation types. The best way is a three-factor solution. The first two factors are of course the same as the two HOMALS dimensions. The third factor is the same as presented in table 4 (below), which explains 'New designs or packaging' (indicator 'Mar1') and 'New goods' (indicator 'Prod1').

The three latent variables explain 55 percent of total variance of the ten indicators. This is the unrotated solution. [xix] However, table 4 shows a rotated solution.

Table 4. Rotated three-factor solution of the innovation constructs in the Netherlands (CIS4)

Item	F1	F2	F3
Prod1 (New goods)			0.66
Prod2 (New services)			
Proc1 (New methods of manufacturing)	0.43		
Proc2 (New methods of logistics, delivery or distribution)	0.67		
Proc3 (New supporting activities)	0.63		
Org1 (New knowledge management systems)		0.58	
Org2 (New organisation of work)		0.73	
Org3 (New relations with others)		0.59	
Mar1 (New design)			0.44
Mar2 (New sales or distribution methods)			
Eigenvalue (original)	3.25	1.29	0.98
Percentage of variance explained (unrotated)	33	12	10

Pattern matrix. Values less than 0.40 are suppressed. Extraction method is principal axis factoring. Rotation method is OBLIQUE; $r_{f1-f2} = 0.47$, $r_{f1-f3} = 0.51$ and $r_{f2-f3} = 0.41$. [xx]

The OBLIMIN (OBLIQUE) was applied is because the 'OBLIMIN rotation' provides a better fit with reality, as the factors should correlate. In the most common 'VARIMAX rotation', factors are orthogonal. This means that factors do not correlate. However, they should, as the correlation matrix (see table 2) demonstrated that all indicators correlate. This means that all constructs correlate. The goal of this FA is to find a separate factor for each construct, that can be interpreted as the construct. Finding such factors would mean that each

latent variable would explain only the variance of the indicators which are part of a specific construct, and not of others. Since all innovation types correlate, such factors should correlate as well, as they represent the constructs. Furthermore, it would be strange if latent innovation factors are not correlated, as any kind of innovation is based on some sort of renewal or change. By definition therefore, such factors always have something in common. However, if the factors are not allowed to correlate (by using VARIMAX), it becomes impossible to find the factors we seek.
For example, a 'VARIMAX rotated factor', which explains only the 'organisational innovation' indicators, is not allowed to correlate with a factor which explains only the 'process innovation' indicators. So this would be a questionable restriction, because we already know that all belonging indicators correlate. Furthermore, organisational change may cause process changes in firms. So, all expected or found factors should correlate. [xxi]
Consequently the indicators cannot, and do not, all load on the first factor – as is the case in the unrotated factor solution – otherwise the factors would not explain only one construct. Therefore, the factors in table 4 differ from the HOMALS dimensions of figure 1, in which the solution is unrotated as well. [xxii]

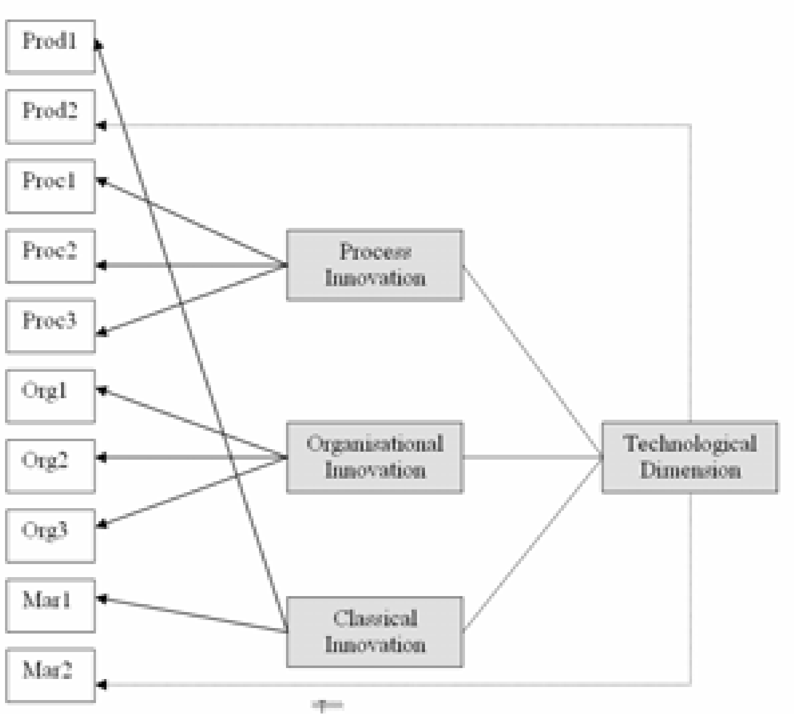
Observations and conclusions:
1. The first factor can clearly be interpreted as a latent cause of the 'process innovation' indicators, the second as causing the variance of the 'organisational innovation' indicators. So both constructs may be treated as scales (see chapter 3), both constructs fit well in the technology dimension and both constructs are one-dimensional;
2. Both factors for 'process' and 'organisational' innovation are meaningful, rewritten versions of the technology dimension;
3. The third factor indicates 'New goods' and 'Changes in design or packaging'. This factor distinguishes firms reporting both types of innovation from firms reporting neither type. This is an important structural aspect of the data. The question is how the third factor can be interpreted. This factor certainly has something to do with goods and aspects of goods, such as design and packaging. The factor actually 'says' that both indicators are too interrelated to be treated separately.

Table 4 shows first that, in the Netherlands, 'marketing innovation' is not a one-dimensional construct. The table confirms again the conclusions drawn so far about 'marketing innovation' Secondly, at the same time the table displays the second part of the combination, which is a solution for this problem: it is empirically advisable to combine the indicator 'New design or packaging' with the indicator 'New goods'. This new combination changes two inconsistent measurements into one meaningful, new construct. FA proves that 'Innovation in goods' and 'Changing design or packaging' (of goods, in most cases) are part of one dimension.
Notice that combining both indicators is an acceptable solution. However, it is only based on the ten indicators of innovation types. Chapter 6 will elaborate on this direction further using all CIS indicators. For the time being, the third factor could be interpreted as 'classical innovation'. Adding several correlated indicators to this new construct may increase the Eigenvalue of the third factor, without the other factors being disturbed. This could lead to a better proposal. Chapter 7 will deal with these proposals.
No factor was found for 'product innovation'; but this was not a strong requirement. 'New goods and services' are both placed under the label 'product innovation', however this label does provide added value, which was already argued above. The other labels do provide added value, as they can be given a

separate scale score, except for 'marketing innovation' – see chapters 3 and 7. So, it is advisable to split 'product innovation' into 'innovation in services' and 'innovation in goods', since the two refer to different dimensions. In test theoretical practice only indicators that are part of one factor could be labelled the same.
HOMALS and FA prove that indicator 'Mar1' (New design or packaging) measures some technological content. It could measure technological content for some groups of firms (see next chapter).

Modelling innovation
The results of HOMALS and FA lead to a model of innovation types in the Netherlands. The arrows represent factor loadings of the three found factors. The factors are in the middle of the figure and cause the variance of the indicators, which are on the left-hand side. On the right-hand side is the 'technology dimension'. This dimension is related to the factors (in a two-dimensional HOMALS solution this itself is a factor), the three factors are related and the indicators are related (their relations are not drawn with arrows). Of course, most of the variance of the remaining two indicators (indicator 'Prod2', New services – and indicator 'Mar2', New sales or distribution methods) is already explained by the three factors, but this did not result in loadings > 0.40 for either item – and only those arrows are plotted in figure 2. However, if the FA had not been rotated, or the HOMALS interpretation had been used to draw the model, both indicators would have been explained very well. To solve this, the technology dimension has two arrows as well. The correlations between indicators and factors are not displayed in the figure.
The structural model justifies two existing constructs – 'process and organisational innovation' – and uses one new construct ('classical innovation'). It can place factors in the context of the 'technology dimension' and it shows interrelations.

Figure 2: Structural model of types of innovation in the Netherlands (CIS4) [xxiii]



Additional remarks:

1. Also, the indicators 'Prod2' (New services) and 'Mar2' (New sales or distribution methods) have large loadings on respectively the first and second factor (0.29 and 0.39). The large loading of indicator 'Mar2' on the second factor means that this factor also explains this indicator, thus making it more 'non-technological'. The loading of indicator 'Prod2' on the first factor makes this factor more 'technological';
2. A plot of factor loadings of the first two (unrotated) factors of the FA leads to a picture that is very similar to figure 1. In fact, the HOMALS space is a plot of the factor loadings, named vectors. The difference is that HOMALS shows the role of the 'marketing innovation' indicators more clearly;
3. The 'OBLIMIN rotation' rotates the HOMALS dimensions (the *x*- and *y*-axes in figure 1), without necessarily rotating by 90° (not orthogonal). If we look at figure 1 again, and rotate the *y*-axis clockwise, we see that this results in larger vectors of the 'organisational innovation' indicators, thus leading to finding the second factor. The *x*-axis rotates clockwise as well, leading to larger vectors of the 'process innovation' indicators, which leads to finding the first factor. The third dimension, in which dispersion in figure 1 still exists, is the distance between indicator 'Prod1' and indicator 'Mar1'– which is the largest distance that remains to be explained, and therefore results in the third dimension. Distance within one quadrant in HOMALS usually implies another, third dimension;
4. Trying to stick to a two-factor solution is defendable, since the Eigenvalue of the third factor is < 1. Still, a three-factor solution is preferred, because the value approaches 1, because the 'scree criterion' supports this choice and because of the improved interpretation of the total solution.

8. NACE and Firm Size

This chapter will demonstrate whether, and how, innovation dimensions differ for groups of firms based on NACE ('Manufacturing', 'Services Sector' or 'Other Sectors') and firm size (firms existing of '10-50 employees', '50-250 employees' or '250+ employees'). Differences between these groups could improve our understanding of innovation dimensions.

If most of the variance of the ten indicators is explained by NACE, then the first unrotated factor of the factor solution – see previous chapter – would explain NACE as well. So, adding NACE and Firm Size in an FA is a way to scan homogeneity quickly. In the Netherlands, this leads to the observation that NACE is part of the second factor. Firm Size is not part of any factor. The unrotated factor solution is used, as the first thing we want to know is how the technology dimension differs for NACE and Firm Size.

Since no third factor explaining 'marketing innovation' was found in the total population, trying to find such a third factor for specific groups would not be a logical step. So, HOMALS, which plots two dimensions, will be used again. However, it should be kept in mind that a third factor may exist.

So, the technology dimension differs for NACE, but how? Also, Firm Size does not affect the 'three-factor solution', but does this mean that Firm Size is irrelevant? To be find answers to these questions is precisely the second reason for using HOMALS again. HOMALS displays factor loadings for categories, while FA can display factor loadings for variables only. So, FA cannot answer the questions of this chapter immediately. Thirdly, HOMALS makes a separate cluster analysis superfluous.

Adding NACE and Firm Size does not change figure 1 very much (see figure 3). [xxiv] Of course positions of all the categories have shifted slightly as new Optimal Scaling was conducted. Again, the 'Yes' answers to the ten types of innovation are labelled with the indicator names instead of the category names, to ease visual interpretation.

Observations and conclusions:

1. NACE and Firm Size categories are closer to the origin than to any other group with respect to both dimensions; [xxv]
2. The category '250+ employees' is displayed close to the positive categories of 'technological innovation' indicators. This means that the pattern of answers of big firms is very similar to the pattern of answers of 'technological innovators': big firms are 'technological innovators introducing new goods';
3. The vector of the category '250+ employees' is much higher on the first dimension than on the second: the fact that big firms are innovative discriminates them from other firms even more than the fact that they are 'technological innovators';
4. The category '10-50 employees' is plotted close to the origin, so this group of firms innovates relatively little, since corresponding CIS responses are quite similar to all 'No' answers to types of innovation. The category '50-250 employees' is located slightly further from the origin;
5. The categories 'Manufacturing' and 'Services Sector' discriminate on the second dimension. So 'Manufacturing' is a 'technological innovating sector', while the 'Services Sector' is more of a 'non-technological innovating sector'; [xxvi]
6. The category 'Manufacturing' is in the bottom half of the figure, 'closer' to the 'technological innovating group' than the category 'Services Sector' is to the 'non- technological innovating group'. So 'technological innovation' is more characteristic for 'Manufacturing' than 'non-technological innovation' is for the 'Services Sector'. [xxvii]

The plot of the discrimination measures, see figure 4, demonstrates that the (whole) variable NACE has a high discriminatory ability in just one dimension, the second, while most – but not all - other indicators discriminate in two dimensions.

The length of the vector indicates its significance. The vector of the variable Firm Size is short, because only one category leads to a higher vector (250+ employees). This category consists of relatively few cases in the Netherlands, which is why the (whole) variable was not detected with the FA.

So, conducting separate analyses for all three categories of Firm Size is unnecessary. However, innovation dimensions clearly differ for big firms (250+ employees) and NACE, thus making separate HOMALS for big firms and NACE ('Manufacturing' and the 'Services Sector') necessary. These plots can be found in the Appendix. The conclusions are:

1. In big firms, the technology dimension is not very clearly applicable. In 'Manufacturing' and the 'Services Sector', the technology dimension is very well applicable;
2. In big firms and 'Manufacturing', indicator 'Mar1' (New design or packaging) has even more technological content than in the total population. In the 'Services Sector', indicator 'Mar1' has some non-technological content. However, in this sector the indicator is still not homogeneous with indicator 'Mar2';
3. In big firms, the difference between 'Manufacturing' and the 'Services Sector' in CIS responses enlarges;

4. In big firms, the most important homogeneous group consists of firms in ‘Manufacturing’ which conduct classical innovation.

Figure 3. Plot of categories of innovation types, NACE, Firm Size, the Netherlands (CIS4)

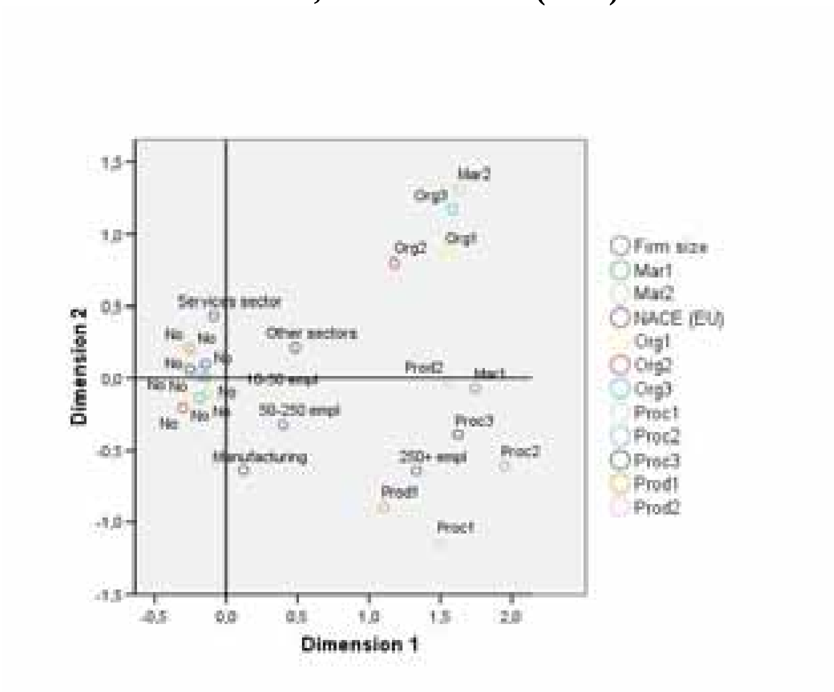
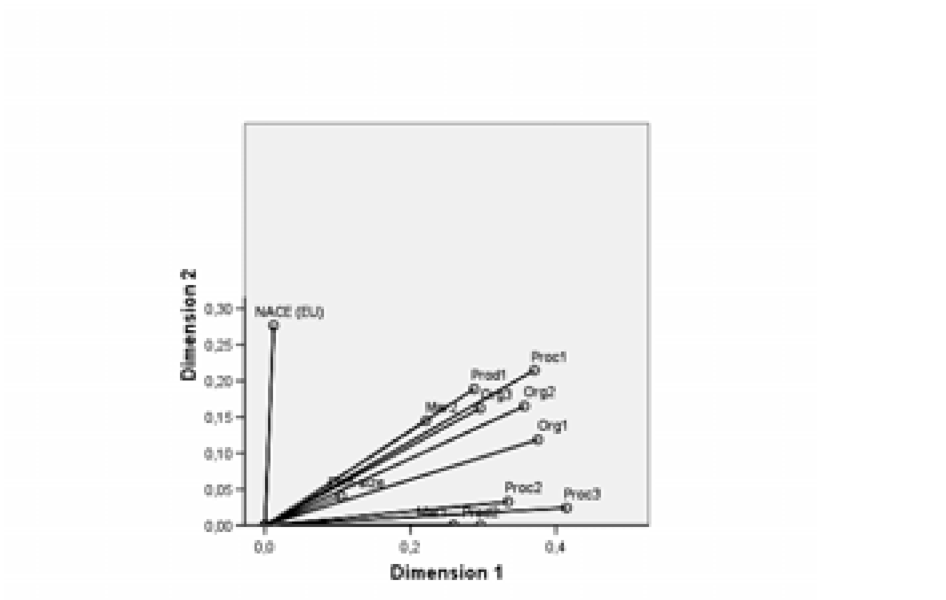


Figure 4. Discrimination measures of innovation types, NACE, Firm Size, the Netherlands (CIS4, plot of variables)



So, while all ‘innovation type’ indicators are more or less independent of the responding types of firms, indicator ‘Mar1’ is not. In the Netherlands, ‘marketing innovation’ measures not only ‘non-technological innovation’ inconsistently, it is even systematically at the ‘technological polar’. This chapter leads to the conclusion that ‘classical innovation’ is especially relevant for big firms (see following chapter).

9. Further details of 'classical innovation'

Chapter 4 revealed that the indicators of 'marketing innovation' do not measure one dimension. Instead, indicators 'Mar1' and 'Prod1' measure one dimension. This new combination indicated the direction in which a search for a new combination would be useful in empirical terms. The Eigenvalue of this direction, which could be identified with classical innovation, is 0.98 (see table 4). So, this new combination could be improved further, leading to an Eigenvalue > 1. This chapter will elaborate on this subject.
To complete the combination, the procedure of factor analytic scaling will be used. Firstly, we need to find all CIS indicators which correlate strongly with indicators 'Mar1' (New design or packaging) and 'Prod1' (New goods). Secondly, the selected indicators need to be factor analysed in the same way as they were in table 4. Thirdly, in the factor solution the selected indicators are only allowed to load on the 'classical innovation factor'. Fourthly, the interpretation of first two factors may not change: the added indicators are not allowed to disturb the factors for 'process and organisational innovation'. Fifthly, the new combination must constitute a reliable scale. Lastly, its validity must be proven. This latter demand is dealt with in chapter 7.

These strong demands make it difficult to find indicators which satisfy all of them. The indicators which correlate with indicators 'Mar1' and 'Prod1' are selected by a minimal value of their correlation with indicators 'Mar1' and 'Prod1'. The mean correlation of an indicator with both must be at least r = 0.20. The reason for this extra demand is the fact that the FA is based on correlations of indicators. So the best correlating indicators should be selected. However, just immediately selecting the best indicators may not lead to the optimal factor solution, as the indicators suffer from multicollinearity. This means that after (semi) partial correction weak correlations may become stronger, and strong correlations may become weaker. On the other hand, as it is not possible to use all indicators in the FA, an arbitrary criterion is unavoidable.
The indicators which correlate with indicators 'Mar1' and 'Prod1' are: 'NEWMKT', 'NEWFRM', 'TURNMAR', 'TURNIN', 'RRDIN', 'RRDEX', 'RMAR', 'RRDINX', 'RRDEXX', 'ERANGE', 'PROPAT', 'PRODSG', 'PROTM' and 'PROCP'.
The indicators which correlate with the indicators 'Mar1' and 'Prod1' with a minimal mean value of 0.20 are: 'NEWMKT', 'NEWFRM', 'TURNMAR', 'TURNIN', 'RMAR', 'RRDINX', 'ERANGE', 'PROPAT' and 'PROTM'.

So these nine indicators are the best candidates to be added to the existing FA. However, they allow 511 combinations of FAs to be performed, which is obviously quite laborious. On the other hand, if all nine indicators are added to the FA, more than three factors will appear and, more importantly, the factors of 'organisational and process innovation' will be disturbed. Actually, the third factor needs only one, maybe two extra indicators, since the Eigenvalue is already near value 1. This means the best way to proceed is to first add one of these nine indicators.
After nine trials, indicator 'RMAR' (Market introduction of innovation) appears to be the best fitting indicator. So this indicator is added to the rest, after which the search for the second indicator continues. However, adding any other indicator to the FA diminishes the result. Also, any other combination of two indicators does not change the conclusion. Adding only indicator 'RMAR' is therefore the best solution.

Indicator ‘RMAR’ measures ‘Activities for the market introduction of your new or significantly improved goods and services, including market research and launch advertising’. The factor solution demonstrates that this activity is closely related to innovation in goods. This fact makes this factor more of a ‘classical innovation factor’: it discriminates firms which realise new goods, adjust design and packaging of existing goods, and market these goods from firms which do not (see table 5).

Table 5. Rotated three-factor solution for ‘classical innovation’ in the Netherlands (CIS4)

Item	F1	F2	F3
Prod1 (New goods)		0.62	
Proc1 (New methods of manufacturing)			0.30
Proc2 (New methods of logistics, delivery or distribution)			0.60
Proc3 (New supporting activities)			0.38
Org1 (New knowledge management systems)	0.56		
Org2 (New organisation of work)	0.79		
Org3 (New relations with others)	0.45		
Mar1 (New design)		0.33	
RMAR (Market introduction of innovations)		0.45	
Eigenvalue (original)	2.14	1.43	1.08
Percentage of variance explained	24	16	12

Pattern matrix. Values less than 0.30 are suppressed. Extraction method is principal axis factoring. Rotation method is OBLIQUE;
$r_{f1\text{-}f2}$ = 0 .18, $r_{f1\text{-}f3}$ = 0.46 and $r_{f2\text{-}f3}$ = 0.07.

The table shows that the Eigenvalue of the ‘classical innovation factor’ is 1.43 and the Eigenvalues of all factors are > 1. Factors 1 and 3 are just a check in this factor solution: they makes sure that ‘classical innovation indicators’ do not measure ‘process and organisational innovation’. This method generally prevents indicators which are empirically part of different dimensions from being combined.
Table 5 thus leads to the combination of indicators ‘Prod1’, ‘Mar1’ and ‘RMAR’, which all refer to ‘classical innovation’. This proposal satisfies all the demands stated above except one: the reliability of this ‘classical innovation scale’ is insufficient. This means that the indicators can only be used in the ‘traditional way’ (combining indicators using or/or-statements). The difference between these two methods is explained further in the next chapter, which examines validity.
It appeared impossible to satisfy all the demands. The FA of table 5 demonstrates that ‘classical innovation’ is one, separate, dimension, but ‘Prod1’, ‘Mar1’ and ‘RMAR’ are not a consistent measurement of that dimension, while the indicators of the other two factors do measure corresponding dimensions consistently. The reason for this lies in indicator ‘Mar1’, which was not ‘designed’ to measure ‘classical innovation’. The next factor solution, table 6, proves this. This solution improves the ‘classical innovation factor’: it distinguishes firms that have performed ‘intramural R&D’ and have innovated goods – resulting in penetrating new markets first and patenting the new good(s) – from firms that have not. Again, the other factors are not disturbed by this factor.
The indicators which are part of this factor constitute a reliable scale. So, this combination makes the measurement of classical innovation empirically consistent. Furthermore, table 6 gives information on the role of indicator ‘Mar1’.
Indicator ‘Mar1’ does not fit in this factor solution very well. It is, however, still related to goods, and previous chapters have demonstrated that it measures

technological content, so why doesn't it fit? The reason is probably that indicator 'Mar1' (New design or packaging) refers to existing goods, while 'classical innovation' refers to new goods.

Table 6. Rotated three-factor solution for 'classical innovation' in the Netherlands (CIS4)

Item	F1	F2	F3
Prod1 (New goods)	0.69		
Proc1 (New methods of manufacturing)			0.32
Proc2 (New methods of logistics, delivery or distribution)			0.57
Proc3 (New supporting activities)			0.38
Org1 (New knowledge management systems)		0.56	
Org2 (New organisation of work)		0.79	
Org3 (New relations with others)		0.46	
Mar1 (New design)	0.23		
RRDIN (Intramural – in house – R&D)	0.44		
NEWMKT (New to your market?)	0.65		
RMAR (Market introduction of innovations)	0.48		
PROPAT (Apply for a patent)	0.36		
Eigenvalue (original)	2.37	1.96	1.10
Percentage of variance explained	20	16	19

Pattern matrix. Values less than 0.30 are suppressed, except for indicator 'Mar1'. Extraction method is principal axis factoring. Rotation method is OBLIQUE; $r_{f1\text{-}f2} = 0.18$, $r_{f1\text{-}f3} = 0.09$ and $r_{f2\text{-}f3} = 0.46$.

The reliability of 'classical innovation' is acceptable: α = 0.65. It sums the five items that are part of the first factor in table 6 – except item 'Mar1' – then divides it by five (scaling, see next chapter). The Alpha shows that the correlation between 'the true classical innovation score' of firms and the 'measured score' is r = 0.81, which is good for survey-based information.
The conclusion is that the proposals resulting from both tables 5 and 6 are convincing solutions. In table 5, indicator 'Mar1' still has a function. In table 6, it has only the function of directing the search to this new combination. The solution of table 6 is preferred, as it satisfies all demands; it also explains more variance.

10. New combinations of existing indicators and validity

Chapters 1 to 6 served one goal: to improve our basic understanding of innovation dimensions. As a result we have found the direction in which proposing new combinations of existing indicators would be empirically meaningful. However, we still have to prove that these new proposals will result in added value.
A new combination of existing indicators has added value if it increases our understanding of innovation. Furthermore, the stronger the empirical requirements, the better founded the proposals are. This study demands the following empirical requirements:
1. The new combination must be explained by one factor. This factor may not explain other innovation types. This requirement means that a new combination is allowed if it is a one-dimensional measurement which does not interfere with existing innovation definitions.
2. The new combination consists of indicators which make up a reliable scale. This requirement means that a new combination is allowed if it is a consistent measurement of one dimension.

3. The new combination is valid.

Validation is a theoretical issue. Absolute certainty about validity of constructs cannot be obtained if indirect observations are used. So validity is plausible, or it is not. Three frequently used definitions of validation are:

1. Content validity: this concerns the content of indicators;
2. Criterion validity: this concerns a comparison of the measurement of a construct with other measurements of the same construct;
3. Construct validity: this concerns the issue of whether the construct fits well in an existing theoretical framework. [xxviii]

Content validity will increase if the content is plausible. Rational analysis of the indicators is the most used method, as content validity is not based on data. Criterion validity will increase if the measurement of the construct correlates very well with another, verified, measurement of the same construct, not generated by another method. Construct validity will increase if tests of hypotheses derived from an existing theoretical framework lead to satisfying results.

Criterion validity will not be included in the proposals put forward in this chapter as no other leading verified measurement of innovation exists. The proposals are based on Factor Analytic and Rational Scale Construction.

So what findings are relevant to the proposals so far? HOMALS led to the discovery of two meaningful dimensions: 'being innovative or not' and the 'technology dimension'. FA found us three factors, of which the third represents 'classical innovation'. The 'OBLIMIN rotation' showed that the first two factors of the FA are a rewritten version of the second HOMALS dimension, so they cannot be treated as different structural aspects of the data at the same time. This leads to four proposals:

1) Combination of the indicators which represent 'classical innovation';
2) Combination of the indicators which represent 'being innovative';
3) Combination of the indicators of process innovation;
4) Combination of the indicators of organisational innovation.

A variant on the second proposal is: addition of the indicators which represent 'being an innovator or innovative active'. This fifth proposal, and the third and fourth, already exists. However, in this report all proposals are scaled combinations. The proposals are not based on the 'traditional' way of computing composite indicators. Making combinations of existing indicators by using 'or/or-statements' refers to this traditional method, which makes poor use of existing information.

For example, innovators are firms which report at least one of the six indicators of 'product or process innovation' or 'ongoing or abandoned innovation activities'. This method is used to calculate the total number of innovators in a country. It leads to two categories: '0: no innovation type reported' and '1: innovation reported'. Instead, scaling innovators leads to seven categories: 0: no innovation type reported', '0.17: one innovation type reported', '0.33: two innovation types reported', and so on up to value '1: all six innovation types reported'. Scaling means that the six items, defining innovation types, are added and then the sum is divided by six. The advantage is that the information on the reported number of innovation types is not ignored; the traditional method put all kind of innovators in one box.

The result of scaling (using ordinal scales) is that innovators are divided into subgroups, which reflect the degree of innovativeness. This is based on the idea that firms which report four types of innovation actually innovate to a larger

extent than firms which report only one. In short, the 'innovator scale' is based on exactly the same information, but it is far more informative.
Ordinal scaling means that only ranking information can been used. Neither the CIS nor the Oslo Manual states which type of innovation type is more important, if any all, so to make it an interval, scales would pose methodological problems. For example, what weights should be used and why? However, this is not a problem if ordinal scales are used (as it would be if interval or ratio scales are used). Ordinal means only, for example, that three is more than one, and ranking is always allowed with scales that are above the nominal level of measurement.
One difference between the 'innovator scale' and the 'innovation scale' – both are proposed below – is that the latter is based on all types of innovation. So whether the 'innovator scale' or the 'innovation scale' is used depends on the purpose of analysis. If the definition of 'innovator' is changed in CIS5, the latter may become more important.

1. Proposal: 'Innovation activity Scale'

RA, HOMALS and FA demonstrated that the ten indicators of innovation types meet the first two requirements. However, the 'innovation activity' is based on six items, so is it reliable as well? The answer is: yes, RA gives $\alpha = 0.69$, which is sufficient. Furthermore, the reliability is larger than the reliability of the separate parts of the scale. All six items should be part of the proposal, as if any were to be deleted, α would not increase.
So can one factor explain these six indicators? Table 7 shows that an 'innovator factor' exists (see below). Of course, HOMALS already demonstrated that all ten indicators are part of the first factor, so it would be strange if six of the same indicators are not, since they all discriminate between 'innovative' or 'not'. Nevertheless, the existence of a second factor had to be excluded.
Again, the unrotated solution is used, as the factor should discriminate only between firms that are innovative and firms that are not. The Eigenvalue of the second factor is too small: 0.89, and the loadings on the second factor do not disturb the expected interpretation. The conclusion is that the first two requirements of this proposal are met.

Table 7. 'Innovator factor' in the Netherlands (CIS4)

Item	F1
Prod1 (New goods)	0.58
Prod2 (New services)	0.47
Proc1 (New methods of manufacturing)	0.63
Proc2 (New methods of logistics, delivery or distribution)	0.59
Proc3 (New supporting activities)	0.61
INONAB (Ongoing or abandoned innovation activities)	0.71
Eigenvalue (original)	2.73
Percentage of variance explained	46

Extraction method is principal axis factoring. Unrotated solution displayed.

Scaling these indicators thus results in the 'innovation activity scale'. Of course, the concept innovation activity already exists and is used when composing the CIS data, but it can be used in a better way if scaled.
The 'innovation activity scale' must be tested to justify this proposal scientifically as well. Is the 'innovator scale' valid, will it provide an added value? It would if it fits in existing innovation theory. Also, if it fits in existing innovating theory better than existing findings, it would be an improvement.

This can be tested by correlating the 'innovation activity scale score' (ISS) with 'R&D expenditure'. A basic assumption of innovation theory is that R&D and innovation are related: firms spend money on R&D, because it will lead to innovation – to some extent. Correlating the two is therefore one way to obtain information on the construct validity of the scale. If there is no positive correlation, the new scale is not valid, or innovation theory would be false. However, the latter possibility is less likely as other – independent - affirmative observations of innovation theory exist.
In addition, if this positive correlation is stronger than the same correlation between 'innovation activity' and 'R&D expenditure'– calculated in the traditional way – scaling innovators would even improve our understanding of innovation.
This immediately proves the usefulness of the 'innovator scale', as it is impossible to calculate the correlation between 'innovation activity' – using the traditional method – and 'R&D expenditure'. One reason for this is that only 'innovation active' firms report 'R&D expenditure'. In CIS, 'innovation inactive' firms (value '0') cannot answer the questions on R&D – understandably – so they all have missing values on R&D. A second reason is that all 'innovation active firms' are part of one group (value '1'), because of the traditional method.
However, it is very important, in policy and science, to improve our understanding of the relationship between R&D and innovation activity. Scaling this makes it possible to calculate a correlation, because it takes differences between 'innovative active' firms into account. This correlation is quite strong: $r_{\text{innovation activity (scale) - R\&D-expenditure}} = 0.34$. [xxix] This means that the more money 'innovation active' firms spend on R&D, the more innovation activity they show, and vice versa. Of course, the correlation does not prove causality.
In the traditional method this kind of interpretation would be tautological (if it could be computed), since a firm spending on R&D could hardly be 'innovation inactive'. However, the advantage of the scaling method is that it reveals information on the degree of activity, which correlates. In this case it says: the more firms spend on R&D, the more they report product and/or process innovation and ongoing or abandoned activities. This degree cannot be expressed if the traditional method is used, since that method cannot compute the correlation ($r_{\text{innovation activity - R\&D-expenditure}}$). This means that other kinds of validation of the 'innovation activity scale' should be conducted. But what other kinds?
One possibility could be eliminating the indicator 'ongoing or abandoned innovation activities'. In that case, innovators consist of firms which report any of the five 'product or process innovations'. This allows a correlation with 'R&D expenditure' to be calculated in the traditional way (as now another category with data emerges). It leads to the correlation ($r_{\text{product or process innovator - R\&D-expenditure}}$) = 0.03. However, the same, scale-based, correlation is much stronger: $r_{\text{product or process innovator (scale) - R\&D-expenditure}} = 0.22$.
This finding proves the construct validity of the innovator scale. [xxx] Or more correctly: it actually proves that the traditional way of calculating innovation activity (or in general: innovators) *reduces* construct validity. The reason for this is that using or/or-statements is not based on information on the number of reported innovation types.
The found construct validity of the 'innovation activity scale' could not be falsified using other techniques to analyse the correlation with 'R&D', nor by making other comparisons, nor by making other combinations. For example, the correlation ($r_{\text{prod1 - R\&D-expenditure}}$) = 0.10. [xxxi] Another possible comparison would be calculating the correlation between 'process innovation' and 'R&D

expenditure', in which 'process innovation' is not scaled. This correlation ($r_{\text{process innovator - R\&D-expenditure}}$) = 0.20, is still weaker than a rank-based correlation.
So the conclusion is that scaling innovation indicators leads to an improved understanding of innovation.

Additional remarks:
1. The 'innovation activity scale' does not assume that some innovation types are more important than others. Of course, this would be possible by weighing some items in the scale. [xxxii]
2. The 'innovation activity scale' does not reveal information on which types of innovation have been realised. However, this output can be generated at any time.

2. Proposal: 'Innovation Scale'

RA, HOMALS and the unrotated FA already demonstrated that the 'innovation scale', which is based on all ten items of innovation types, meets the requirements for proposals. This scale is similar to the 'innovation activity scale'. It differs in that it does not (only) refer to the definition of innovation activity, as it consists of 'organisational and marketing innovation items' as well. Another difference is that the indicator 'ongoing or abandoned innovation activities' is not part of the scale. Of course, this could be the case, if desired. If so, the reliability of such a scale would increase slightly, to $\alpha = 0.79$ in the Netherlands.
So the 'innovation scale' is an ordinal measurement of the reported number of innovation types. Again, it does not matter which types are reported: it is just a simple measure of the degree of renewal. Its correlation with 'R&D expenditure' is $r_{\text{innovation (scale) - R\&D-expenditure}} = 0.22$, which is weaker than $r_{\text{innovation activity (scale) - R\&D-expenditure}}$. This is quite a logical observation, as it is not very likely that 'R&D expenditure' would relate very strongly with, for example, 'changing the organisation of work', in terms of, say, 'changes in the management structure'. Note that the correlation of the 'innovation scale' is still much stronger than the correlation calculated in the traditional way, which is not logical in this case ($r = 0.02$). [xxxiii]
The correlation of the 'innovation scale' and 'R&D' therefore demonstrates that the more innovation types they report, the more firms spend on R&D.
In the Netherlands, indicators 'Proc1' (New methods of manufacturing) and 'Mar1' (New design or packaging) contribute the most of all indicators to the total correlation with 'R&D expenditure'. So all disadvantages of indicator 'Mar1' – as demonstrated in this report – turn out to be an advantage at the same time. This feature means that this indicator is still very useful.

3. Proposal: 'Process innovation scale'

The RA, HOMALS and FA showed that scaling 'process innovation' is very well possible. The correlation ($r_{\text{process innovator - R\&D-expenditure}}$) = 0.20, while scaling results in a slightly better correlation ($r_{\text{process innovation scale - R\&D-expenditure}}$) = 0.22; the gain is mainly the possibility to use information on the number of reported types of 'process innovation'.

4. Proposal: 'Organisational innovation scale'

The RA, HOMALS and FA showed that scaling 'organisational innovation' is very well possible. The scale makes analyses possible in which information on the number of reported types of 'organisational innovation' can be used. The correlation ($r_{\text{organisational innovator - R\&D-expenditure}}$) = 0.12, while scaling results in a correlation ($r_{\text{organisational innovation scale - R\&D-expenditure}}$) = 0.14; this is a small improvement as well.

5. Proposal: 'Classical innovation scale'

The fifth and last proposal is based on the remaining third factor (see chapter 4, table 4) and leads to a new construct, which is labelled 'classical innovation' (see chapter 6, table 6). Chapter 6 led to the conclusion that two satisfying constructs of 'classical innovation' exist. However, the second solution satisfies all demands, so this construct is preferred.

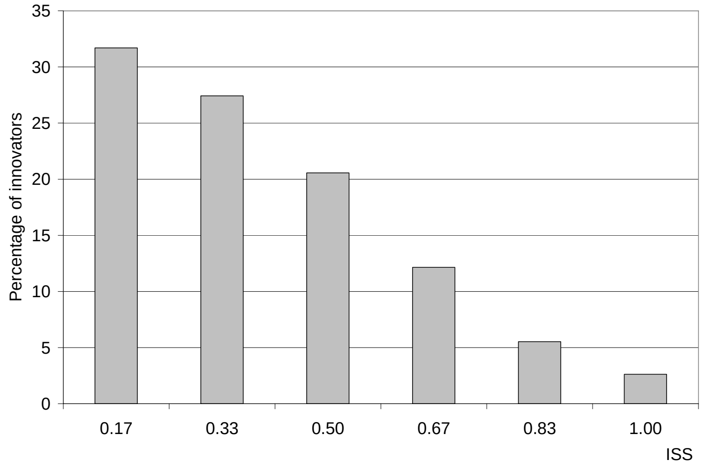
'Classical innovation' scales the indicators 'New goods' (Prod1), 'New to the market' (NEWMKT), 'Market introduction of innovations' (RMAR), 'Intramural R&D' (RRDIN) and 'Patent apply' (PROPAT) ordinally. The scale score distinguishes firms by their 'degree of activity in new goods'.
'Classical innovation', calculated in the 'traditional way' correlates with 'R&D expenditure': r = 0.17. Scaling 'classical innovation' nearly doubles this correlation to r = 0.32. This is an impressive improvement, because the correlation is almost equal to the correlation between innovation activity and R&D-expenditure: $r_{\text{innovation activity (scale) - R\&D-expenditure}} = 0.34$.
Of course, scaling is only an option. 'Classical innovation' could be also used by just combining the indicators using or/or-statements. For example, to establish the total number of 'classical innovators' in a country, the traditional method could be used.

11. Applications of scaling innovation

How can scaling of innovation indicators be used? One possibility is by calculating the mean (or perhaps better: the median) 'R&D expenditure' for each 'innovation activity scale score' (ISS). For example, firms that have realised two types of technological innovation may have spent less on R&D than those that have realised four types (true in the Netherlands). Another application is relating ISS to the share of firms which have realised 'product innovations'. A third application could relate ISS to the effects of innovation. The results of the latter two suggestions are presented in this chapter. First, the distribution of ISS is given.

ISS consist of the five types of innovation ('product and process innovation') and the item 'ongoing or abandoned innovation activities'. All are part of the definition of innovation activity. Scaling these items results in seven categories. The ISS for 'innovative active' firms are displayed in figure 5. This means that ISS=0 is not part of the scale (of course, it could also be displayed). ISS=0.17 represents firms that report one type of innovation (which defines 'innovation activity'), score 1 refers to firms that report all six types. So, figure 5 displays the degree of innovation activity of firms in the Netherlands, instead of the traditional yes-no distribution.

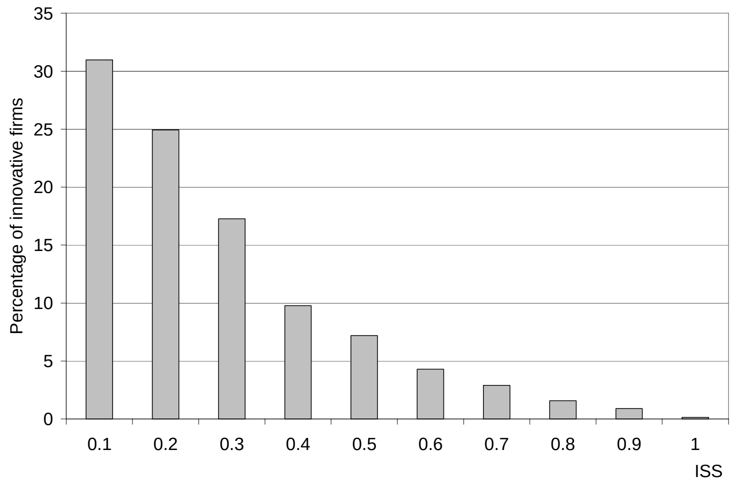
Figure 5. Distribution of the Innovation activity Scale Scores (ISS), the Netherlands (CIS4)



Observations:

1. Of all innovators (to be more precise: of all ‘innovation active’ firms), nearly 3 percent reported all type of innovations (i.e. relevant for this scale). This group of firms is extremely innovative. However, most innovators – nearly one third of the total – reported only one type of innovation;
2. Around 20 percent of all innovators reported more than three innovations.

Figure 6. Distribution of Innovation Scale Score (ISS), the Netherlands (CIS4)



The distribution of the ‘innovation scale’ (instead of ‘innovation activity scale’) is presented in figure 6. Innovative firms on the *y* axis now consist of firms having reported at least one of the ten innovation types defining the indicators. ‘Innovative’ means ISS > 0, which is the reason why, again, ISS=0 is not part of the figure. So the *y* axis is not intended to distinguish ‘innovators’ from ‘non-innovators’, although it is intended to distinguish between innovators.

Observations:
1. ISS=1 represents the group of firms which report all four types of innovation (product, process, organisational and marketing innovation, according to the corresponding ten indicators). This share is, not surprisingly, very small (but more than 0 percent);
2. Nearly one third of innovative firms, as defined in figure 6, reported one indicator of innovation. One quarter of innovative firms reported two indicators;
3. Around 10 percent of innovative firms reported more than five innovations.

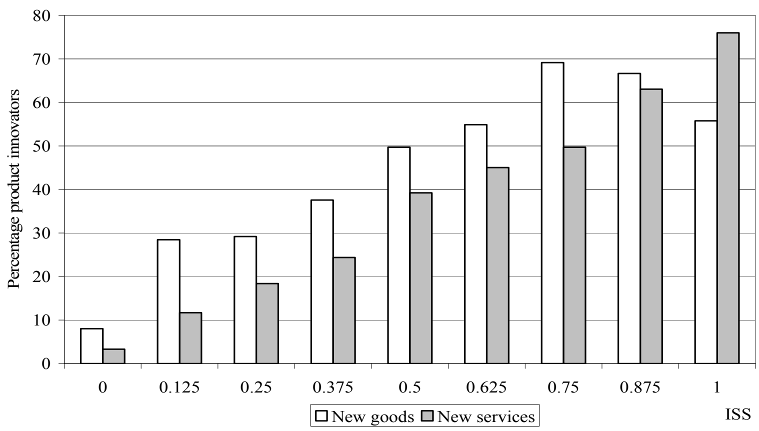
Using ISS

Figures 5 and 6 show the distribution of both scale scores. The first way to use ISS was already demonstrated in the previous chapter: they make it possible to correlate with R&D expenditure. But ISS can also be used as follows.

If the 'product innovation indicators' ('Prod1' and 'Prod2') are deleted from the 'innovation scale score' – in which case the ISS would consist of eight items – would the ISS still correlate with 'product innovation'? In other words: "Does a relationship exist between 'innovation in goods and services' on the one hand and the other innovation types on the other? Table 2, the correlation matrix, already demonstrated that this relationship does exist, but scaling will lead to new insights. Figure 7 displays this relationship.

Once again the ISS is on the *x* axis. However, scale score 1 now refers to firms that have realised all eight indicators of innovation types: it is the highest value. Score 0.5 represents firms that report four out of eight indicators of innovation types. Notice that value 0 is now also included in the figure. Both indicators of 'product innovation' are displayed on the *y* axis. Of all CIS4 firms in the Netherlands 19 percent reported 'New goods' (Prod1); this fact is not revealed by the figure.

Figure 7. Innovation scale score (ISS) and share of product innovators, the Netherlands (CIS4)



A clear pattern is visible: the higher the ISS, the larger the share of 'product innovators'. Firms with ISS=0.75 clearly more frequently innovate in terms of their goods: around 69 percent of this group. The share of 'goods innovators' in this group is 3.5 times the average. This finding suggests that types of innovation other than 'product innovations' are mainly supporting and facilitating 'product innovations'.

Another way to use ISS could be to relate them to effects of innovation. In CIS4, effects of innovation are measured with the aid of nine indicators. These effects can be clustered into ‘product-oriented’, ‘process-oriented’ and ‘other’ effects. Furthermore, firms are asked to report the ‘degree of the observed effects’ as well.
On behalf of this analysis, ISS – based on the ten items – are split into two groups: ISS < 0.5 and ISS ≥ 0.5; once again, ISS = 0 is not included in the analysis. [xxxiv] This results in two observations.
First, firms with ISS < 0.5 report all effects of innovation less frequently than firms with ISS ≥ 0.5. Second, the latter group reports all observed effects to a significantly higher degree as well.
Many other applications may exist. This chapter only presents some examples which cannot be calculated using the traditional method of combining indicators. Information on the ‘degree of being innovative’ may help to understand other topics as well, for example ‘productivity’.

12. Discussion

A first proposal of this report is to scale innovation indicators. This is perfectly possible with the proposed ‘innovation activity scale’, ‘innovation scale’, ‘process innovation scale’, ‘organisational innovation scale’ and ‘classical innovation scale’. A ‘marketing innovation scale’ and a ‘product innovation scale’ are not allowed in the Netherlands, according to test theoretical standards.
Such standards are important statistical tools for assessing survey information, like CIS. For example, they can be used to calculate reliability and validity of the main innovation definitions and corresponding items in the questionnaire. It is advisable to use such measures before using any survey data at all, as innovation researchers, for example, need to know whether the data are reliable and valid. A test theoretical approach is particularly advisable if indicators are to be combined.
Of course, differences exist between tests and survey information. For example, a test measures a test score which refers to a real score of a person or firm, while survey-based information actually states directly: this is the real score. Both contain error and both use different ways to generalise. However, both methods have to deal with reliability and validity issues. Testing techniques offer such tools to a much larger extent than standard survey procedures.

The CIS4 data of the Netherlands reveal that combining ‘New design or packaging’ (indicator ‘Mar1’) and ‘New goods’ (indicator ‘Prod1’) makes more sense empirically than combining ‘Mar1’ and ‘Mar2’. Together, ‘Prod1’ and ‘Mar1’ are part of one factor, which directed the present study to ‘classical innovation’. Efforts to scale other CIS indicators – in order to try to find a potential ‘marketing innovation scale’ – are likely to fail, as ‘New design or packaging’ is strongly related to ‘New goods’. This makes it very difficult to find one unique factor justifying such a new combination.
However, this study has used strict demands for the proposal of new combinations of existing indicators. For example, the demand that any proposal must consist of one factor and the demand that such a factor may not disturb the other factors of innovation types. There was a reason for this strict demand: combining CIS indictors is actually making new definitions, in an empirical way. Therefore, it has to be ascertained that the proposed new combinations do

not interfere with existing innovation definitions, at least with those consisting of one dimension, and themselves consist of one dimension.

This approach revealed that the measurement of marketing innovation could be improved. Neither indicator of 'marketing innovation' measures one dimension and neither are consistent measurements, while the other constructs of innovation types do satisfy the demands, or expectations – in the Netherlands.

What kind of improvement is needed is another question, and is not covered by this report. Institutions like Merit, the OECD, Eurostat and NESTI are working on and improving science and technology indicators, like CIS.

Another demand was that the proposals must use the available information as well as possible, which was the reason for scaling the new combinations. Subsequently, another demand emerged: the proposal must constitute a reliable scale.

The consequence of all these demands is that only a few indicators can be combined. The validity of the scales needed to be proven as well, which was the final demand. This enabled, for example, the calculation of the correlation between innovation activity and 'R&D expenditure' ($r = 0.34$). If not scaled, such a correlation is impossible to calculate since inactive firms do not spend on R&D. The resulting correlation contains information on the degree to which the two are related, which is impossible using the traditional method.

The question of whether the typology 'technological' and 'non-technological' innovation may be oversimplified can be answered with 'No' for the Netherlands. At least, all indicators of the basic definitions of *innovation types* fit very well in this typology. Indicator 'Mar1' (New design or packaging) is the only exception found. However, this indicator is still useful in the Netherlands, as it correlates well with 'R&D expenditure'.

It is easy to understand that 'New design or packaging' (Mar1) is strongly related to 'product innovation'. Firms which modify their goods sometimes actually innovate, while sometimes they just change the design or packaging. Another way of reasoning is that implementing 'changed design or packaging' may force firms to innovate their technology. If this is true, it would lead to innovation of, for example, the – possibly already existing – assembly line. It is not relevant for the present discussion how often this may be the case. Both arguments suggest technological aspects and are therefore a rational counterfeit of 'Mar1' as an indicator reflecting 'non-technological content'. So, both technological and non-technological *knowledge* could be needed when innovation is applied in any form.

Classical innovation can be used both as a scale and in the traditional way. Furthermore, all proposals are applicable immediately and in any country, as it is irrelevant what factor structure exists in a given country; only one demand must be met to use the proposed scales in other countries: the scale must be reliable. Since every proposed scale is based on a combination of indicators which are part of one dimension, verifying these dimensions again is unnecessary. For example, if different innovation dimensions exist in a country, this will reduce the reliability of the scales.

Just calculating the RA of the scale, therefore, is a quick way to test whether innovation dimensions are equal to those of the Netherlands. Of course, if any RA result appears insufficient new innovation dimensions will have to be found. This study has tried to demonstrate a way to do this.

However, when proposing combinations of existing indicators which concern specific types of innovation and which are not scaled, this may not be the case. Such combinations could suffer from international differences, especially if there is no check on whether the proposal disturbs other innovation definitions.

So the reliability of the proposed scales may differ between countries. In the Netherlands, all ten innovation types, for example, are part of the ‘innovation scale’. It is possible that in another country the removal of any item will make the scale more reliable.
Scaling innovation indicators may be useful for researchers and policymakers. For example, suppose the share of innovators in the Netherlands is equal to the share of innovators in Germany; at first sight one would say both countries are equally innovative.

Endnotes

[i] The Oslo Manual is an internationally accepted handbook of Innovation.

[ii] The weighting factor used, differs from the ‘normal’ weighting factor. First, the ‘CIS population for EU’ differs from the ‘Dutch CIS population’; the latter consists of more NACE categories (called ‘SBI’ in the Netherlands), so the total number of firms is almost twice as high (as is the sample). Converting to the ‘CIS population for EU’, makes an EU comparison possible, so only the ‘EU required NACE’ are used. However, if only national interests are relevant, all NACE may be used. The Dutch CIS4 dataset for ‘EU purpose’ contains 5,927 cases and after using the ‘normal’ weighting factor this converts to 27,412 firms. However, converting to this total would inflate error of measurement, which is undesirable. The goal is not to raise to a certain population. On the other hand, just basing the analysis on the sample results without using the normal weighting factor would mean no correction for the bias of the sample, as the ‘normal’ weighting factor corrects this as well. So, the weighting factor used in the analyses of this report is first weighing the data to the total ‘CIS EU population’ – to correct for representativeness – and second, ‘weighing back’ the data to the original sample size. This results in a sample in which the cases have the right weight with respect to each other, without blowing up errors in measurement. Multiplying the ‘normal’ CIS EU weighting factor by the fraction n/N leads to this new weighting factor. In this case, n/N = 5 927/27 412 (sample size/sum of weights). The sum of weights of the new weighting factor is now 5,927 again, which is the sample size. Another advantage is that it makes a better approximation of exact testing. After ‘normal’ weighing you would already know the parameters in the total population, so testing would not make sense. The correlations displayed in the table can now be tested.

[iii] Since the indicators consist of two nominal categories (0 and 1), Spearman and Pearson r are the same.

[iv] Calculating (semi) partial correlations accounts, just as FA, for multicollinearity.

[v] Scaling assumes that firms can be given a true innovation score, which are, in these cases, ordinal scores.

[vi] Only if random error is perfectly spread around the ‘true score’, would increasing N increase the validity; in all other cases, increasing N would just decrease random error, not bias.

[vii] The mean mutual correlation between all items of a scale is the best predictor of α (using the Spearman-Brown formula). The scale and item means are not given in the table, because they have little relevance (any transformation of 0 and 1 is permitted).

[viii] In surveys, these conditions of α are applied; stricter conditions are used if direct measurement is possible; using a survey is no direct measurement, so more random error has to be tolerated.

[ix] Using scaling techniques does not mean that items should be treated as part of a scale, if the scale is sufficient. This result may be explained by the concept ‘degrees of freedom’ of responses.

[x] Item ‘Mar1’ provides added value in this scale as well, since it measures renewal: it is part of the first dimension, see the next chapter.

[xi] Compare it with the items of the ‘process innovation scale’, which is reliable and which is explained by one unique factor (see next chapter). Suppose these items did not measure this construct consistently either, and no one factor could be found. Would the solution be to say that the three items of ‘process innovation’ measure three different dimensions of the construct ‘process innovation’ – and that actually this fact causes the poor reliability and is the reason for not finding one, unique factor? No, because then it would be unclear why all indicators are labelled ‘process innovation’: the construct could, in that case, consist of three random items, which do not measure one dimension either. It would further lead to the existence of n dimensions of any construct, where n = the number of indicators of a construct. Instead, it is only logical to cluster indicators which refer to one single dimension. This reasoning makes the construct ‘marketing innovation’ existing of two dimensions invalid. Furthermore, what would the second dimension be and mean, if it existed? Also, the label ‘marketing innovation’ cannot account for the variance of the two items ‘Mar1’and ‘Mar2’ in the Netherlands. Moreover the number of dimensions of a construct, when it is defined, should be known before measurement. Also, if two dimensions did exist, two items for each dimension are needed – which is not the case – and if not, calculating reliability would be impossible. So, the indicators of ‘marketing innovation’ should be treated as separate measurements in the Netherlands. The next chapter will demonstrate that the construct is invalid in the Netherlands, since no factor exists which explains ‘Mar1’ and ‘Mar2’. This means that the two indicators are not part of one dimension, which makes it logical that such a supposed scale is not allowed.

[xii] Now the categories of indicator are ‘0: No’ and ‘1: Yes’. However, these scores are actually random. They could be replaced by any other two values. HOMALS transforms all categories of all variables in a way that correlations between the variables will become optimal, so variance is detected best. ‘Optimal’ means that the total fit of the solution is best. The choice for this technique is based on the fact that all variables used in this analysis are of nominal level. A (supposed) binary FA or PCA could be used as well. Since all variables used in this analysis consist of two categories, they can be considered as being interval variables (since a straight line exists between any two points in space), which is necessary to use FA or PCA; however HOMALS treats the variables as if they are of nominal level of measurement, which they obviously are, so no assumptions are needed, since a difference exists between ‘real’ interval variables – such as the variable body weight – and variables consisting of two categories. HOMALS transforms the indicators, using different multiple quantifications for each dimension. A requirement could be to use single quantifications for all dimensions, which is PRINCALS. This technique is the best option if the indicators consist of a mix of nominal and ordinal data. It does not plot categories, however, only variables. HOMALS plots categories. Another difference between FA and HOMALS is that the latter reveals information about categories as well, which is very useful, as the next chapter will show. FA results in factor loadings of variables, not of categories of variables.

[xiii] The object scores of figure 1 did not contain outliers (so these plots are not given) and there are no missing data.

[xiv] Eigenvalues in HOMALS are not added, as FA uses means. The reason for this is that multiple quantifications are used. Eigenvalue has a different meaning in HOMALS. It describes, or accounts for, the total variance, instead of explaining it. However, adding them or not is just a theoretical question of what is permitted, since

the multiple quantifications do not increase the separate Eigenvalues, compare them to the FA Eigenvalues, given in table 4. However, it would be strange if the summed Eigenvalues of HOMALS were larger than the summed Eigenvalues of the FA, since HOMALS is used for data of lower level measurement. The total fit is therefore always better using FA, while in most cases it leads to the same solution and interpretation. However, in this case the purpose is only to demonstrate the technology dimension; plotting a third dimension is not necessary. FA will deal with that later on.

[xv] The Oslo Manual uses only the former 'technological innovation indicators' and not the former 'non-technological innovation indicators' in the definition of innovators. In practice this means that this distinction is still applicable. However, after 2009, at the time of CIS5 (reference period 2006-2008), both types of innovation may and probably will be used to define innovators.

[xvi] Or PCA, depending of the exact goal of the analysis; examining whether items add up to one scale is best conducted with a PCA. However, an FA assumes causality, at least on theoretical grounds, so this is a stronger requirement and therefore an FA is preferred. Using an FA assumes interval level, which is possible. Conducting HOMALS again, generating three dimensions, leads to the same result.

[xvii] The FA of the innovation types will lead to three factors and in SPSS, 3-dimensional plots are difficult to visualise.

[xviii] Although not a very strong validation; the constructs are better validated if content, criterion and construct validity is found as well (see chapter 7).

[xix] More matrix solutions exist.

[xx] When using OBLIMIN, a pattern and structure matrix is given; the procedure is to choose the one that can be interpreted best. Note that the Eigenvalue of the third factor is < 1. This makes a two-dimensional interpretation (see HOMALS), in which the dimension 'technological – non-technological' is most important, defendable.

[xxi] Or any other 'non-orthogonal rotation method', but OBLIMIN has the best credits.

[xxii] Another consequence of rotating with OBLIMIN is that the better fit has a price: in this rotation method, all loadings are always slightly, compared to VARIMAX. However, this is not a great disadvantage, as the total variance explained by the solution is the same, and the interpretation of indicators which are relevant is most important. The fact that a loading has a value of 0.89 or 0.72, for example, is less relevant. In this case, there is no choice, since VARIMAX cannot be used.

[xxiii] Notice that to create the same figure were to be created using a PCA, the arrows would have to be plotted in reversed direction, since this technique does not assume causality.

[xxiv] The position of the positive category-point of indicator 'Prod2' has changed, becoming technologically neutral (see figure 3). However, this is merely a result of the plotting of category 'Services Sector', as the other figures will show later on in this chapter. In figure 1, all NACE (EU) data have been used as well. However, in figure 1 NACE was not directly part of the analysis, so the NACE categories were not plotted. HOMALS is now forced to recalculate the optimal category quantifications, and to plot NACE categories. Of course, the shift of the positive category of item 'Prod2' is real, meaning that the 'Services Sector' and indicator 'Prod2' are related.

[xxv] Notice that a vector < 0 does not mean there is no innovation.

[xxvi] HOMALS can show this very well; most studies calculate percentages of innovation types for both sectors, thus not revealing the whole pattern.

[xxvii] SPSS tabulates the vectors, so it is possible to calculate how much ‘closer’ a category-point is to another point.

[xxviii] Multi Trait Multi Method Matrix is also a way to check test properties of a survey, and many more definitions, or types, of validation exist.

[xxix] ‘R&D expenditure’ consists of ‘Intramural R&D’, ‘Extramural R&D’, ‘Acquisition of machinery, equipment and software’, ‘Acquisition of other external knowledge’. ‘R&D expenditure’ is extremely skewed. For this reason Spearman r was used.

[xxx] Note that this correlation is not based on the fact that scale score ‘0’ has no ‘R&D expenditure’. Actually the correlation is based on the scale values > 0, as scale value ‘0’ has missing values on ‘R&D expenditure’, just as value ‘0: no innovator’ using the traditional method. Notice that replacing all missings by value ‘R&D = 0’ would be incorrect. This indeed enables the calculation of a correlation between innovators and R&D in the traditional way. However, the resulting correlation will be extremely large (r_{false} is about 0.85), as false covariance is created by this action: missings cannot be treated as zeros, even not under assumption. For example, this action could be defended as follows: in CIS ‘non-innovators’ cannot report ‘R&D’, as CIS does not measure this. So if such a firm did – in reality – spend on ‘R&D’, it must be treated as an R&D institution, and therefore be excluded from the CIS population. This would allow the assumption that all CIS non-innovating firms have an R&D of 0. However, this is just a smart trick to let all definitions fit theoretically; it is not based on measurement, instead it is based on arbitrary arguments. For example, an equally strong arbitrary argument is this: such supposed R&D institutions may also change their processes, which would be an argument to treat them as innovators, and thus exclude them from the population of R&D institutions. In that case, the same R&D firm becomes part of CIS again. This reasoning ‘proves’ that the missings cannot be replaced by zero. The assumption ‘non-innovating firms spending on R&D are R&D institutions’ is merely a practical choice. Incidentally, even if all R&D missings are replaced by ‘0’, a scale-based correlation is still stronger if the same action is performed with the scaled value ‘0’ which suffers from the same missings as well.

[xxxi] Note that this is a correlation based on a single indicator, no combining or scaling is used. However, this correlation could falsify the result as well, which is not the case, and had therefore to be calculated.

[xxxii] For example, if - in a given study - ‘product innovation’ is important and this has to be reflected in the innovation scale, this would lead to weighing ‘Prod1’ and ‘Prod2’. Of course, which weighting factor should be used best is another issue, which is not relevant here.

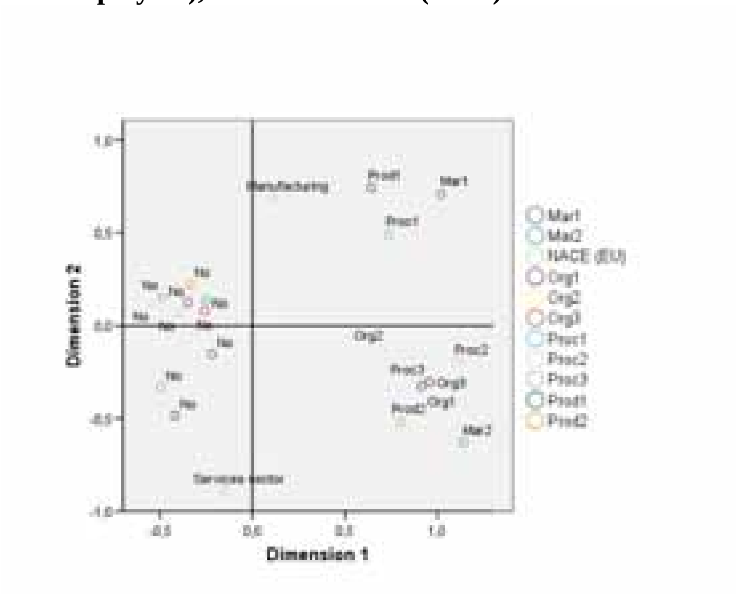
[xxxiii] In this scale, it would be possible to add ‘ongoing or abandoned innovation activities’ again, which improves the correlation (r = 0.29). Only scaling enables this calculation. If, for example, ‘ongoing or abandoned innovation activities’ is added to the ten types of innovations, using or/or-statements, calculation of the correlation with ‘R&D-expenditure’ will again not be possible, because all innovators (category “1’) would be in that calculation (category ‘0’ has no R&D, as is the case with the other non-technological innovation types).

[xxxiv] Defining other groups, for example ISS < 4, gives comparable results.

Appendix A: HOMALS plots of relevant NACE and Firm Size selections

Big firms (250+ employees)

Figure 8. Plot of categories of innovation types, NACE; only big firms (250+ employees), the Netherlands (CIS4)



Observations and conclusions for big firms:

1. The difference between ‘Manufacturing’ and ‘Services Sector’ on the second dimension has increased substantially compared with figure 3; at first sight, it seems that the two sectors differ most on the technology dimension. However, it is not yet certain that this dimension still has the same meaning. Indeed one of the reasons for conducting HOMALS for the subgroups is to establish this. It is certain that big firms in the two sectors respond very differently to CIS, but how?
2. Notice that the second dimension has flipped; this is a result of finding other fitting category quantifications for this group of firms, it has no other meaning;
3. Just as in figure 3, in figure 8 the category ‘Manufacturing’ is closest to the positive categories of indicator ‘Prod1’ (New goods) and indicator ‘Proc1’ (New methods of manufacturing). However, an important difference between the two figures is that indicator ‘Mar1’ (New design or packaging) now discriminates on both dimensions, while first it did not; so indicator ‘Mar1’ is especially different for big firms, though it is difficult to say where this difference lies;
4. Once again the ‘non-technological indicators’ are located in the bottom right-hand corner of figure 8. However, some ‘technological indicators’ have now also become part of this group. So the second dimension is not a simple typology of ‘technological or non-technological innovation’ where big firms are concerned;
5. Again, the first dimension is ‘being innovative or not’; the second dimension needs another interpretation. It isolates mainly one more or less homogeneous group: big firms in ‘Manufacturing’ conducting ‘classical

innovation’. The pattern of big firms in the ‘Services Sector’, and of the group in the bottom right-hand corner in the figure, is difficult to interpret, though the fit of the solution is the still the same. Big firms in ‘Manufacturing’ are best characterised as goods producers. Big firms in the ‘Services Sector’ are best described as innovators of services. This is an important difference with figure 3;

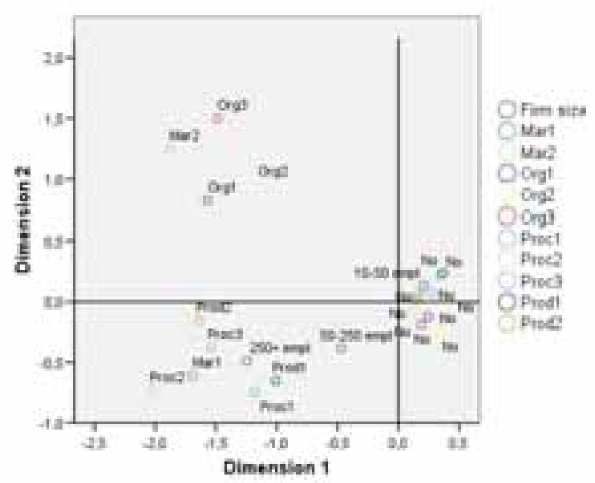
6. Big firms represent 3.4 percent of the total CIS (EU) population in the Netherlands. However, big firms are most important where ‘R&D expenditure’, for example, is concerned.

Manufacturing

Observations and conclusions for ‘Manufacturing’:

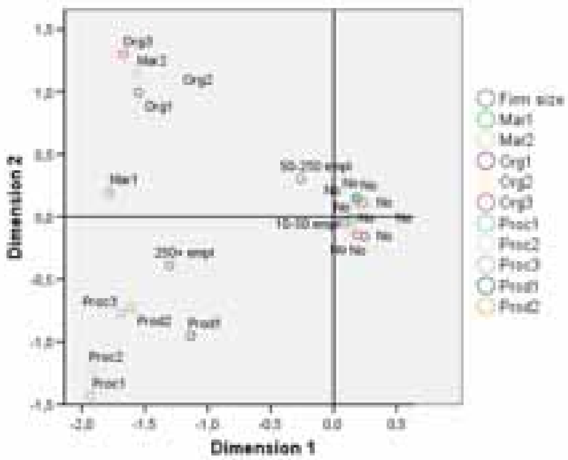
1. The plot is rather similar to figure 3, though the origin is now on the right-hand side. However, this is has no significance, so in total ‘Manufacturing’ factors of innovation are quite the same as in the total population. In ‘Manufacturing’ all ‘technological’ and ‘non-technological’ innovation indicators are homogeneous, as expected, except for indicator ‘Mar1’;
2. However, one difference is that the technological relevance of the positive category of indicator ‘Mar1’ has increased, which is displayed on the second dimension. So both in total ‘Manufacturing’ and big firms, indicator ‘Mar1’ has technological relevance, which is related to innovation of goods. However, in figure 3 the vector of the positive category of indicator ‘Mar1’ was near zero. In the ‘Services Sector’, as we will see later on, indicator ‘Mar1’ ‘behaves’ slightly more ‘non-technologically’.

Figure 9. Plot of categories of innovation types, Firm Size; only ‘Manufacturing’, the Netherlands (CIS4)



The ‘Services Sector’

Figure 10. Plot of categories of innovation types and Firm Size; only ‘Services Sector’, the Netherlands (CIS4)



The ‘problem indicator Mar1’ is now on the ‘non-technological side’ of the plot. However, it is not close enough to the other ‘non-technological indicators’ to consider it as a homogeneous group.
This difference in indicators ‘Mar1’ between ‘Manufacturing’ and the ‘Services Sector’ (figures 9 and 10) is caused by the fact that ‘innovating in goods and services’ (indicators ‘Prod1’ and ‘Prod2’) is reported together with ‘changing design or packaging’ (indicator ‘Mar1’) more frequently in ‘Manufacturing’ than in the ‘Services Sector’. Figures 3 and 8 prove this. A closer look shows that the distances between the categories ‘Manufacturing’ and ‘Services Sector’ on the one hand, and indicator ‘Mar1’ on the other are largest between the latter sector and indicator ‘Mar1’. The larger the distance, the less homogeneity exists. So knowing more about this is knowing more about how indicators ‘Prod1’ and ‘Prod2’ ‘behave’, in relation to indicator ‘Mar1’:

- In figure 3, indicator ‘Prod2 is close to indicator ‘Mar1’; however, both vectors are near zero on the second dimension, which makes this distance irrelevant;
- In figure 8, indicator ‘Prod2’ is plotted at a large distance from indicator ‘Mar1’;
- In figure 9, indicators ‘Prod1’ and ‘Prod2’ are plotted at nearly the same distance from indicator ‘Mar1’;
- In figure 10, indicator ‘Prod2’ is slightly closer to indicator ‘Mar1’.

So figure 8 is the crux. For this group of firms (250+ employees) in ‘Manufacturing’, innovation of goods must (nearly) always imply ‘changing design or packaging’. This proves that the third rotated factor (of the FA, see table 4) plays a major role for these types of firms. As this group of firms accounts for a very small percentage of the total CIS population, nearly all of them need to have reported indicator ‘Mar1’, otherwise this could not have been concealed in the HOMALS plots. Furthermore, the Eigenvalue of the third factor would be > 1 if ‘product innovations’ and indicator ‘Mar1’ in the ‘Services Sector’, and in smaller firms, correlated more.

Reliability of constructs for NACE and Firm Size
Is a supposed ‘marketing innovation’ scale reliable in the ‘Services Sector’, since in this sector indicator ‘Mar1’ has some ‘non-technological relevance’? Indeed, the reliability of the scale appears to be slightly better in the ‘Services Sector’. However, it is still far from sufficient (and therefore α is not given).
Calculation of all other RAs, for groups based on NACE and Firm Size, does not lead to changes in the main conclusions of this report.

Appendix B: Making the HOMALS-plot, syntax

```
MULTIPLE CORRES
VARIABLES=Prod1 Prod2 Proc1 Proc2 Proc3 Org1 Org2 Org3 Mar1 Mar2
/ANALYSIS=Prod1(WEIGHT=1) Prod2(WEIGHT=1) Proc1(WEIGHT=1)
Proc2(WEIGHT=1)
Proc3(WEIGHT=1) Org1(WEIGHT=1) Org2(WEIGHT=1) Org3(WEIGHT=1)
Mar1(WEIGHT=1)
Mar2(WEIGHT=1)
/MISSING=Prod1(PASSIVE,MODEIMPU) Prod2(PASSIVE,MODEIMPU)
Proc1(PASSIVE
,MODEIMPU) Proc2(PASSIVE,MODEIMPU) Proc3(PASSIVE,MODEIMPU)
Org1(PASSIVE
,MODEIMPU) Org2(PASSIVE,MODEIMPU) Org3(PASSIVE,MODEIMPU)
Mar1(PASSIVE
,MODEIMPU) Mar2(PASSIVE,MODEIMPU)
 /DIMENSION=2
 /NORMALIZATION=VPRINCIPAL
 /MAXITER=100
 /CRITITER=.00001
 /PRINT=DISCRIM
 /PLOT=JOINTCAT( Prod1 Prod2 Proc1 Proc2 Proc3 Org1 Org2 Org3 Mar1
Mar2 )
 (20) DISCRIM(20) .

* Notice that the normal codes such as INPDGD (New goods), INPDSV (New
services) etc. are used  * in the  syntax as: Prod1 Prod2 Proc1 Proc2 Proc3
Org1 Org2 Org3 Mar1 and Mar2. This is a result
* of recoding. The normal codes (usually) consist of values “0” and “1”.
HOMALS cannot deal with
* value “0”, so, Prod1 Prod2 Proc1 Proc2 Proc3 Org1 Org2 Org3 Mar1 and
Mar2 consist of values
* “1”and “2”, done by:

AUTORECODE
  VARIABLES=INPDGD INPDSV INPSPD INPSLG INPSSU ORGSYS
ORGSTR ORGREL MKTDES
  MKTMET  /INTO Prod1 Prod2 Proc1 Proc2 Proc3 Org1 Org2 Org3 Mar1 Mar2
  /GROUP
  /PRINT.
```

* So if necessary the autorecode-command precedes. Running HOMALS may take some time. The
* resulting plot will be a mess, as the value and variable labels have not yet been adjusted. To do so:
* Change value labels of Prod1 Prod2 Proc1 Proc2 Proc3 Org1 Org2 Org3 Mar1 Mar2 into:
* "No" for category “1” and "Variable name" for category “2” of each indicator in the data file. For
* example, the second category of 'Prod1' is assigned value label "Prod1", so this label will be plotted.
* Repeat the MULTIPLE CORRES command.
* Click the plot to add x- and y-axes (use 0,0).

****
* Reweighing the data, as is explained in note 2 in document.
****
* To do this, use your created normal weight factor (use only the EU NACE cases for data delivery to * EUROSTAT, in order to assure comparability).

```
COMPUTE corweigh = WEIGHT04 * 5927 / 27412.
EXECUTE.
WEIGHT
  BY corweigh.
```

* Explanation: WEIGHT04 is the normal weight factor in Dutch CIS4-file, sum = 27,412 (sum
* only EU NACE cases). This sum must be replaced by the sum of weights of your country.
* Sample size, number of cases (CIS, EU NACE) = 5,927 (Netherlands). To be replaced by the
* number of cases of your country (again, only EU NACE cases).
* Repeat the MULTIPLE CORRES command.

****
* Other variables in the HOMALS plot, such as NACE, Firm Size and R&D.
****

* If you want to add NACE and FIRM SIZE in the plot: 1) You can put these variables in the MULTIPLE
* CORRES syntax under the subcommands VARIABLES, ANALYSIS, MISSING and PLOT, or 2) You * can use the menu: go to ANALYZE / DATA REDUCTION / OPTIMAL SCALING and now: choose the * two bullets (this is default): first: “All variables multiple nominal’ and second: ‘one set’. After this click * DEFINE. Put all desired variables in ‘Analysing variables’. Go to the ‘button’ OUTPUT and choose
* ‘Discrimination measures’ (only this one of the five tables is needed). Click CONTINU. Go to the
* button OBJECT: nothing is needed here, while probably ‘object point’ is ‘on’, if so, remove this. Go to * the button VARIABLE and select all variables and drop them into the box ‘joint category plots’. Click
* CONTINU. Click OK.

* If you want to add ‘R&D expenditure’ (RTOT) in the HOMALS plot, by categorising ‘R&D’, it is quite a
* puzzle to search for the best fitting categories. The ‘best’ fitting categories are the categories which
* lead to the plots which are expected in theory. In the Netherlands, the plots are best when, at least, * isolating the top R&Ds, e.g. the 5 percent highest

R&D-spenders (as R&D is extremely skewed), and * the rest e.g. above median, below median. It is also a solid option to first exclude R&D = 0, since * many firms report this value; then finding the best R&D categories may lead to better results.

****

Appendix C: Additional possible improvements of CIS

This appendix suggests some general improvements that do not belong to the purpose of this Microdata project.

1. Why not examine the relationship between types of innovation in CIS? This will improve our knowledge of *innovation flows*. Such a flow is a cluster of interdependent innovations. Interdependent means A causes B. The ISS consists of innovations flows and innovations which are not interdependent. For example, an innovation flow could be: changing the methods of manufacturing ('Proc1') as a result of innovation in goods ('Prod1'). So, in this case the firm had to adjust some processes to be able to produce the new good. Many interdependent combinations may exist and they may differ between NACE classes and Firm Sizes. Adding a simple question for firms, about which of the reported innovation types were interdependent, would provide this information;
2. Indicators 'Prod1' and 'Prod2' measure the market introduction of a new, or significantly improved, good or service. In CIS, firms which have introduced one or three new goods, or services, are treated in the same way. Is this justified? Being able to distinguish firms on this matter would lead to an indication of '*product innovation productivity*'; a simple question regarding the number of 'product innovations' would solve this. This information would also lead to new insights;
3. This report has demonstrated that in the Netherlands the first indicator of 'marketing innovation' does not measure 'non-technological' content. However, the second indicator of the construct can be improved as well. This indicator – 'changed sales or distribution method' – is quite similar to the second indicator of 'process innovation' – 'new logistics, delivery or distribution'. Both indicators use the term 'distribution', so where should a firm report an innovation of distribution? A simple adjustment would solve this problem. The same survey is used in CIS4.5.

Bibliography

Hays, W.L. (1994), "Statistics", International Edition, University of Texas at Austin, fifth edition, Hartcourt Brace College Publishers.

Heus, P. de, Leeden, R. van der & Gazendam, B. (1995), "Applied Data Analyses. Techniques for Social Science", Utrecht: Lemma BV Publishers.

Meerling (1989), “Methods and Techniques of Psychological Research, part 1. Model, Observation and Decision, Meppel: Boom, Leiden: Faculty of Social Sciences, Department of Psychology.

Meerling (1989), “Methods and Techniques of Psychological Research, part 2. Data Analyses, and Psychometrics, Meppel: Boom, Leiden: Faculty of Social Sciences, Department of Psychology.

Walsh, W.B. & Betz, N.E. (1995), “Test and Assessment”, Ohio State University, third edition, Prentice Hall, Englewood Cliffs, New Jersey.