The internal differentiation of the KIBS sector: empirical evidence from cluster analysis

Daniel Horgos*

Helmut Schmidt University
University FAF Hamburg, Germany
E-mail: horgos@hsu-hh.de
*Corresponding author

Andreas Koch

Institute for Applied Economic Research
Tübingen, Germany
E-mail: andreas.koch@iaw.edu

Abstract: The sector of Knowledge-Intensive Business Services (KIBS) is characterised by high rates of firm fluctuation, rapid changes in technological progress (*e.g.*, in the software industry) as well as high interdependencies between subsectors (*e.g.*, consultancy and engineering). These features give rise to fuzzy internal and external sectoral boundaries and make it difficult to apply common, basically output-oriented industry classifications. Although conventional taxonomies, such as the NACE or the ISIC, are indispensable in many respects (*e.g.*, for comparative studies), it is worth considering classifications falling back on alternative criteria. In the present paper, we perform a cluster analysis on a sample of 547 German KIBS firms. The study takes into account fundamental characteristics of these firms, *e.g.*, interaction patterns and innovation behaviour. The resulting classification (seven distinct groups) reveals that the examined service firms can also be differentiated by using firm-internal attributes rather than the services provided. Finally, the new taxonomy is tested with several regression models showing that our classification is a viable alternative to conventional industry classifications.

Keywords: knowledge-intensive business services; KIBS; classification; cluster analysis; typology; innovation; spatial proximity.

Reference to this paper should be made as follows: Horgos, D. and Koch, A. (2008) 'The internal differentiation of the KIBS sector: empirical evidence from cluster analysis', *Int. J. Services Technology and Management*, Vol. 10, Nos. 2/3/4, pp.190–210.

Biographical notes: Daniel Horgos has been a Research Associate at the Helmut Schmidt University, University of the Federal Armed Forces, Hamburg, Germany since March 2005. He studied Business Administration and Computer Sciences at the University of Applied Sciences in Pforzheim and Economics at the University of Tübingen, focusing on quantitative methods and economic policy. His current research focuses on international trade, international outsourcing and labour market effects.

Andreas Koch has been a Research Associate at the Institute for Applied Economic Research (www.iaw.edu) in Tübingen, Germany since March 2003. He studied geography, urban planning, and public law at the Universities of Hannover, Tübingen, and Rio de Janeiro focusing on economic geography and development studies. In his dissertation (2006), he analysed the role of employee start-ups in the German KIBS sector. Currently, he focuses on entrepreneurship research, regional development and innovation studies.

1 Introduction

Within the last decades, the service sector has gained increasing importance in most of the world's economies and has itself undergone a very dynamic development. In Germany, for example, the service sector contributed by 69.1% to the nation's Gross Value-Added (GVA) in 2006, whereas in 1970 this rate was still below the 50% margin. At the same time, the manufacturing sector's contribution to the German GVA diminished from 48.4% in 1970 to 29.9% in 2006. Equally, significant rises in the number of service firms, and, even more accentuated, in the number of service employees can be stated. Between 1970 and 2006, the number of employees in the service sector in Germany increased from 12 million to 28.3 million persons (Federal Statistical Office of Germany, 2007).

These more general economic developments of the service sector go hand in hand with fundamental and continuous changes of its internal structure: First, the service industries are undergoing a process of continuous diversification; second, within the service sector, there is a shift from personal services to business (producer) services; third, services and service firms are characterised by growing degrees of sophistication, differentiation and an increasing significance of innovation activities; and, fourth, existing subsectoral boundaries are often broken up by firms acting at the interfaces between subsectors and providing services and goods from different areas, for example, software development *and* business consulting. In the course of these developments, particularly the sector of Knowledge-Intensive Business Services (KIBS) has gained increasing economic weight and has recently attracted more and more scientific attention (see, for example, Czarnitzki and Spielkamp, 2003; EMCC, 2005; Freel, 2006; Koch and Stahlecker, 2006; Koch and Strotmann, 2006; Koch and Strotmann, 2008; Miles *et al.*, 1995; Martinez-Fernandez, 2006; Muller and Zenker, 2001; Toivonen, 2004; Wiig Aslesen and Jakobsen, 2007).

The outlined intense and impetuous changes entail a series of consequences most of which cannot be discussed in the present paper. Yet, from an analytical point of view, one severe implication is the growing intricacy of statistically accounting and systematising the rapidly and constantly shifting and diversifying sectoral and subsectoral structures and boundaries of the service sector. Existing approaches of industry classification as the *International Standard Industry Classification* (ISIC, cf. United Nations, 2002) or the *Nomenclature Générale des Activités Économiques dans les Communautés Européennes* (NACE, cf. European Commission, 2007), are not fully capable of covering the evolving structures and of accompanying the rapid pace of change in these new sectors.

The ISIC, which is also the basis for various other classifications as the NACE, claims to differentiate economic sectors along *economic activities,* defined as "the combination of actions that result in a certain type of products […] Thus, an activity is characterised by an input of resources, a production process and an output of products" (United Nations, 2002, p.12) and the following three aspects are stated to be the basis of the classification scheme: "(a) the character of the goods and services produced, (b) the uses to which the goods and services were put, and (c) the inputs, the process and the technology of production" (United Nations, 2002, p.14). Nevertheless, outputs produced by firms seem to be the most important ingredient of the classification, as "the principal activities of the unit in general can be determined from the goods that it sells or ships or the services that it renders to other units or consumers" (United Nations, 2002, p.21).

Besides focusing mainly on the outputs of goods and services, another principal problem of standard industry classifications is that, like for example in Germany, most industry statistics refer to enterprises as basic units of analysis and existing data does not allow for differentiation of separate units of economic activity within these enterprises. Furthermore, standard industry classifications are rather static regarding both the limits between existing economic activities as well as the incorporation of new economic activities, as intervals between revisions are not able to keep up with developments in reality. The probably best-known example illustrating the problems of the conventional classifications is the sector of Information and Communication Technology (ICT): both the external limits and the internal structure of this sector recently have been a constantly recurring subject of scientific discussion (cf. Atzema, 2001; OECD, 2004). Similarly, the KIBS sector is not consistently defined and systematised in the literature (see, for example, Miles *et al*., 1995; Nählinder and Hommen, 2002; for a recent overview, see Koch and Stahlecker, 2006).

The contribution at hand will address these problems and shortcomings of existing industry classifications using the KIBS sector as an example. Based on data of the *KIBS Foundation Survey* – a data set of nearly 550 KIBS firms in three German regions – we will outline an alternative way of classifying new firms in this sector by performing a cluster analysis.[1] Our analysis is based on other than the traditional output characteristics to classify the firms, addressing particularly the central characteristics of the KIBS sector, namely, the roles of knowledge, innovation, and interaction. It is not the aim of this paper to establish a new universal classification scheme for industrial sectors; we rather want to point out an alternative way of identifying types of firms in this dynamic new service sector.

The remainder of the paper is structured as follows: Section 2 will outline the basic characteristics of the KIBS sector and address the key features differentiating the firms of this sector. The potentials and in particular the shortcomings of the identification and the internal structure of the KIBS sector in the NACE will be clarified. Section 3 then introduces the data set to be analysed and describes the empirical classification process and its results. Section 4 comprises econometric regression models to test whether the resulting cluster structure is a possible alternative to the conventional industry classification. Finally, Section 5 concludes the paper.

2 The KIBS sector: characteristics, boundaries and differentiation

The sector of KIBS is – like the ICT sector – not explicitly displayed in conventional industry classifications. Nevertheless, the external boundaries of the KIBS sector according to conventional industry classifications are rather consistently defined across different studies. The mainstream of existing research includes the sectors displayed in Table 1 (*e.g.,* Miles *et al*., 1995; Nählinder and Hommen, 2002; Freel, 2006).

Table 1 The KIBS sector in the NACE (rev. 1, 1993)

Technical KIBS	
721	Hardware consultancy
722	Software consultancy and supply
723	Data processing
724	Data base activities
725	Maintenance and repair of office, accounting and computing machinery
726	Other computer related activities
731	Research and experimental development on natural sciences and engineering
742	Architectural, engineering and other technical activities
743	Technical testing and analysis
Professional KIBS	
732	Research and experimental development on social sciences and humanities
741	Legal, accounting, bookkeeping and auditing activities; tax consultancy; market research and public opinion polling; business and management consultancy; [holdings*]
744	Advertising

Note: * Not included in our definition of the KIBS sector.

Source: European Commission (2007), own compilation

However, this classification of KIBS includes quite heterogeneous types of services transcending by far the displayed differentiation between technical KIBS and professional KIBS, which dates back to Miles *et al.* (1995). Particularly, it has to be mentioned that sectoral boundaries are rather fuzzy and that quite a lot of firms operate at the interfaces between different subsectors. For example, many technical consultants also provide business consultancy or vice versa and bookkeeping and software development are strongly intertwined activities often performed by one single firm and not separable in the balance sheets. Another fact limiting the practicability of the conventional industry classification is that the KIBS sector is characterised by high rates of innovation and a rapid pace of technological change. Notwithstanding, the firms in the KIBS sector have important commonalities, which are not adequately reflected in the internal structure of the given classification.

Some of the most central characteristics of KIBS firms are:

- the outstanding significance of knowledge – both codified and tacit – and the resulting realisation of *innovative activities* (knowledge intensity)[2]

- the orientation of the services towards other firms or organisations and the resulting high relevance of interactivity (defined here as *functional integration*), and, conjoint with the previous
- the importance of *spatial proximity* between KIBS firms and their providers and clients (cf. Illeris, 1994; Koch and Strotmann, 2006; Koch and Strotmann, 2008; Miles *et al*., 1995; Toivonen, 2004).

These issues stand in the centre of the present analysis and they are the constituting factors of our cluster analysis.

Regarding *innovation*, not only the rapid development of the new information and communication technologies, but also changing management paradigms, globalisation and serious shifts in production processes and in the division of labour have led to an increasing significance of innovation and innovative capabilities in service firms. The KIBS sector is not only an outcome of these developments, but also one of the driving forces behind this trend. Recently, it has been claimed that KIBS are not only external knowledge sources for their clients, but that they are increasingly becoming independent innovators (Cainelli *et al*., 2006; Czarnitzki and Spielkamp, 2003; Gallouj and Weinstein, 1997).[3] Innovation, *i.e*., the provision of new-to-the-market or new-to-the-firm services, is thus of crucial importance in the KIBS sector and different degrees of innovative performance can be regarded as a distinguishing factor of firms in this new sector.

It has often been claimed that both innovative activity and the provision of business services require specialised knowledge and cumulative learning processes, which can only be realised by intense *interaction* between service suppliers and clients (Johannisson, 1998; Lundvall, 1988; Strambach, 2002). Furthermore, as KIBS mostly provide highly application-oriented services, implicit knowledge plays an important role. For the acquisition of this type of knowledge, cooperation, trust, communication and face-to-face contacts are very important (Howells, 2002). Thus, intense interaction between service providers and clients as well as between service providers and other knowledge sources is believed to be a second distinctive factor in the KIBS sector.

Owing to the high importance of innovation and interaction in the KIBS sector, firms frequently seek vicinity to their customers and/or suppliers and often locate in *spatial proximity* to their customers (Illeris, 1994).[4] With regard to the outstanding roles of specialised, applied and particularly implicit knowledge, spatial proximity frequently roots in the history of firms in the KIBS sector. Many firms in this sector are initiated by persons who have previously been working for other (often manufacturing) firms and start their new firm close to the place they have been working and living before; in many cases, their former employer is an important client in the initial stages of their new firm. Therefore, spatial proximity can be seen as the third distinguishing factor between firms in the KIBS sector.[5]

These outlined characteristics may be more instructive for internally differentiating the KIBS sector than the traditional classification schemes resulting in the ostensive, but rather theoretical distinction between technical KIBS and professional KIBS outlined in Table 1 and applied in wide parts of the literature. In the following, we will thus present an alternative typology of firms in the KIBS sector based on the criteria of innovation, functional integration and spatial proximity.

3 Classification of firms in the KIBS sector

The present section describes the empirical classification of the KIBS firms in our sample. We start by briefly describing the applied dataset (3.1) and motivating the selection of variables that will be taken into account in the clustering process (3.2). In Section 3.3, we present the classification process. Thereby, the ‘mixed-data problem’ arises. To solve this problem, we focus on the most appropriate solution, the quantification of all variables before going into the clustering process. Finally, we figure out and discuss the results of the discovered classification (3.4).

3.1 Data: the KIBS foundation survey

The analyses of the paper at hand are based on the KIBS Foundation Survey, a data set consisting of young KIBS firms in three German agglomeration regions (Bremen, Munich and Stuttgart).[6] The firms have been selected from a sample consisting of address data provided by the Chambers of Industry and Commerce in the respective regions including all firms classified under the NACE-Codes 72, 73 and 74 (see Table 1),[7] which have been founded between 1996 and 2003.[8] We considered only genuine foundations listed in the trade registers and, thus, subsidiaries, branch offices, firms arising from mergers and acquisitions; firm reformations were excluded from our survey.

Based on these definitions, the population size in the three regions was 7714 firms. A random sample of 2108 firms, stratified on the 3-digit sectoral level was drawn and their actual managers (which mostly had started the firm) were questioned in telephone interviews. Altogether, 547 interviews could be successfully conducted resulting in a quite satisfactory rate of return of 26%. The interviews were based on a standardised questionnaire, which covers a large variety of detailed questions concerning individual attributes of the firm manager (*e.g.*, context of business idea, former occupation and location of workplace, skills, *etc.*), initial characteristics of the firm at its origin as well as its development over time.

In addition to the address information, the original firm data includes also the *sectoral affiliation* of the respective firms on a 3-digit level according to the German national industry classification (WZ93), which is consistent with the NACE in the surveyed sectors. As the questionnaire of the KIBS Foundation Survey also included an open question about the main *activities* of the surveyed firms, we are able to compare the official classification provided by the Chambers of Industry and Commerce with the self-assessment given by the interviewees. Table 2 illustrates the low rates of concordance between the two classifications.

The self-assessment of the surveyed firms and the official classification are quite different with rates of concordance of only 47% on the 3-digit level and 70.1% on the 2-digit level. Furthermore, quite a few of the interviewed firm leaders stated that they acted at the interfaces between two or more of the subsectors of the official classification. Most common are combinations between software development (722) and consulting activities (741) or any kinds of combinations inside the sector of computers and related activities (72). These facts point to a weakness of the official sectoral classification of the firms in the KIBS sector and are a further argument for developing an alternative typology of the firms in this new sector.

Table 2 NACE-classification and self-assessment of KIBS firms (N = 542)

NACE-code	*Percentage of firms according to official classification*	*Percentage of firms according to self-assessment*	*Percentage of congruence between official classification and self-assessment*
721	7.50	1.46	4.88
722	17.00	16.64	41.94
723	3.29	1.65	22.22
724	1.83	0.37	0.00
725	0.55	5.85	33.33
726	7.13	1.10	0.00
731	4.57	1.65	32.00
732	0.18	0.91	100.00
741	22.30	36.01	69.67
742	17.00	11.70	53.76
743	1.83	4.75	40.00
744	16.45	17.37	70.00
Total	100	100	46.98

Source: KIBS Foundation Survey, own calculations

3.2 Selection of variables

To classify the firms in the KIBS sector based on the above-mentioned economic characteristics rather than on the criteria employed by the standard industry classifications, we apply the statistical tool of a cluster analysis. KIBS provide knowledge-intensive, innovative services for other private or public organisations. As outlined above in Section 2, the central characteristics of firms in the KIBS sector are their innovative behaviour, the interactive nature of their activities and the high significance of spatial proximity to clients and providers, which is particularly relevant in the early years of the firms' development. (see, for example, Illeris, 1994; Wiig Aslesen and Jakobsen, 2007). The cluster analysis below, designed to test whether significant classes of KIBS can be identified, is based upon these characteristics.

First, it is necessary to identify the variables that may serve as indicators for these three factors and, thus, can be measured within our classification process (see Table 3). In the case of *innovation*, we chose three dummy variables resulting from a question asking the firms whether they provide totally new services, improved services or none of the two. Moreover, we have information on R&D expenditures,[9] which can be seen as an input factor for innovative behaviour. Two variables indicate the integration of a KIBS firm into its environment. As the dataset does not contain variables directly indicating the *functional integration* and interaction of a firm with other firms, we use both the professional background of a new firm's founder as well as the utilisation of results from activities prior to the firm foundation as proxies for the interactive activities of a firm. It is assumed that firm founders with prior experience in the private sector (as employees, freelancers, or self-employed) represent a higher integration than founders with a background from the public sector, *e.g.*, from universities). Moreover, we suppose that the utilisation of ideas, technologies or finished products developed in earlier job positions implies a higher degree of integration of a KIBS firm into its environment. Last but not least, five variables indicate the role of *spatial proximity*: the share of turnover effectuated in the district of the firm's location, the location where a firm purchases the majority of services, the number of regional partners, the importance of a regional lead customer and the regional provenance of the firm's founder.[10]

Table 3 Variables applied for the clustering procedure

Variable	*Scale*	*Values*
Innovation		
Development of totally new services	dummy	0 = no/1 = yes
Improvement of existing own services	dummy	0 = no/1 = yes
No own innovations	dummy	0 = no/1 = yes
R&D-expenditures (share of total turnover in %)	metric	min 0/max 400
Functional integration		
Professional background of the founder (activity of the founder immediately before start-up)	ordinal	1 = public sector
		2 = private sector
		3 = freelancer, self employed
Utilisation of results within the firm founding process	ordinal	0 = no
		1 = ideas
		2 = technologies, patents
		3 = finished products
Proximity		
Share of regional turnover (in % of total turnover)	metric	min 0/max 100
Purchase of knowledge intensive services	ordinal	0 = no
		1 = yes, mostly from outside the region
		2 = yes, mostly from the region
Number of regional partners	metric	min 0/max 7
Importance of a regional lead customer	ordinal	0 = no regional lead customer
		1 = regional lead customer available
		2 = important regional lead customer available
Regional background of the founder	ordinal	1 = from abroad
		2 = from the federal republic
		3 = from the federal state (the German 'Länder')
		4 = from the region

Note: * The term 'region' or 'regional' refers to the Planning Region (*Raumordnungsregion*) where the respective firm is located.

One of the strengths of the cluster analysis is the possibility of considering a large number of variables. However, this causes several empirical problems, one of which is known as the 'mixed-data problem', which also arises in our analysis. The variables we selected are differently scaled and can therefore not be directly considered for clustering. Thus, we tested three ways to solve this problem. However, we reject the way of generating several dummy variables, since this strongly increases the number of variables while a lot of information gets lost. Owing to the lack of possibilities in statistical software packages, we also reject a second way of solving for the mixed-data problem – the consideration of all the different scale levels. Thus, we decided to quantify the variables as a third solution. The details of this commonly used procedure are described below.

3.3 Cluster analysis

To generate quantitative variables, all variables that are not dichotomously or metrically scaled have to be marginally modified. Nominally scaled variables are transformed to an ordinal scale level. Owing to the natural order of these variables with respect to the three dimensions of innovation, functional integration and spatial proximity, the additionally required information can be assumed. For example, the values of the variable 'regional origin of the founder' have been reversed so that higher values indicate a higher significance of spatial proximity within the start-up process of the firm. Moreover, some variables have been aggregated ('Importance of regional lead customer' and 'Contribution of results in the founding process') and others have been condensed ('Professional background of the founder'). With the assumption of additional information, the newly created variables are dichotomously, ordinally or rationally scaled. As Bacher (1996, pp.186, 232) notices, with respect to cluster analysis, all these variables can be handled like quantitative variables and thus, the proximity of these types of variables can be measured using a quantitative method. Table 3 presents the 'quantitative' variables that will be considered for clustering.

Owing to measurement differences, the generated variables have to be standardised to become usable within an object-orientated cluster algorithm. Thus, new variables with a mean of 0 and a variance of 1 are formed. Since some variables have undefined upper limits, the theoretical z-transformation, which is mostly recommended by the literature, cannot be applied in our case. Instead, we adopt an empirical z-transformation to achieve standardised variables (cf. Aldenderfer and Blashfield, 1984, p.20; Everitt *et al*., 2001, p.51f; Gordon, 1981, p.24f).

Since variables with missing values have the same codification, the clustering process would lead to strongly biased results if they were included into the analysis. Thus, in the case of cluster analysis, an adequate treatment of missing values is very important. One possibility is to exclude all cases with missing values in any variable. The disadvantage of this method is the reduction of the sample. However, compared with other methods, it seems appropriate for our data situation.[11] Owing to these restrictions, our sample reduces from 547 to 390 firms.

To identify possible outliers, we use the single-linkage method as a first step of our classification process. The single-linkage method, developed by Sneath (1957), clusters the two firms with the smallest 'distance' in each step. Afterwards, the new distances between the remaining objects or clusters and the new cluster are calculated, using the smallest value of the separate distances. Thus, in the resulting clusters, each object is forced to have at least one nearest neighbour. This is a minor assumption concerning the homogeneity within the clusters. Owing to the fact that the single-linkage method generates few big groups and several small groups, outliers can be identified easily. One disadvantage of the single-linkage method is the property of building chains. Thereby, different groups will be combined to a cluster if two single representatives have a big similarity. We implement the single-linkage algorithm after calculating the distances as Euclidean distances. Thereby, two objects are identified as outliers and will be eliminated before starting the next steps (cf. Aldenderfer and Blashfield, 1984, p.38ff; Kaufman and Rousseeuw, 1990, p.47f; Bailey, 1994, p.55ff; Gordon, 1981, p.34ff).

Owing to the described disposition of the single-linkage algorithm to build chains we only use this method to identify the outliers, but not for the main clustering process. Since the cluster analysis is a method for discovering structures, no information and no ideas about the resulting typology are available in advance. Thus, it is difficult to test which cluster algorithm yields the most appropriate result. In such a situation, the agglomerative Ward Algorithm leads to very good classification results since certain assumptions, *i.e.*, the elimination of outliers before clustering, the existence of quantitative variables or expected similar cluster sizes, are met. As these preconditions are fulfilled in our case, we will apply the Ward Algorithm for the main clustering process (cf. Bergs, 1981, pp.83–106).

This method, developed by Ward (1963), is widely accepted in the economic literature. It differs from other clustering methods, *inter alia,* in the procedure of finding the two objects or clusters that are going to be merged. Therefore, it combines the two objects or clusters that cause the fewest enlargement of a specific heterogeneity measure, similar to the sum of the squared errors. Thus, the Ward Method generates very homogeneous clusters. For our main analysis, we use the Ward Method and measure the distances with the squared Euclidean distances (cf. Anderberg, 1973, pp.42ff and 142–145; Bacher, 1996, pp.217, 222; Gordon, 1981, pp.39–53).

Since the growth of the heterogeneity measure is quite high from seven to six groups, we decide to extract seven clusters as our main result, while a solution with three or ten clusters would also be possible with respect to the development of the heterogeneity.[12] The different cluster solutions can be seen as different aggregation levels. However, in the following, we focus on describing and interpreting the result with seven groups in an economical as well as a statistical way, using the cluster mean values and calculated cluster z-values with respect to the eleven selected variables. With the z-values it can be shown whether a cluster is situated above or beneath the mean of any specific variable (cf. Bacher, 1996, p.184f).[13] In the following section, we will briefly describe the resulting seven clusters that have been discovered in the course of our cluster analysis.

3.4 Seven clusters to differentiate types of KIBS

The Cluster 1 (C1) contains 114 firms. It is the biggest cluster with an ordinary to high degree of innovation, a high significance of spatial proximity (the highest throughout all clusters), but only moderate degrees of functional integration. A big share of the respondents stems from the region, significant parts of the turnover are effectuated within the region, knowledge-intensive services are purchased from the region and there exist many regional partners. The moderate functional integration is indicated by large numbers of interviewees coming from the public sector and by the fact that only ideas (and no products or technologies) from former employment are involved into the process of establishment of the firms.

A total of 65 firms constitute the Cluster 2 (C2). These firms are characterised by a high degree of innovativeness as well as extraordinary importance of both functional integration and spatial proximity. Although R&D expenditures are relatively low, own services are improved, and new, own services are being developed. The high significance of functional integration of cluster C2 is reflected by the fact that the private sector is the professional background of most of the firm leaders. The solutions involved in the founding process frequently are finished products. Most of the respondents of these firms start their venture in the region they have been living or working before. Also, the purchase of knowledge intensive services, the amount of the partners in the region and the importance of a regional lead customer are above the average. Anyhow, the firms show only low shares of regional turnover.

Cluster 3 (C3) contains 69 firms. These are ordinarily innovative and neither spatial proximity nor functional integration do play significant roles. Also, with respect to spatial proximity, the firms lie below the average.

Cluster 4 (C4) is the most ‘normal’ cluster. The 47 firms in this group are characterised by an ordinary innovativeness as well as by average values regarding functional integration and spatial proximity. Most firm leaders have been working in the public sector before and many of them transferred technologies or patents from the former employment into their new firm.

Ordinary innovation patterns, an average level of functional integration, and a low significance of spatial proximity are the main characteristics of the 54 firms included in Cluster 5 (C5). Own innovations exist in the form of an improvement as well as a development of own new services. The leaders of these firms do not hail from the region. This indicates the low spatial proximity, which is also reflected by the other regional variables.

Cluster 6 (C6) is a very small cluster with only two firms. These are innovative with a very high R&D share of turnover. They show very low values regarding the role of spatial proximity and they are characterised by a low significance of functional integration. It is important to mention that the two firms in C6 are not outliers in respect to our definition. The two firms have not been merged at the end of the clustering process with a sharp increase of the heterogeneity measure. They were identified as an independent cluster during the classification process.

Last but not least, Cluster 7 (C7) contains 37 firms. These are particularly characterised by not performing any innovations. Moreover, spatial proximity plays an average role and functional integration is rather high. The majority of the respondents stems from other regions than the one in which their firm is located. The professional background of the firm leaders is the private sector and in many cases technologies or patents have been transferred from former activities. Table 4 summarises the results of the cluster analysis.

To complete the classification procedure we test whether the extracted cluster structure is correlated with the traditional NACE classification scheme or if these two typologies are totally uncorrelated. Applying contingency analysis we obtain that the typologies are correlated and thus do not refer to completely separable situations. Since the aim of our contribution is not to refute the traditional classification but to extract an alternative typology, this result is highly desirable. However, despite the correlation, there still exist important differences.[14] Thus, the next section compares the quality of our classification result with the quality of the traditional NACE classification.

Table 4 Characteristics of the clusters

Cluster	*Size*	*Innovation*	*Functional integration*	*Spatial proximity*
C1	n = 114	▲	▼	▲
C2	n = 65	▲	▲	▲
C3	n = 69	▶	▼	▼
C4	n = 47	▶	▶	▶
C5	n = 54	▶	▶	▼
C6	n = 2	▲	▼	▼
C7	n = 37	▼	▲	▶

Notes: ▲ = high significance, ▶ = ordinary significance, ▼ = low significance.

Source: KIBS Foundation Survey, own calculations and illustration

4 Quality tests of the new classification

The idea of this section is to prove whether the result of our cluster analysis represents an adequate classification of the surveyed firms, and whether it is consequently a possible alternative to the conventional output-orientated industry classification. A ‘good’ classification is characterised by firms similar within but different between groups with respect to important economic variables, *e.g.,* the growth of turnover or the development of employment. Thus, we investigate if economically important firm characteristics allow a separation of the discovered clusters and we compare the results of the quality tests with the results achieved with the traditional NACE classification.

The conventional NACE classification is represented by the variable ‘nace’. This variable consists of the categories ‘Software’, ‘Other Electronic Data Processing’, ‘Technical Services’, ‘Consulting’ and ‘Advertising’.[15] As variables of our classification result, we consider ‘cluster3’, ‘cluster7’ and ‘cluster10’, which are the results of extracting three, seven or ten clusters. Since these different cluster results present different aggregation levels of one typology (similar to 2-, 3-, or 4-digit industry classification) we consider all of them to achieve further information on the robustness of our econometric analysis. Since the group C6 of ‘cluster7’ and ‘cluster10’ contains only two firms, we exclude it in the analysis below in order to avoid estimation bias.

For the important economic variables, which should be able to distinguish the different groups of the classification results, we consider employment (in 2003), the annual growth of employment, the number of freelancers (in 2003), the annual growth rate of the number of freelancers, the turnover (in 2002) as well as the annual growth rate of turnover. Since all these variables are metrically scaled, we apply several multinomial logistic regression models, with the classification results considered as endogenous and the different economic variables as exogenous variables.

First, we apply a likelihood-ratio-test (also called ‘model chi-square-test’) to investigate the goodness of fit of the different logit models. With the traditional NACE classification (p-value 0.02) as well as with our cluster results of seven groups (p-value 0.04) and ten groups (p-value 0.02), the selected variables can significantly distinguish between the different groups and thus, the overall model is well specified. With a p-value of 0.22, the result of the three-cluster solution shows no significant impact of the economic variables to distinguish between the groups. However, since the results below turn out to be quite interesting, and since the three-groups classification is only a different aggregation level of our extracted typology, we will continue presenting the econometric estimation results for the three-cluster solution as well.

After testing the goodness of the overall models, we apply a likelihood-quotient-test (an extension of the likelihood-ratio-test) to investigate the ability to distinguish between the different clusters for each individual economic variable. The H0-hypothesis of this kind of estimations states that the considered variable does not significantly influence the classification at stake. The results of our logit estimation are presented in Table 5.

Table 5 Likelihood-Quotient-Test

Characteristics	*nace*	*cluster3*	*cluster7*	*cluster10*
Employment in 2003	0.52	0.12	0.24	0.27
Annual growth rate of employment	0.70	0.10	0.07(*)	0.04(**)
Number of freelancers 2003	0.60	0.41	0.64	0.84
Annual growth rate of number of freelancers	0.84	0.60	0.30	0.47
Turnover in 2002	0.63	0.05(**)	0.08(*)	0.14
Annual growth rate of turnover	0.39	0.05(**)	0.03(**)	0.02(**)
Share of employees with university degree	0.01(***)	0.04(**)	0.15	0.02(**)
Constant term	0.08(*)	0.00(***)	0.00(***)	0.00(***)

Note: P-values (significant at * = 10%, ** = 5%, *** = 1%).
Source: KIBS Foundation Survey 2003, own calculations

The table shows that the typology resulting from our cluster analysis is quite a good alternative to the traditional output-orientated NACE classification. All three cluster results, and thus aggregation levels, have a stronger correlation with the considered economic variables than the traditional classification. The estimation shows for the NACE only a significant influence of the amount of employees with university degree. Important firm characteristics such as the development of the amount of employees, the actual value of turnover or the development of the value of turnover cannot distinguish between the different groups. By contrast, the results of our cluster analysis show a high statistically significant relation with important economic variables, such as the turnover or the development of employment. On the classification with three groups, the variables ‘employees with university degree’, ‘turnover in 2002’ and ‘development of turnover’ have significant impact, indicating that these variables have good properties in distinguishing the different cluster groups. Important economic variables, namely, the ‘development of employment’, the ‘turnover in 2002’ as well as the ‘development of turnover’, can also distinguish between the groups from the classification in seven clusters. For the classification of ten groups, the variables ‘development of employment’, ‘number of employees with university degree’ and ‘development of turnover’ show a significant impact.

5 Conclusions

Conventional industry classifications such as the NACE are not unconditionally adequate to classify all subsectors of the economy. Within the paper at hand, we examine the sector of Knowledge-Intensive Business Services (KIBS) in this respect. This relatively new economic sector is characterised, *inter alia*, by large numbers of firm start-ups, a high interactivity of firms with their environment and an outstanding pace of technological and economic development. While the output of the firms in the KIBS sector (*i.e.*, services and products) is quite heterogeneous – even on the level of a single firm – the commonalities of the firms are high degrees of innovative behaviour and interactivity, as well as an outstanding relevance of spatial proximity to clients, suppliers, and partners. As conventional industry classifications are mainly output-oriented and classify firms using their products, difficulties occur, at least in the KIBS sector.

We address this problem by proposing an alternative way of classifying the firms in the KIBS sector. The basis therefore is a cluster analysis of the KIBS Foundation Survey, a data set of nearly 550 German KIBS firms. Instead of falling back on the output of the firms, we use the rather collective and persistent characteristics of the firms, *i.e.*, their innovativeness and interaction behaviour, as well as the role of spatial proximity as the basis of our classification. Thereby, a structure which groups the firms of our data set into clusters with different characteristics is discovered. The validity of this structure of classification can be confirmed with multinomial logit models, taking into account different important economic characteristics. The tests confirm that the classification result of our analysis is a possible alternative to the conventional industry classification. Important firm characteristics such as the development of the employment or the development of the turnover have a strong correlation with our cluster results and, thus, are adequate to distinguish between the different clusters.

With the dissociation of the clustering process from the supply structure of KIBS and the consideration of the three outlined dimensions, the new classification is not so strongly dependent on actual changes as the traditional classification. Additionally, since KIBS can be definitely allocated, problems of overlapping can be solved with the new cluster structure. With the new classification the possibility arises to forecast the development of economic variables in the first years after the KIBS founding process. The firms could get assigned to a specific cluster to forecast tendencies in the development of employment or the development of turnover.

The present analysis can be seen as a first step in developing alternative, more adequate ways to classify new economic sectors, particularly within the service industries. It has to be noted that the new classification should not replace the traditional classification. The SIC follows an international standard and is important *e.g.*, to compare structures of industries between different countries or regions. Notwithstanding the mentioned problems, the traditional classification provides important information. The developed cluster structures from this analysis should be seen as an additional possibility for getting further information with respect to specific questions in the high innovative

KIBS sector. This new knowledge is also of political relevance, *e.g.*, when considering measures for the assistance of the service sector, the KIBS sector in particular or newly founded firms in these areas. To reach the desired effect of special policy interventions, the knowledge of the dissimilarity of newly founded KIBS with respect to different employment or turnover patterns, is of high political importance. It is helpful to know that firms in the KIBS sector are rather differentiated by their internal characteristics such as innovative or cooperative behaviour than by their outputs in order to design specific support actions.

For future studies and for future attempts of reforming industry classifications, the specific structure of sectors should be better considered. Thus, it can be contemplated to base the initial assignment of a newly founded firm or firm unit on criteria other than the traditional output-oriented measures. In the KIBS sector, as well as in other service sectors and, for example, in intermediary sectors such as the ICT industry, identifying more persistent and stable firm characteristics should be taken into consideration, as the boundaries between products and services get more and more blurred.

Acknowledgements

We owe our thanks to the German Research Foundation (DFG) for funding the project *The Foundation of Knowledge-intensive Business Service Firms in the Context of Industrial Core Regions* (grant No. RO 534/6), in which data have been collected and the fundamentals of the present paper have been laid. The Chambers of Industry and Commerce (IHK) in the regions of Bremen, Munich and Stuttgart are gratefully acknowledged for providing the firms' addresses for our survey. Participants of the CREPS Conference in Lüneburg, Germany, November 2005, where an earlier version of the paper has been presented, are gratefully acknowledged. We are indebted to Gerd Ronning and Harald Strotmann for their comments and support. The helpful suggestions of an anonymous referee are gratefully acknowledged. The remaining errors and all opinions expressed in the paper are the sole responsibility of the authors.

References

Aldenderfer, M.S. and Blashfield, R.K. (1984) *Cluster Analysis*, Beverly Hills: Sage.

Anderberg, M.R. (1973) *Cluster Analysis for Applications*, New York: Academic Press.

Atzema, O. (2001) 'Location and local networks of ICT firms in the Netherlands', *Tijdschrift voor Economische en Sociale Geografie*, Vol. 92, No. 3, pp.369–378.

Bacher, J. (1996) *Clusteranalyse: Anwendungsorientierte Einführung*, 2nd ed., München, Wien: Oldenbourg.

Bailey, K.D. (1994) 'Typologies and taxonomies: an introduction to classification techniques', *Sage University Paper Series on Quantitative Applications in the Social Sciences*, Thousand Oaks, CA: Sage, pp.7–102.

Bergs, S. (1981) 'Optimalität bei Clusteranalysen. Experimente zur Bewertung numerischer Klassifikationsverfahren', Dissertation, Münster: Westfälische Wilhelms-Universität.

Brixy, U. and Grotz, R. (2006) 'Regional patterns and determinants of new firm formation and survival in Western Germany', IAB discussion papers, Nürnberg: IAB, No. 5.

Cainelli, G., Evangelista, R. and Savona, M. (2006) 'Innovation and economic performance in services: a firm-level analysis', *Cambridge Journal of Economics*, Vol. 30, No. 3, pp.435–458.

Czarnitzki, D. and Spielkamp, A. (2003) 'Business services in Germany: bridges for innovation', *The Service Industries Journal*, Vol. 23, pp.1–30.

European Commission (2007) 'List of NACE codes', http://ec.europa.eu/comm/competition/mergers/cases/index/nace_all.html (download 12 June 2007).

European Monitoring Centre on Change (EMCC) (2005) 'Sector futures. The knowledge intensive business services sector', Dublin: European Foundation for the Improvement of Living and Working Conditions, http://www.eurofound.europa.eu/emcc/content/source/eu05016a.html (download 20 May 2007).

Everitt, B.S., Landau, S. and Leese, M. (2001) *Cluster Analysis*, 4th ed., New York: Oxford University Press.

Federal Statistical Office of Germany (2007) 'Data on employment and labour market', http://www.destatis.de/themen/e/thm_erwerbs.htm (download 25 May 2007).

Freel, M. (2006) 'Patterns of technological innovation in knowledge-intensive business services', *Industry and Innovation*, Vol. 13, No. 3, pp.335–358.

Gallouj, F. and Weinstein, O. (1997) 'Innovation in services', *Research Policy*, Vol. 26, pp.537–556.

Gordon, A.D. (1981) *Classification: Methods for the Exploratory Analysis of Multivariate Data*, London, New York: Chapman and Hall.

Howells, J.R.L. (2002) 'Tacit knowledge, innovation and economic geography', *Urban Studies*, Vol. 39, Nos. 5–6, pp.871–884.

Illeris, S. (1994) 'Proximity between service producers and service users', *Tijdschrift voor Economische en Sociale Geografie*, Vol. 85, No. 4, pp.294–302.

Johannisson, B. (1998) 'Personal networks in emerging knowledge-based firms: spatial and functional patterns', *Entrepreneurship & Regional Development*, Vol. 10, No. 4, pp.297–312.

Kaufman, L. and Rousseeuw, P.J. (1990) *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley & Sons.

Kaufman, R.L. (1985) 'Issues in multivariate cluster analysis. Some simulations results', *Sociological Methods & Research*, Vol. 13, pp.467–486.

Koch, A. and Stahlecker, T. (2006) 'Regional innovation systems and the foundation of knowledge intensive business services', *European Planning Studies*, Vol. 14, No. 2, pp.123–146.

Koch, A. and Strotmann, H. (2006) 'Impact of functional integration and spatial proximity on the post-entry performance of knowledge intensive business service firms', *International Small Business Journal*, Vol. 24, No. 6, pp.610–634.

Koch, A. and Strotmann, H. (2008) 'The determinants of innovation in newly founded knowledge intensive services', *Economics of Innovation and New Technology,* forthcoming.

Kronthaler, F. (2005) 'Economic capability of East German regions: results of a cluster analysis', *Regional Studies*, Vol. 39, No. 6, pp.739–750.

Lundvall, B-Å. (1988) 'Innovation as an interactive process: from user-producer interaction to the national system of innovation', in G. Dosi, C. Freeman, R. Nelson, G. Silverberg and L. Soete (Eds.) *Technical Change and Economic Theory*, London, NewYork: Pinter.

Martinez-Fernandez, M.C. (2006) 'Introduction to special issue of "International Journal of Services Technology and Management" on Knowledge Intensive Service Activities (KISA)', *International Journal of Services Technology and Management*, Vol. 7, No. 2, pp.109–114.

Miles, I., Kastrinos, N., Flanagan, K., Bilderbeek, R., den Hertog, P., Huntink, W. and Bouman, M. (1995) *Knowledge-Intensive Business Services. Users, Carriers and Sources of Innovation*, Brussels: European Commission.

Muller, E. and Zenker, A. (2001) 'Business services as actors of knowledge transformation: the role of KIBS in regional and national innovation systems', *Research Policy*, Vol. 30, pp.1501–1516.

Nählinder, J. and Hommen, L. (2002) 'Employment and innovation in services: knowledge intensive business services in Sweden', Report prepared for the final meeting of AITEG.

OECD (2004) *OECD Information Technology Outlook 2004*, Paris: OECD.

Peneder, M. (2005) 'Creating industry classifications by statistical cluster analysis', *Estudios de Economía Aplicada*, Vol. 23, No. 2, pp.451–463.

Roberts, J. (1998) *Multinational Business Service Firms. The Development of Multinational Organisational Structures in the UK Business Service Sector*, Aldershot: Ashgate.

Sneath, P.H.A. (1957) 'The application of computers to taxonomy', *Journal of General Microbiology*, Vol. 17, pp.201–226.

Strambach, S. (2002) 'Change in the innovation process: new knowledge production and competitive cities – the case of Stuttgart', *European Planning Studies*, Vol. 10, No. 2, pp.214–231.

Tether, B.S. (2003) 'The sources and aims of innovation in services: variety between and within sectors', *Economics of Innovation and New Technologies*, Vol. 12, No. 6, pp.481–505.

Toivonen, M. (2004) 'Expertise as business: long-term development and future prospects of knowledge-intensive business services (KIBS)', Doctoral Dissertation Series 2004/2, Helsinki University of Technology, Department of Industrial Engineering and Management, Espoo.

United Nations (2002) 'International Standard Industrial Classification of all economic activities (ISIC), Revision 3.1', *Statistical Papers Series M No. 4, Rev. 3.1*, New York: United Nations.

Ward, J.H., Jr. (1963) 'Hierarchical grouping to optimize an objective function', *Journal of the American Statistical Association*, Vol. 58, pp.236–244.

Wiig Aslesen, H. and Jakobsen, S-E. (2007) 'The role of proximity and knowledge interaction between head offices and KIBS', *Tijdschrift voor Economische en Sociale Geografie*, Vol. 98, No. 2, pp.188–210.

Notes

1 For recent examples of the use of cluster analysis in economic research see, for example, Kronthaler (2005) or Peneder (2005).

2 The knowledge intensity is commonly defined either by input factors such as the qualification structure of the employees (high proportion of skilled labour) or the R&D expenditures, or by output factors such as innovations or patents (see for example, Miles *et al*., 1995; Toivonen, 2004).

3 For a more detailed discussion of the definition and the measurement of innovative activity in the service sector see Cainelli *et al*. (2006), Gallouj and Weinstein (1997), Koch and Strotmann (2008) or Tether (2003).

4 Indeed, at least in Germany, most firms and firm foundations in the KIBS sector concentrate in the major urban agglomerations (Brixy and Grotz, 2006), where important potential clients are also located. However, the role of proximity may well vary not only from firm to firm, but also between different subsectors of the KIBS sector (Czarnitzki and Spielkamp, 2003).

5 However, it has to be noted that the role of spatial proximity for firms in the KIBS sector may decrease as firms develop, grow and standardise their products. It is of outstanding relevance in contexts with high degrees of interaction, tacit knowledge and innovation (see also Roberts, 1998).

6 These three German metropolitan regions were chosen owing to their comparability regarding political functions (all are federal state capitals) and owing to the distinctiveness of their industrial structures (for a detailed assessment, see Koch and Stahlecker, 2006).

7 Some 74 subsectors have been excluded. For example, the firms classified as 'Management Activities of Holding Companies' (74.15) – up to 40% of the regional samples – have not been considered as KIBS. Likewise, some subsectors of the advertising industries (*e.g.*, call centres) have been excluded.

8 We are aware of the fact that the restriction of our samples to young firms may be a limiting factor to our analysis as young firms may behave differently from incumbents. However, the KIBS sector is a young industry and it is thus characterised by young firms as well as above-average rates of entry and exit (Brixy and Grotz, 2006). Since we base our analysis on a dataset focusing on young firms, we are able to take into account this special aspect of the KIBS sector.

9 In the questionnaire of the KIBS Foundation Survey, a broad definition of R&D expenditures has been applied with respect to the particular situation in the service sector (*e.g.*, investments in human resources). R&D was broadly defined as "not order-bound investments in the development of products/services and/or qualification and training of staff". In some, particularly very young firms, these R&D expenditures exceed turnover, leading to R&D shares of more than 100%.

10 Certainly, some of the latter variables subsumed under the factor 'spatial proximity' could also serve as indicators for integration, as the two are highly correlated. For the sake of clarity, we decided to assign all variables with a spatial dimension to the category of spatial proximity.

11 Other possibilities mentioned in the literature are to exclude the single variables of a firm where missing values occur. A major problem is that this leads to different cases of diverse variables. Another possibility is the completion of missing values with estimated values (*e.g.*, mean of the correct answered variables). However, this is questionable in an analysis including binary variables (Bacher, 1996, p.231) and problematic because of the unequal dimension of the variables. Kaufman (1985) analyses different effects of various alternatives on the Ward algorithm.

12 For detailed information on the development of the heterogeneity measure during the last 20 steps of the Ward algorithm, see Appendix 1.

13 For detailed information about the different z-values of each cluster, see Appendix 2.

14 For details on the contingency analysis, see Appendix 3.

15 These categories have been created by aggregation of the 3-digit subsectors of the NACE (see Table 1). 'Software' is equal to sector 722, 'Other Electronic Data Processing' comprises the sectors 721 and 723–726, 'Technical Services' includes 731 and 742, whereas 'Consultancy' contains the sectors 732 and 741. 'Advertising' corresponds to sector 744.

Appendix 1

Since the increase of the heterogeneity is very high from seven to six groups, it is justified to stop the classification procedure at seven clusters (C7). Additional to the seven-cluster solution, a result with three clusters (C3) or ten clusters (C10) seems to be adequate. However, since the Ward algorithm is a hierarchical cluster algorithm, the different results can be interpreted as different aggregation levels of the one classification.

Development of the heterogeneity (Ward algorithm)

Number of clusters	*Heterogeneity measure*	*Development of heterogeneity*	*Growth of the development (%)*
20	1779.280	36.617	1.96
19	1817.104	37.825	3.30
18	1856.995	39.890	5.46
17	1906.774	49.749	24.72
16	1958.782	52.038	4.60
15	2011.667	52.886	1.63
14	2065.229	53.562	1.28
13	2127.099	61.870	15.51
12	2195.165	68.066	10.01
11	2264.941	69.776	2.51
10	2337.530	72.589	4.03
9	2421.404	83.875	15.55
8	2520.596	99.192	18.26
7	2631.319	110.723	11.63
6	2773.019	141.700	27.98
5	2961.967	188.948	33.34
4	3165.770	203.803	7.86
3	3370.971	205.201	0.69
2	3709.485	338.515	64.97
1	4177.150	467.665	38.15

Appendix 2

To statistically interpret the different clusters, we calculate the arithmetic mean of each cluster with respect to the selected variables and compare these values with the *z*-values, using:

$$z_{gj} = \frac{\overline{x}_{gj} - \overline{x}_j}{\overline{s}_{gj}}$$

with z as the *z*-value of cluster g in variable j, calculating the deviation of the cluster (represented by the mean value of the cluster g in variable j $\overline{x}_{gj}$) from the total mean of each variable $\overline{x}_j$, related by the standard deviation $\overline{s}_{gj}$. Thus, the *z*-values indicate if one cluster is above or below the mean of one specific variable.

Cluster mean values (and cluster *z*-values in parentheses)

Cluster	*C1*	*C2*	*C3*	*C4*	*C5*	*C6*	*C7*	*Total*
Size (number of firms)	114	65	69	47	54	2	37	388
Innovation								
No own innovations	0.982 (0.610)	1 (+)	1 (+)	1 (+)	1 (+)	1 (+)	0.027 (–5.323)	0.902
Improvement of existing own services	0.535 (–0.167)	0.969 (2.015)	0.478 (–0.279)	0.851 (0.646)	0.778 (0.379)	0.5 (–0.168)	0 (–)	0.619
Development of totally new services	0.982 (1.724)	1 (+)	0.928 (0.660)	0 (–)	0.926 (0.646)	1 (+)	0 (–)	0.755
R&D-expenditures (share of turnover in %)	20.581 (0.056)	11.962 (–0.676)	29.841 (0.321)	8.043 (–1.190)	17.565 (–0.097)	330 (3.139)	8.027 (–0.622)	19.243
Functional integration								
Professional background of the founder	2.167 (–0.044)	2.338 (0.255)	1.942 (–0.438)	2.149 (–0.083)	2.185 (–0.025)	2 (–)	2.649 (0.925)	2.201
Utilisation of results within the firm founding process	0.877 (–0.467)	2.923 (5.587)	0.681 (–0.656)	1.574 (0.108)	1.611 (0.135)	1 (–0.299)	1.405 (–0.012)	1.423
Spatial proximity								
Regional background of the founder	3.956 (1.525)	3.938 (1.470)	3.913 (0.997)	3.851 (0.646)	1.907 (–2.999)	1.5 (–2.945)	3.405 (–0.175)	3.582
Share of regional turnover	63.623 (0.522)	44.738 (–0.013)	15.087 (–1.353)	59.043 (0.341)	25.704 (–0.593)	30 (–0.359)	57.378 (0.304)	45.227
Purchase of knowledge intensive services	1.175 (0.236)	1.169 (0.234)	0.696 (–0.324)	1.805 (0.131)	0.593 (–0.520)	0.5 (–0.656)	0.865 (–0.108)	0.964
Number of regional partners	1.675 (0.304)	1.754 (0.340)	0.667 (–0.768)	1.255 (–0.009)	0.685 (–0.538)	1 (–0.188)	1.135 (–0.133)	1.265
Importance of a regional lead customer	1.167 (0.369)	0.985 (0.148)	0.362 (–0.834)	0.894 (0.057)	0.500 (–0.422)	0.5 (–0.488)	0.973 (0.167)	0.845

Note: In the cases where the *z*-value is indicated with + (–), the exact value could not be calculated owing to a standard deviation or the means of ‘0’. Thus, they indicate a very high (low) value.

Appendix 3

NACE classification	*Classification in 7 clusters*						*Total*
	1	*2*	*3*	*4*	*5*	*7*	
Software	24	13	16	7	8	4	72
	21.3	12.1	12.9	8.8	10.1	6.9	72
Other electronic data processing	21	12	15	12	9	12	81
	23.9	13.6	14.5	9.9	11.3	7.8	81
Technical services	24	15	17	13	12	12	93
	27.5	15.7	16.6	11.3	13.0	8.9	93
Consulting	16	14	9	12	21	7	79
	23.3	13.3	14.1	9.6	11.1	7.6	79
Advertising	29	11	12	3	4	2	61
	18.0	10.3	10.9	7.4	8.5	5.8	61
Total	114	65	69	47	54	37	386
	114	65	69	47	54	37	386

Notes: Pearson $X^2(20) = 37.097$, $Pr = 0.011$, Cramer's $V = 0.155$.