<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *International Journal of Social Research Methodology*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Waxell, A. (2009)

To capture an emerging industry: using industrial standards to identify the biotech industry.

International Journal of Social Research Methodology, 12(5): 441-456

<http://dx.doi.org/10.1080/13645570802573465>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Link to the publishers version: <http://www.tandfonline.com/doi/abs/10.1080/13645570802573465>

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-111227>

To cite:

Waxell, Anders (2009). To capture an emerging industry: using industrial standards to identify the biotech industry, *International Journal of Social Research Methodology*, 12(5), pp. 441-456.

[Page numbers in original printed version cited in brackets throughout the paper]

To Capture an Emerging Industry: Using Industrial Standards to Identify the Biotech Industry

ANDERS WAXELL

Centre for Research on Innovation and Industrial Dynamics (CIND),
Department of Social and Economic Geography, Uppsala University,
P.O. Box 513,
SE-751 20 Uppsala, Sweden

ABSTRACT

Standard industrial classification systems are the cornerstone of quantitative research in several fields. Data collected and coded on the basis of such systems often provide the most complete and accessible material with which to investigate, quantify and map the industrial landscape, and understand the economy at large. However, as relatively static classifications, they present researchers attempting to capture new, emerging and growth industries with a number of problems. Our understanding of the dynamics of economic development is largely based on aggregated firm data, often categorised and identified using an intertwined system of national and international industrial classification standards. This paper uses the example of the biotech industry in testing the ongoing appropriateness of one such industrial classification system, namely the Swedish industrial classification system.

Keywords: industrial standards; industry studies; biotech industry; Sweden

INTRODUCTION

Economic activities are often described in terms of industries, and using industry analyses in building an understanding of how the economy is structured is nothing new. Several standards for classifying industries, such as the United Nations' ISIC standard and the European Union's NACE standard, have been formulated and revised in recent decades in order to collect economic data. However, given their relatively static nature, these standards are accompanied by some problems that are especially apparent to researchers attempting to capture and understand new, emerging and growth industries, and to policy-makers aiming to devise measures to boost the economy.

Here we primarily deal with the critical methodological issue of how to identify dynamic economic activities and classify them into static entities. As processes of industrial transformation and growth unfold over time and space, firms and industries emerge in the intersection of new and old ones. However, the ways we observe, sort and classify these firms or industries do not change at the same pace, making existing industry nomenclatures and standards a blunt tool with which to understand such dynamics. To understand and analyse the underlying mechanisms of industrial transformation and growth there is obviously a need to systemise information and data about economic activities. This includes a need to identify actors and categorise them [end of p. 441] in an orderly fashion. The best alternative is not always to use existing standards or nomenclatures, but to use new alternative methods or at least a combination of new methods and traditional standards. Nonetheless, information about, for example, change processes, growth rates and quantities still needs to be systematised. This leaves the main question unanswered: How are we supposed to analyse or understand industrial transformation and growth when we are unable to capture those actors that are part of new and emerging industries?

The emerging biotech industry is in this sense often cited as a highly prioritised economic growth area. This may also be exemplified by the numerous cluster initiatives focusing on biotech-related areas set up around the globe, not least in Sweden (Lundequist & Power, 2002; Sölvell, Lindqvist, & Ketels, 2003). The key issue this paper deals with is the nature of the Swedish industrial classification system and its usefulness in capturing new areas of economic growth, such as the biotech industry.

This paper thus intends to discuss and critically assess the use of industrial classification standards and nomenclatures, especially in relation to their ability – or lack thereof – to capture new and emerging firms or types of

industries. The main purpose is to demonstrate the difficulties of tracking and identifying firms in biotech-related activities using official industrial standards. Here, this will be exemplified using a population of firms in the Swedish biotech industry as defined and analysed by the Swedish Agency for Economic and Regional Growth (NUTEK). Furthermore, our aim is to make industry scholars aware of the main issues and problems related to official industrial data and how industrial activities are classified. Several problems and shortcomings related to the use of official industrial standards and nomenclatures will be highlighted. Furthermore, the paper will discuss other possible methods for identifying and capturing industrial actors.

From a broader perspective, the paper can be viewed as a contribution to the growing literature challenging the traditional view of industries and industry affiliations. In this literature, the economy is more or less understood as related or interlinked systems delimited by either technological or geographical denominators. The trend in contemporary industry studies in general is to use concepts linked to various industrial systems to analyse and understand processes of innovation and economic growth. Doing so involves concepts such as clusters, national or regional innovation systems, technological systems and networks (cf. Carlsson, 1997; Cooke, Uranga, & Etzebarria, 1997; Lundvall, 1992; Porter, 1990). In Sweden, as in many other countries, this approach has received particular attention from various official policy organs. However, these concepts, especially the cluster concept, are widely criticised. First, the concepts are viewed as somewhat ‘fuzzy’, due to both inconsistent definitions and a lack of common methodological understanding. Second, the criticism focuses on the question of how policy-makers and practitioners use the concepts, and whether they should be used in this context at all (for a more thorough discussion, see, e.g. Lundequist & Power, 2002; Malmberg, 2002; Markusen, 1999; Martin & Sunley, 2003). For the sake of the argument expressed in the present paper, it should be noted that the criticism is in many ways legitimate. More importantly, however, what is seldom discussed is that more traditional industry studies, using current standards and nomenclatures, are hardly better alternatives, at least not in terms of studying new and emerging industries.

The structure of the paper is as follows. First, the difficult balance between industrial dynamics and the static standards that industrial statistics often involve will be discussed, not least in terms of international comparisons. Then the Swedish industrial [end of p. 442] classification system – the SNI standard – and its ability to capture the evolving biotechnology sector will be used to illustrate the problems we encounter when attempting to capture

emerging industries. The paper concludes by arguing that we need careful thought about adjustments and alternatives to contemporary classification systems. Nonetheless, implicit throughout the paper is an acknowledgement of the importance of such classification systems to the ability of both academia and policy to understand the factors and mechanisms underlying economic growth.

NOMENCLATURE IN OFFICIAL INDUSTRIAL STATISTICS: PROBLEMS AND PRACTICAL EXAMPLES

Studying industries and their characteristics obviously calls for industry statistics. According to a 1979 document from the Swedish Ministry of Economy, industrial statistics should be used to convey useful information about conditions concerning production and resources related to various manufacturing activities, and about the structure of production and its spatial distribution within the nation (Ds E, 1979:2). In practice, there is a will and a need to measure and classify the production and supply of goods and services, including the measurement of costs involved in making a product (i.e. inputs, energy, labour and fixed capital). These data are then used as input in analysing economic transformation and growth. However, the economy and the industrial landscape comprise activities that are both similar/related and diverse. This is evident not only if we analyse the activities performed, but also if we look at the range of products actually produced. The need to systemise and categorise activities and products that are closely related or akin to each other into industries and sectors is thus prominent. The division of industry into sectors is in other words an instrument used to clarify and make penetrating analyses of the economy of a territorially delimited spatial area, such as a nation state, region or city.

Four general problems with industry studies

On a more general level, there are four main problems related to industry statistics. These are more or less built into the statistics used, and are thus part and parcel of many industry studies.

First, information about economic activities can be gathered in numerous ways. Usually, the information or data capture conditions in the firm or workplace. This means that it is not always evident what is, or what ought to be, the true or actual information-bearing unit. Information is normally gathered and sorted according to different types of production units, which in

real life could be workplaces, the firm as a legal entity, authorities and/or similar units. These production units are functionally and geographically delimited and are the smallest, physically independent units where the production of goods or services takes place (cf. Maskell, 1992). The biggest challenge for the researcher in this respect is to be aware of and acknowledge the differences between a workplace and a firm, and to choose the appropriate unit for analysis. By 'firm', we often mean the legal entity encompassing the economic activity. In many business registers, whether public or private, economic data usually refer to the 'firm' as a legal entity or as the firm's main office. On the other hand, a firm is always made up of at least one workplace, though it could also consist of several workplaces that might even be spatially separated. By 'workplace' we mean 'every address, building or group of buildings where a firm performs an activity' (Statistics Sweden, 2003b, translated from Swedish). [end of p. 443]

The distinction between firm and workplace can be important, depending on the type of study to be performed. If the study aims to analyse the spatial division of labour and its frequency, the workplace is the correct unit for analysis and not the firm (cf. Fritz, Alvstam, & Korhonen, 1995). The result of the analysis may in other words be skewed if the wrong unit for analysis is chosen. For example, a firm may be legally domiciled in a certain location, although most of its goods or services are produced in workplaces situated elsewhere; this fact would have implications for all the data gathered about the firm.¹ The differences between workplace and firm should not least be apparent to researchers doing statistical analyses based on location quotients, regional employment and establishment counts, regional impact multipliers and regional growth relative to national growth (cf. McCann, 2001).

Second, it is not easy to classify economic activities or, especially, to define clear-cut boundaries between industries (cf. Lundmark & Malmberg, 1988). The problem here is how to move from the production unit to an aggregated industry level: on what attributes is the aggregation based? Normally, the production unit is classified based on the products of the unit, which, according to the Ministry of Economy, actually means the 'goods' produced. Every good is thus assigned to an industry (Ds E, 1979:2); in other words, the

¹ For example, the former Pharmacia Corporation (today Pfizer) was legally domiciled in Sweden in Stockholm and had approximately 3400 employees in 2002. If we instead chose the workplace as unit for analysis, it would become evident that approximately 1750 of the total number of employees worked in Stockholm while approximately 1400 worked in the city of Uppsala and the remainder elsewhere in Sweden. [on p. 453]

starting point has been to group and classify goods. As the economy has progressed over time, however, the production of services has increased, which in turn has made the classification process even more complex. There is thus a need to distinguish between the production of goods and services; meanwhile, the line dividing the two has become increasingly uncertain and in some cases has even faded away (cf. Alvstam, 1998; Alvstam & Erlandsson, 1995; Fritz et al., 1995).

The division of production units into industry sectors in Sweden and Europe is thus based on a product-oriented way of defining economic activities. That is to say, how a production unit is sorted and classified into an industry is dependent on the end product produced. This method has not gone uncontested. Over the past decade, several scholars have instead recommended that it would be more illuminating to focus on similarities in processes and activities related to the production process rather than the end product itself; for example, this view has been applied prominently in discussing and analysing the biotech industry (cf. Eliasson & Eliasson, 1997; Liebeskind, Oliver, Zucker, & Brewer, 1996). Since 1997, this perspective has also been incorporated into the North American standard (see subsequent section). Theoretically, there is a range of ways to define and order economic activities into industries, such as focusing on products, processes, inputs, technology platforms or other innovative alternatives.

The third problem is related to the constant tension between dynamic and static conditions or objects. It is problematic to capture and study processes of change over time using more or less static industrial standards. Industrial transformation and growth are parts of broader processes that not only lead to economic development (via innovation and improved competitiveness, for example), but also cause new firms and industries to emerge and others to fade away and disappear. The biotech industry is a good illustration of how existing industrial standards cannot capture such activity, as will be highlighted in the paper. It is only natural that industries, industry sectors and their definitions should change and be revised over time, and that this should pose a largely intractable analytical problem. On the one hand, one can choose to make the industrial classification structure or standard more or less permanent, making comparisons over time more readily available. On the other hand, one can choose to revise [end of p. 444] the classification structure and standards according to the changes occurring in the industry. The first scenario will be more appealing in contexts in which studies and comparisons over time are needed. The latter scenario will make it harder to perform

longitudinal studies, while it would be more likely to recognise and capture industrial transformation and related processes of change.

The fourth problem involves the difficulties of making comparative studies between different countries where specific national standards apply. Distinctive national features are often built into nomenclatures and standards originally based on international standards. This is often due to ambiguity resulting from a national notion that international standards do not adequately capture nation-specific industrial structures and conditions, while there is a concurrent need for benchmark studies comparing nations. One way to make comparative studies between nations is to use the NACE standard on a four-digit level. In cases in which no international standards are available, a comparative approach is of course even more difficult.

All four problems discussed above are more or less embedded in contemporary industry statistics and data. This does not mean that data gathered about firms and their industrial affiliations are not to be taken seriously, but merely that researchers and practitioners should at least be aware of the problems and take account of them.

The Swedish Standard Industrial Classification (SNI) and its international counterparts

Statistics Sweden (SCB) is the central government administrative authority responsible for gathering, co-ordinating and maintaining official statistics in Sweden. Since 1963, SCB has, via its Business Register, gathered data on every firm and workplace in both the public and private sectors in the nation (Statistics Sweden, 2002).

Given the need to categorise and classify economic activities, SCB has created the Swedish Standard Industrial Classification (SNI), in which the economy is divided into industry sectors or segments. The standard provides every economic activity (both the firm as legal entity and workplaces) with a five-digit classification code that is supposed to reflect the main activity or activities undertaken. The standard has been revised several times. The first version was established in 1969 (SNI 69) and was based on the United Nations' International Standard Industrial Classification of All Economic Activities (ISIC Rev. 2) of 1968 (Statistics Sweden, 2003b).

A revised version was established in 1992 (SNI 92), this time based on the industrial standard of the European Union, i.e. the Nomenclature Générale des Activités Economiques dans les Communautés Européennes (NACE Rev.

1), which came into force in 1990. SNI 92 consists of five hierarchical levels, the first four being identical to the NACE standard while the fifth is a Swedish supplement level. A new standard was later introduced, SNI 2002, based on NACE Rev. 1.1 (Statistics Sweden, 2002).

It should be noted that both the NACE and ISIC standards share similarities with the Standard Industrial Classification (SIC) established in the USA in 1937, which since 1997 has largely been replaced by the North American Industry Classification System (NAICS). A big difference between the SIC- and the NAICS-based classification is that the former classifies establishments that produce similar products while the latter groups establishments based on their production processes (Kelton, Pasquale, & Rebelein, 2008).

In early 2008, yet another revised version will be introduced in Sweden, SNI 2007, which is integrated with NACE Rev. 2; changes have been made so as better to [end of p. 445] incorporate activities in the information and communications technology sector (Statistics Sweden, 2007).

SNI classification in practice

From a critical perspective, it is important to recognise the processes by which economic activities become classified into SNI codes and the uncertainties that follow from this. While SCB has compiled the standard, the actual classification of economic activity is done at the local tax authority when a firm is registered. Newly established firms as well as reconstructed and taken-over firms must complete a form that allows entrepreneurs themselves to define their new firm's activity area. If the firm is a trading partnership (handelsbolag, HB), limited partnership (kommanditbolag, KB) or limited-liability company (aktiebolag, AB), the firm should also be registered at the Swedish Companies Registration Office.

The actual classification according to the SNI standard takes place at the local tax authority and is based on the information supplied in the form provided by the tax authority. The information in the form should briefly state the line of business: for example, manufacturing, trading (wholesale, retail or trading) and consultancy. The completed form should also include information about the kinds of goods or services to be produced: for example, pharmaceuticals manufacturing. If the firm is a service company, a short description of the business activity should be provided. In the next step, the local tax authority assigns the firm its correct SNI classification code based on the information provided in the form. In cases in which a firm performs more than one

activity, the level of each activity must be presented in percentage terms (Swedish Tax Agency, 2002a, 2002b, 2003). The activity level is normally based on turnover or amount of labour put into each business activity. The SNI code originally represents the legal unit of the firm. In cases in which a firm consists of more than one workplace, SCB sends out a complementary survey to each workplace in the firm (Statistics Sweden, 2003b). Of course, the above process does not always run flawlessly. For the entrepreneur starting up a new firm, the process of being assigned a correct SNI code is perhaps not the most important task on the agenda. Moreover, how the firm's activity is described may also be influenced by how the entrepreneur defines the activity (i.e. what we want to do is not always what we actually do), and the entrepreneur's self-image may not coincide with the tax authority's definition of the activity when assigning it a SNI classification. Finally, it may be hard in practice to define firms operating on the boundary between existing activity fields and industries.

METHODS AND MATERIALS

The present analysis and discussion are based on the Swedish Standard Industrial Classification 1992 standard (SNI 92),² according to which the economy is categorised at a detailed five-digit level, based on the economic activity performed in business establishments. The material used here is a pre-defined population of biotech firms that will serve as a control population. The population was first identified by NUTEK in a study of the Swedish biotechnological innovation system (Backlund, Markusson, Norgren, & Sandström, 2000).³ The study uses a number of different sources to identify existing biotech activities. These activities have moreover been analysed and classified using activity descriptions from the Internet, literature and [end of p. 446] personal contacts with experts in the fields or firms in question.⁴ Of course, it may be problematic to use pre-defined populations, given that we

² Also abbreviated as SE-SIC 92. The standard has been revised twice since the present material was gathered. Given the nature of the material, this study will use the SNI 92 standard. [on p. 454]

³ The study, part of a larger project, is now under the aegis of the Swedish Governmental Agency for Innovation Systems operations (Vinnova). [on p. 454]

⁴ The study is part of a larger project that has been presented in English in the report, *The Swedish biotechnology innovation system, VINNOVA Innovation in Focus VF 2001:2*; the report provides a more thorough discussion of material than the earlier Swedish version (see Sandström et al., 2001, pp. 5–6). Some data have been updated in the final English report. This paper uses the population presented in the Swedish version, which lists firms active in 1998. [on p. 454]

do not know of possible shortcomings and limitations associated with the identification of the material. However, the method used here provides a comprehensive picture of an industry that otherwise is statistically hard to identify and capture.

The population is complemented by industry classification data (SNI 92) and activity descriptions (according to each firm's articles of association). These data were manually collected from the Företagsfakta public database provided by AffärsData in 2001. The total population in the NUTEK study comprises 193 firms. These firms represent the larger biotechnology innovation system, so the population incorporates firms that are not actual biotech firms. In the total population, approximately 137 firms are biotech oriented, according to the definition of biotech provided by NUTEK, which, briefly stated, includes 'those actors that develop, produce, analyse, or use biological systems on a micro, cellular, or molecular level' (Backlund et al., 2000, p. 6, translated from Swedish). Here the population analysed consists of 135 firms, as 2 firms became missing from the sample due to bankruptcy or mergers during the time of the analysis.

Moreover, of these 135 firms, 3 lacked industrial classification codes, while 31 had more than one code. In the latter case, this occurred because the first classification code represented the more dominant activity of the firm, which is why the first classification code is used here for the analysis. According to the SNI standard, the dominant activity determines how the firm should be classified (see Statistics Sweden, 2003a).

EMERGING INDUSTRIES, INDUSTRIAL TRANSFORMATION AND INDUSTRIAL STANDARDS: THE CASE OF THE BIOTECH INDUSTRY

Industrial transformation can in short be described as a struggle between old and new that will lead to processes of change, which in turn will stimulate economic development (cf. Dahmén, 1980). The basis of these changes is most often related to innovations/innovativeness and/or organisational changes within the firm. Products and methods of production are constantly developed and refined in interplay between firms and between their pool of labour and their customers. The contemporary biotech industry is in its very essence a result of industrial transformation, although the history of using biotech processes in industrial contexts is long and not just a recent phenomenon. Broadly speaking, mankind has been relying on biotech

processes for over 6000 years, for example, when it comes to cross-fertilisation and the use of fermentation processes related to making bread, beer, wine, cheese and yoghurt. Fermentation processes are industrially still a very lucrative part of the biotech sector, especially in the food sector but also in the chemical and pharmaceutical sectors, where the development and manufacturing of, for example, acetone, glycerol and penicillin are important activities (Waxell, 2005; Zucker & Darby, 1997). Although the history of many biotech processes extends far back in time, what many researchers understand as the contemporary biotech industry comprises activities that are essentially different from fermentation, for example. It is therefore common to distinguish between classic and modern biotech activities.

Modern biotech and its industrial spin-offs have their background in a number of groundbreaking scientific discoveries of the first half of the 1970s. The largest breakthroughs were the discovery and development of recombinant DNA technology (rDNA, sometimes referred to as 'gene technology') and the manufacturing of [end of p. 447] monoclonal antibodies (mAb technology). These discoveries were also path breaking from an industrial standpoint, as they aroused industrial and commercial interest around the globe. In a sense, these discoveries were to lay the groundwork for and shape the modern biotech industry (for a more thorough description, see Zucker, Darby, & Armstrong, 1998). Some studies even call the introduction of these technologies 'the biotech revolution'. Although 'revolution' is a strong word, we can conclude that biotech activities have become an important input into the economy, at least in some locations (cf. McCutchen & Swamidass, 1996; Waluszewski, 2003; Waxell & Malmberg, 2007; Zucker & Darby, 1997). The modern biotech industry has thus developed through processes in which industrial transformation is the focus, and in which processes of innovation can be viewed as stimulating factors. This also means that dynamic development may often be hard to describe using existing standards and nomenclatures. New and emerging industries are per definition hard to capture solely using existing statistical boundaries and delimitations.

Biotech industry according to the SNI nomenclature

Since biotech has no official coding in the SNI standard, biotech firms are hard to identify. In other words, industrial applications of biotech activities are part of several previously defined industries. These activities are usually found in industries in the healthcare, agricultural and food, environment and energy, and machinery and instruments sectors, and in a number of process

industries, i.e. the mining, chemical, and pulp and paper industries (see Audretsch, 2001; Eliasson & Eliasson, 1997; OECD Observer, 1999; SIND, 1991).

According to the SNI 92 standard, the production of goods and services is divided into a number of main sections, A to Q. The division is based on an aggregation of the SNI standard on a two-digit level in so-called subsections and main groups (see Statistics Sweden, 1992b). The firms that NUTEK has identified as biotech firms, thus forming part of the control population, are according to this division spread throughout the different subsections (see Table 1). Overall, it could be stated that the biotech sector actually cuts across the economy in a variety of industries.

Most biotech firms are found in section K (i.e. real estate, renting and business activities). This section contains the main group Research and development (SNI code 73), which is a common activity of many biotech firms.

The main factor underlying the uneven distribution of biotech firms is that these firms are classified based on the products they produce, i.e. the product-oriented approach to understanding economic activities. This implies that a firm that uses biotech processes to manufacture a food product, for example, a functional food product such as bioactive yoghurt, would be classified as belonging to section D and subsection DA (manufacture of food products, beverages and tobacco), even though the firms in question may have more in common with other biotech firms than with a firm that manufactures hot dogs, for example. In other words, it is the end product that determines how the firm is classified.

Biotech industry according to the two-digit SNI level

As already mentioned, the classification into sections is based on a grouping of similar activities on a two-digit level. It is therefore interesting to examine this level a bit closer. [end of p. 448]

Table 1. The Swedish industry structure according to Swedish Standard Industrial Classification (SNI 92) and the number of biotech firms in each section.

Sector	Sections of the Swedish standard industrial classification (SNI 92)	SNI code (two-digit level)	No. of biotech firms
I	A: Agriculture, hunting and forestry	01-02	3
	B: Fishing	05	-
	C: Mining and quarrying	10-14	-
II	D: Manufacturing	15-37	44
III	E: Electricity, gas and water supply	40-41	-
	F: Construction	45	1
	G: Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods	50-52	11
	H: Hotels and restaurants	55	-
	I: Transport, storage and communication	60-64	-
	J: Financial intermediation	65-67	1
	K: Real estate, renting and business activities	70-74	71
	L: Public administration and defence; compulsory social security	75	-
	M: Education	80	-
	N: Health and social work	85	1
	O: Other community, social and personal service activities	90-93	-
	P: Private households with employed persons	95	-
	Q: Extra-territorial organisations and bodies	99	-
Total			132

Source: Statistics Sweden (1992a) and Backlund et al. (2000).

On the two-digit level, the SNI 92 standard is very rough. The reason why this level is often used in many studies is that it comprises so-called pre-named main groups. Biotechnology is not a given main group, and according to Figure 1, most biotech firms are found in the main groups Research and development (SNI code 73) and Manufacture of chemicals and chemical products (SNI code 24). The likelihood of a biotech firm being categorised into either one of these two main groups seems fairly high. However, judging by the names of the other main groups in Figure 1, it is not as obvious that these would also contain biotech firms. Of the 135 firms included in this study, 11 are found in the main group Wholesale trade and commission trade, except of motor vehicles and motorcycles (SNI code 51). However, on closer examination, 5 of these 11 firms have the full five-digit coding Wholesale of

pharmaceutical goods (SNI code 51460). The miscellaneous category in Figure 1 includes main groups that only contain one biotech firm.

The rough two-digit level provides poor insight into the biotech industry. In practice, this means that if one does not have access beforehand to a pre-selected population of biotech firms, it is hard, if not impossible, to trace and identify them using the SNI standard on a two-digit level. To illustrate this problem even further, we can use the coding from the main group Wholesale trade and commission trade, except of motor vehicles and motorcycles (SNI code 51) to identify firms in the AffärsData public database that match this two-digit level. This results in a total [end of p. 449]

Figure 1. Number of biotech firms identified by Nutek according to the two-digit level of the SNI standard (code in parentheses). Source: Backlund et al. (2000) and AffärsData (2001).

number of 29,042 limited-liability companies, 8135 trading partnerships and limited partnerships, and 2631 other enterprises active in Sweden. To find and identify 11 biotech firms among all these would be tedious and not very productive.

Biotech industry according to the five-digit SNI level

Although more finely tuned than the two-digit level, the five-digit level – the most detailed level – also proves to be poor in terms of identifying firms in the biotech industry. According to Table 2, the 135 biotech firms included in the NUTEK study are assigned as many as 39 different classification codes on the five-digit level. On the one hand, and in relation to the size of the population, the wide range of classification codes must be considered somewhat conspicuous. On the other hand, this result is not unexpected, since the level of detail is supposed to be high. However, if one looks at the titles of the categories it becomes quite clear that a lot of them are still quite general, such as that labelled by classification code 74849, Various other business activities (a residual group) or 01300, Growing of crops combined with farming of animals (mixed farming).

The most common classification codes, however, do present a quite good picture of the biotech industry in general. The largest number of biotech firms, 29 firms, is found in Research and development on medical sciences (SNI code 73103), which [end of p. 450]

Table 2. Number and percentage of biotech firms and respective SNI codes at the five-digit level.

SNI 92	Text	No.	(%)
73103	Research and development on medical sciences	29	(22)
24420	Manufacture of pharmaceutical preparations	12	(9)
73105	Interdisciplinary research and development, predominantly on natural sciences and engineering	11	(8)
74202	Construction and other engineering activities	9	(7)
24410	Manufacture of basic pharmaceutical products	7	(5)
24660	Manufacture of other chemical products n.e.c.	7	(5)
33101	Manufacture of medical and surgical equipment and orthopaedic appliances except artificial teeth, dentures, etc.	5	(4)
51460	Wholesale of pharmaceutical goods	5	(4)
73102	Research and development on engineering and technology	5	(4)
73101	Research and development on natural sciences	4	(3)
33200	Manufacture of instruments and appliances for measuring, checking, testing, navigating, and other purposes, except industrial process control equipment	3	(2)
74140	Business and management consultancy activities	3	(2)
24140	Manufacture of other organic basic chemicals	2	(2)
24200	Manufacture of pesticides and other agro-chemical products	2	(2)
73104	Research and development on agricultural sciences	2	(2)
74112	Advisory activities concerning patents and copyrights	2	(2)

74849	Various other business activities	2	(2)
15512	Other dairy production	1	(0.7)
24130	Manufacture of other inorganic basic chemicals	1	(0.7)
24150	Manufacture of fertilisers and nitrogen compounds	1	(0.7)
24520	Manufacture of perfumes and toilet preparations	1	(0.7)
25210	Manufacture of plastic plates, sheets, tubes and profiles	1	(0.7)
33102	Manufacture of artificial teeth, dentures, dental plates etc.	1	(0.7)
45240	Construction of water projects	1	(0.7)
51210	Wholesale of grain, seeds and animal feeds	1	(0.7)
51230	Wholesale of live animals	1	(0.7)
51380	Wholesale of other food including fish, crustaceans and molluscs	1	(0.7)
51550	Wholesale of chemical products	1	(0.7)
51651	Wholesale of measuring and precision instruments	1	(0.7)
51659	Wholesale of machinery for industry, trade and navigation n.e.c.	1	(0.7)
67120	Security broking and fund management	1	(0.7)
71340	Renting of other machinery and equipment n.e.c.	1	(0.7)
72202	Software supply	1	(0.7)
73202	Research and development on humanities	1	(0.7)
74702	Disinfecting and exterminating activities	1	(0.7)
85140	Other human health activities	1	(0.7)
01410	Agricultural service activities	1	(0.7)
01300	Growing of crops combined with farming of animals (mixed farming)	1	(0.7)
01233	Raising of piglets and swine for slaughter, mixed	1	(0.7)
(empty)	-	3	(2)
Total		135	(100)

corresponds to 22% of the population. The runner-up is Manufacture of pharmaceutical preparations (SNI code 24420) followed by Interdisciplinary research and development, predominantly on natural sciences and engineering (SNI code 73105).

One of the more serious problems with the SNI codes is that the margin of error related to the classification process is relatively high. This may be due to a faulty classification procedure at the local tax authority or because the activity of the firm is hard to classify, since it may perform multiple activities or produce several products of different character. According to Larsson (1998), as many as one of four firms allegedly part of the Swedish machinery industry was on closer examination wrongly classified. Here, in the control population, the coding Manufacture of perfumes and toilet preparations (SNI code 24520) raises some suspicion. In this case, the firm could be developing biochemical products or the like for sanitation use. Nevertheless, the category Research and development on humanities (SNI code 73202) calls for more explanation.

The same dilemma is apparent on the five-digit level as on the two-digit level when it comes to identifying biotech firms. A search of AffärsData identifies 671 limited-liability companies, 364 trading partnerships and limited partnerships, and 57 other enterprises active only in Research and development on medical sciences (SNI code 73103). If one only considers the limited-liability companies, the biotech firms in the control population only represent approximately 4.3% of the total population in Sweden with this SNI classification.

To generate a clear-cut group of biotech firms only using SNI codes proves impossible in practice. In other words, we face a problem of over-coverage, since a large number of firms surface that are not part of the target population (cf. Körner, Ek, & Berg, 1984). We may thus conclude that no classification codes in the 1992 standard would solely comprise or identify biotech firms. In summary, it is important that scholars involved with statistical analyses based on industry standards be aware of the above issues and problems, especially as related to the classification process, as they could well affect the results and conclusions of such research.

CONCLUDING DISCUSSION

This paper demonstrates that contemporary industry standards, such as the Swedish Standard Industrial Classification 1992, are lacking in ability to identify and capture firms and various business activities that make up new and emerging industry formations, such as the biotech industry. As a complement to standard classifications, however, alternative methods can be used to identify new and emerging industries. This would ultimately generate a clearer and better picture of the contemporary [end of p. 452] economic landscape. In this respect, a more unconventional method of identifying and

grouping similar economic activities is to use the activity description written in the firm's articles of association, which are submitted to the authorities when registering a firm (i.e. as a limited-liability company, limited partnership or trading partnership). The articles of association include a detailed description of the business activity/-ies of the firm and the industry segment in which it is active. The articles of association are also a legal document protecting the firm name, and the firm has a legal obligation to report any changes concerning firm name, activity focus, address or board members (Bolagsverket, 2007; PRV Bolag, 2002), which in a sense ensures the currency of the data. In Sweden, the descriptions contained in the articles of association are publicly available through various public databases. By searching these

databases for common keyword search terms, firms with a certain range of activities, such as those in the biotech industry, may be grouped and identified. This method, however, is not foolproof and is largely dependent on the design of the keyword search. A closer examination of the activity descriptions furthermore reveals that the activities are largely described in general terms and often according to a standardised formulation, which also affects the result in terms of generating relevant populations. In our case, this method could only find 42% of the control population. Unlike the search using the SNI standard, we now face an under-coverage problem, since a great number of pre-identified firms were missing in the population generated. The method is thus not perfect and would require more thorough analysis of the activity description in the articles of association as well as a more methodical and open-ended way to find the right search words. Although the method might be considered somewhat problematic, it should be considered better than just using the SNI standard in terms of finding and identifying biotech firms.

In conclusion, new and alternative methods are needed in order to study and better understand industrial transformation and growth, especially in cases in which new industries are emerging. Contemporary industrial standards are failing in their duties in this, leaving a window of opportunity open for new and alternative methods for identifying and categorising economic and industrial activities. In this respect, various industrial systems approaches, such as the cluster approach or regional innovation approach, may fill an important gap. In answer to criticisms of these approaches, as presented earlier, it should be noted that these methods and perspectives should be judged in relation to existing alternatives, i.e. methods relying on available industry standards and structures in official statistics.

Acknowledgements

This research has been financed by Centre for Research on Innovations and Industrial Dynamics (CIND) and Jon Wallanders och Tom Hedelius Stiftelse and Tore Browaldhs Stiftelse. I am grateful to Anders Malmberg and Claes-Göran Alvstam for their constructive and encouraging comments on earlier versions of the paper, and Dominic Power who provided valuable insights into the later versions. Usual disclaimers apply.

REFERENCES

- AffärsData. (2001). *Företagsfakta [Business facts]*. Retrieved October 15, 2001, from <http://www.ad.se/nyad/index.phtml>
- Alvstam, C.G. (1998). Lokaliseringsteori för produktion av tjänster: Ueber den Mangel eines Standorts der Dienstleistungen. *Svensk Geografisk Årsbok*, 74, 67–80.
- Alvstam, C.G., & Erlandsson, U. (1995). Kommersiella tjänster: Inledning [Commercial services: Introduction]. In C.G. Alvstam (Ed.), *Sveriges nationalatlas: Industri och service [National Atlas of Sweden: Industry and Service]* (pp. 120–121). Stockholm/Höganäs: Sveriges nationalatlas (SNA)/Bra Böcker.
- Audretsch, D.B. (2001). The role of small firms in US biotechnology clusters. *Small Business Economics*, 17, 3–15.
- Backlund, A., Markusson, N., Norgren, L., & Sandström, A. (2000). Det svenska biotekniska innovationssystemet: Drivkrafter och hinder för innovationer och tillväxt [The Swedish biotechnology innovation system: Drivers and barriers for innovation and growth] (No. Innovationspolitiska analyser), NUTEK (Närings- och teknikutvecklingsverket).
- Bolagsverket. (2007). (Swedish Companies Registration Office website). Retrieved November 21, 2007, from http://www.bolagsverket.se/in_english/
- Carlsson, B. (Ed.). (1997). *Technological systems and industrial dynamics*. Boston, MA/London: Kluwer.
- Cooke, P., Uranga, M.G., & Etxebarria, G. (1997). Regional innovation systems: Institutional and organisational dimensions. *Research Policy*, 26, 475–491.
- Dahmén, E. (1980). Hur studera industriell utveckling? In E. Dahmén & G. Eliasson (Eds.), *Industriell utveckling i Sverige: Teori och verklighet under ett sekel [Industrial development in Sweden: Theory and reality during a century]* (pp. 25–62). Uppsatser till ett IUI-symposium i anledning av Marcus Wallenbergs 80-årsdag i Saltsjöbaden, Grand Hotel, October 8–9. Stockholm: Industriens Utredningsinstitut/Almqvist & Wiksell International.
- Ds E. (1979:2). Industristatistiken – uppgiftskrav och användbarhet. [Industrial statistics – data requirements and usefulness] Delbetänkande avgivet av utredningen om statistiska centralbyråns uppgiftssamling. Stockholm: Ekonomidepartementet.
- Eliasson, G., & Eliasson, Å. (1997). The pharmaceutical and biotechnological competence bloc and the development of Losec. In B. Carlsson (Ed.), *Technological systems and industrial dynamics* (pp. 139–168). Boston, MA/London: Kluwer.
- Fritz, M., Alvstam, C.G., & Korhonen, J. (1995). Svenskt näringsliv i historiskt perspektiv. In C.G. Alvstam (Ed.), *Sveriges nationalatlas: Industri och service [Swedish industry in a historical perspective]* (pp. 10–23). Stockholm/Höganäs: Sveriges nationalatlas (SNA)/Bra Böcker.
- Kelton, C.M.L., Pasquale, M.K., & Rebelein, R.P. (2008). Using the North American Industry Classification System (NAICS) to identify national industry cluster templates for applied regional analysis. *Regional Studies*, 42(3), 305–321.
- Körner, S., Ek, L., & Berg, S. (1984). *Deskriptiv statistik [Descriptive statistics]*. Lund: Studentlitteratur.

- Larsson, S. (1998). Lokal förankring och global räckvidd: En studie av teknikutveckling i svensk maskinindustri [Local embeddedness and global reach: a study of technological development in the Swedish machinery industry] (Vol. 35). Uppsala: Kulturgeografiska Institutionen.
- Liebeskind, J.P., Oliver, A.L., Zucker, L.G., & Brewer, M.B. (1996). Social networks, learning, and flexibility: Sourcing scientific knowledge in new biotechnology firms. *Organization Science*, 7, 428–443.
- Lundequist, P., & Power, D. (2002). Putting Porter into practice? Practices of regional cluster building: Evidence from Sweden. *European Planning Studies*, 10, 685–704.
- Lundmark, M., & Malmberg, A. (1988). Industrilokalisering i sverige: Regional och strukturell förändring [Industrial location in Sweden: regional and structural change] (Vol. 19). Uppsala: Kulturgeografiska Institutionen vid Uppsala Universitet.
- Lundvall, B.-Å. (Ed.). (1992). National systems of innovation: Towards a theory of innovation and interactive learning. London: Pinter.
- Malmberg, A. (2002). Klusterdynamik och regional näringslivsutveckling: Begreppsdiskussion och forskningsöversikt [Cluster dynamics and regional industrial growth: a conceptual discussion and research overview] (No. A2002:008). Stockholm: Institutet för Tillväxtpolitiska Studier.
- Markusen, A. (1999). Fuzzy concepts, scanty evidence, policy distance: The case for rigour and policy relevance in critical regional studies. *Regional Studies*, 9, 869–884.
- Martin, R., & Sunley, P. (2003). Deconstructing clusters: Chaotic concept or policy panacea? *Journal of Economic Geography*, 3, 5–35.
- Maskell, P. (1992). Nyetableringer i industrien: Og industristrukturens udvikling [Industrial startups: and the development of the industrial structure]. Copenhagen: Handelshøjskolens Forlag/Nyt Nordisk Forlag Arnold Busck.
- McCann, P. (2001). *Urban and regional economics*. Oxford: Oxford University Press.
- McCutchen, W.W.J., & Swamidass, P.M. (1996). Exploring larger biotech research firm strategies: Projections from a comparison of small and larger firms. *Technological Forecasting and Social Change*, 51(2), 173–184.
- OECD Observer. (1999, March). Biotechnology and industry: A union of promise. *OECD Observer*, 216. Retrieved September 29, 2008, from <http://www.oecdobserver.org/news/fullstory.php?aid=2>
- Porter, M.E. (1990). *The competitive advantage of nations*. London: Macmillan.
- PRV Bolag. (2002). Aktiebolag: Nyregistrering (No. PRV 13), Patent- och registreringsverket Bolag [limited-liability companies: registering].
- Sandström, A., Backlund, A., Häggblad, H., Markusson, N., Norgren, L., & Westerlund, L. (2001). The Swedish biotechnology innovation system. VINNOVA Innovation in Focus VF 2001:2, Stockholm: VINNOVA – Verket för innovationssystem/The Swedish Agency for Innovation Systems.
- SIND. (1991). Biotekniken i svensk industri: Bokslut och prognos inför 90-talet [Biotechnology in Swedish industry: financial statements and projections for the 1990s]. Stockholm: Statens Industriverk/Allmänna Förlag.
- Sölvell, Ö., Lindqvist, G., & Ketels, C. (2003). *The cluster initiative greenbook*. Stockholm: Bromma Tryck AB.

- Statistics Sweden. (1992a). *SE-SIC 92: Swedish Standard Industrial Classification 1992* (No. MIS 1992:5). Stockholm: Statistiska Centralbyrån.
- Statistics Sweden. (1992b). *SNI 92: Standard för svensk näringsgrensindelning 1992* (No. MIS 1992:4). Stockholm: Statistiska Centralbyrån.
- Statistics Sweden. (2002). *Statistisk årsbok för Sverige 2003* [Statistical Yearbook of Sweden 2003] (Vol. 89). Stockholm: Statistiska Centralbyrån.
- Statistics Sweden. (2003a). *Företagsregistret* [Business Register]. Retrieved March 20, 2003, from http://www.scb.se/foretagsregistret/innehall/sni2002_sni2002sort.pdf
- Statistics Sweden. (2003b). *Företagsregistret: Variabelbeskrivning* [Business Register: description of variables]. Retrieved May 9, 2003, from <http://www.scb.se/foretagsregistret/innehall/fdbvariabel.asp#s>
- Statistics Sweden. (2007). *Ny näringsgrensindelning 2007* [New standard industrial classification 2007]. Retrieved November 21, 2007, from http://www.scb.se/templates/Standard_147807.asp
- Swedish Tax Agency. (2002a). Information från skattemyndigheten. Dags att starta eget? [Information from the Swedish Tax Agency. Time to start a new business?] (No. RSV 462 utgåva 2). Stockholm: Riksskatteverket.
- Swedish Tax Agency. (2002b). *Skatte-och avgiftsanmälan* [Tax and fee registration] (No. RSV 4620 utgåva 22). Stockholm: Riksskatteverket.
- Swedish Tax Agency. (2003). Information från skattemyndigheten. Så fyller du i 'Skatte-och avgiftsanmälan' [Information from the Swedish Tax Agency. How to fill in the form 'Tax and fee registration'] (No. RSV 418 utgåva 16). Stockholm: Riksskatteverket.
- Waluszewski, A. (2003). Ur Pharmacias aska reste sig ingen Fågel Fenix [No Pheonix bird rising from the ashes of the firm Pharmacia]. *Axess*, 3, 25–28.
- Waxell, A. (2005). Uppsalas biotekniska industriella system: En ekonomisk-geografisk studie av interaktion, kunskapsspridning och arbetsmarknadsrörlighet [The Uppsala biotech cluster – economic-geographical studies of interaction, knowledge spillover and labor mobility]. Uppsala: Department of Social and Economic Geography/Kulturgeografiska Institutionen.
- Waxell, A., & Malmberg, A. (2007). What is global and what is local in knowledge-generating interaction? The case of the biotech cluster in Uppsala, Sweden. *Entrepreneurship and Regional Development*, 19, 137–159.
- Zucker, L.G., & Darby, M.R. (1997). Present at the biotechnical revolution: Transformation of technological identity for a large incumbent pharmaceutical firm. *Research Policy*, 26, 429–446.
- Zucker, L.G., Darby, M.R., & Armstrong, J. (1998). Geographically localized knowledge: Spillovers or markets? *Economic Inquiry*, XXXVI, 65–86.