Knowledge diversity and firm growth:

Searching for a missing link

Markus Grillitsch (markus.grillitsch@circle.lu.se)
CIRCLE, Lund University, Sweden

Torben Schubert (torben.schubert@circle.lu.se)
CIRCLE, Lund University, Sweden & Fraunhofer ISI, Germany

Martin Srholec (martin.srholec@circle.lu.se)
CIRCLE, Lund University, Sweden & CERGE-EI, Czech Republic

Papers in Innovation Studies

Paper no. 2016/13

This is a pre-print version of a paper that has been submitted for publication to a journal.

This version: April 2016

Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE)
Lund University
P.O. Box 117, Sölvegatan 16, S-221 00 Lund, SWEDEN
http://www.circle.lu.se/publications

WP 2016/13

Knowledge diversity and firm growth: Searching for a missing link

Markus Grillitsch, Torben Schubert, Martin Srholec

Abstract The link between knowledge and firm growth has been a core topic in economics of innovation for a long time. However, despite strong theoretical arguments, empirical evidence remains inconclusive. One important reason for this conundrum may be the failure of standard indicators to comprehensively capture firm innovation activities. We contribute to overcoming this limitation by zooming in on the knowledge processes that drive variegated forms of innovation and aim thereby to establish a solid relationship with firm growth. The paper draws on the differentiated knowledge base approach, distinguishing between analytical, synthetic, and symbolic knowledge, and measures these types of knowledge with detailed longitudinal linked-employer-employee micro data from Sweden. Econometric findings indicate positive relationships between the three knowledge types, in particular combinations thereof, and firm growth. These relationships remain robust in a wide range of models. Our analysis therefore suggests that the seemingly weak relationship between firm growth and innovation may be explained by the narrow measurement concepts that have dominated in this literature so far.

Keywords: Knowledge; innovation; firm growth; micro data; Sweden

JEL: C33; D22; O12; O32; O33

Knowledge diversity and firm growth: Searching for a missing link

Markus Grillitsch
Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE), Lund University, P.O. Box 117, SE-221 00 Lund, Sweden.

Torben Schubert
Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE), Lund University, P.O. Box 117, SE-221 00 Lund, Sweden. Fraunhofer ISI, Breslauer Str. 46, 76135 Karlsruhe, Germany

Martin Srholec (corresponding author)
Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE), Lund University, P.O. Box 117, SE-221 00 Lund, Sweden; CERGE-EI, Economics Institute, Czech Academy of Sciences, P.O. Box 882, Politických vězňů 7, Prague 1, 111 21, Czech Republic; phone: 046-222 00 00; e-mail: martin.srholec@circle.lu.se.

Abstract

The link between knowledge and firm growth has been a core topic in economics of innovation for a long time. However, despite strong theoretical arguments, empirical evidence remains inconclusive. One important reason for this conundrum may be the failure of standard indicators to comprehensively capture firm innovation activities. We contribute to overcoming this limitation by zooming in on the knowledge processes that drive variegated forms of innovation and aim thereby to establish a solid relationship with firm growth. The paper draws on the differentiated knowledge base approach, distinguishing between analytical, synthetic, and symbolic knowledge, and measures these types of knowledge with detailed longitudinal linked-employer-employee micro data from Sweden. Econometric findings indicate positive relationships between the three knowledge types, in particular combinations thereof, and firm growth. These relationships remain robust in a wide range of models. Our analysis therefore suggests that the seemingly weak relationship between firm growth and innovation may be explained by the narrow measurement concepts that have dominated in this literature so far.

Keywords: Knowledge; innovation; firm growth; micro data; Sweden.

JEL codes: C33, D22, O12, O32, O33.

1. Introduction

Despite the great efforts to understand the drivers of firm growth the results have often remained contradictory and ambiguous varying largely also with the observation period. Early studies on this topic have consistently argued in line with Gibrat's law (1931). This law argues that firm growth is essentially random (Geroski et al, 1997) because too many factors chip in with little overall impact of any. On the contrary, more recent evidence indicates that there are at least some recognizable patterns. In particular, Coad and Rao (2008) sparked renewed interest in the study of the relationship between innovation and growth by finding that R&D and patenting is crucial for the top growers. Likewise Lee (2010), Demirel and Mazzucato, (2012), Deschryvere (2014) and Triguero and Córcoles (2014) followed the suit by showing that although innovativeness is not related to the growth of most firms; there are segments of the population of firms for which the positive link holds.

Generally speaking, however, this expanding empirical literature has been unable to establish that innovation is a robust driving force behind firm growth. Coad (2007) calls for recognizing the paradox that despite the widespread agreement on the positive relationship, many empirical studies have difficulties in verifying the link. One reason for this ambiguity may be that traditional research-centred indicators such as R&D, patents and technological innovation counts only partially measure the relevant processes that boost growth. Hence, rather than focusing on the traditional innovation indicators, we propose to zoom in on the relevant types of knowledge and their combinations which drive different forms of innovation and thereby firm growth.

The argument on the importance of different knowledge types and knowledge combinations for innovation processes can be traced back to Schumpeter's classical writings (1911) and has been an important topic since (e.g. Fleming, 2001, Jensen et al. 2007). In this paper, we use the knowledge base approach introduced by Laestadius (1998) and Asheim and Gertler (2005) in order to capture the different types of innovation-relevant knowledge. This approach is particularly useful, because it explicitly establishes the connection between the different types of knowledge, learning processes of firms and innovation outputs. It distinguishes between three types of knowledge that map on three different modes of innovation: analytical knowledge represents the traditional science-based modes of innovation that build on knowledge about natural laws, synthetic knowledge is more tacit, experiential and applied to concrete problem solving, while symbolic knowledge creates meaning, aesthetic value, brands, and design (Asheim 2007, Asheim et al. 2007). This approach is more encompassing than the measurement concepts based on e.g. patents or R&D, because the latter almost exclusively focus on analytical science-based and possibly to some limited degree synthetic engineering-based modes of innovation. Symbolic modes remain neglected.

Using detailed and extensive microdata from Sweden, we operationalize the knowledge base approach empirically on the basis of evidence on employees' job occupations in firms. The employer-employee occupational data is merged with business registry and financial indicators creating a unique large longitudinal firm-level panel dataset with about one million observations over the period 2004-2011. Unlike most of the existing studies on this topic, the econometric results show that the relationship between innovation-relevant knowledge and firm growth is strong and robust among all types of firms (in contrast to e.g. just high growth firms). Furthermore, the capability of firms to combine different types of knowledge turns out to be highly growth enhancing. Nevertheless, there are limits to the positive relationship, as the results indicate that increasing investment in specific knowledge bases beyond a certain point leads to declining (absolute) returns in terms of growth.

The paper is organized as follows. We present the theoretical background and previous evidence on the topic in Section 2. The database, variables, and identification strategy are explained in Section 3. The empirical analysis is presented in Section 4. Conclusions are derived in Section 5.

2. Theory and previous results

2.1 The relationship between knowledge, innovation and firm growth

Dating back more than 80 years Gibrat (1931) conducted one of the first systematic studies on firm growth resulting in "Gibrat's law", which states that growth rates are independent of firm size. From a theoretical lens, Gibrat's has soon come under attack. In particular, evolutionary economists have argued that innovation is an activity that creates asymmetries in firm capabilities bestowing the innovating firms with a competitive advantage that allows them to grow (Dosi 1988). This asymmetry works through two channels. First, innovation leads to differential product or service characteristics lending a competitive advantage to firms with superior goods (e.g. Dasgupta, 1985). Second, innovation implies organizational learning (Phillips 1971), which will strengthen dynamic capabilities (Teece et al. 1998, Eisenhardt and Martin 2000) and unique knowledge (Grant 1996), which is hard to imitate and therefore creates a competitive advantage (Barnes 1991). In this respect innovation creates growth-inducing asymmetries both on the level of the product and service characteristics as well as in the abilities to create future innovations. Besides the evolutionary asymmetry-argument, there is the view that innovation tends to produce entry barriers that limit the number of competitors and thus leads to market concentration and growth of firms (Del Monte and Papagni, 2003).

Nevertheless, empirical evidence shows a mixed picture. Review articles by Sutton (1997), Caves (1998) and Geroski (1995) indicate that earlier studies focusing primarily on large firms gave support to Gibrat's law by showing that growth appeared to be largely independent of various firm level characteristics. However, once studies emerged that used data on a broader spectrum of firms it became apparent that there are systematic patterns in the growth of firms (Sutton and Caves, 1998, Geroski, 1995). Following these results most authors dealing with firm growth tend to regard Gibrat's law as superseded due to its weaknesses in theory and the contradicting evidence. Based on this notion more recent studies continue to analyze firm specific factors that can contribute to explaining differences in firm growth. While one line of this literature has focused on demographic variables such as age, size, or sector (e.g. Audretsch et al. 1999, Wagner, 2001, Almus and Nerlinger, 2000), a second line embarked upon insights from evolutionary economics and gave considerable attention to innovativeness of firms.

Despite the strong theoretical arguments for a positive relationship between innovation and growth, empirical evidence remains far from conclusive (Coad, 2009, Audretsch et al. 2014). Some authors established the predicted positive association (e.g. Geroski and Machin, 1992, Yasuda, 2005, Coad and Rao, 2007), others have found a non-significant or even negative effect (e.g. Almus and Nerlinger 1999, Lööf and Heshmati, 2006, Freel and Robson, 2004). More recent studies have shown that the positive link is highly conditional on other firm-level characteristics such as patenting (Demirel and Mazzucato, 2012), the persistence of innovation (Deschryvere, 2014, Triguero and Córcoles, 2014), or the internal vs. external R&D (Segarra and Teruel, 2014). Hence, the results suggested that the relationship depends on detailed characteristics of the firms' innovation behavior. By relying on R&D, patent and technological innovation counts, however, this literature has largely ignored the variegated forms of innovation and underlying knowledge processes. We will review some of these insights now.

2.2 The multidimensional nature of innovation and the differentiated knowledge base approach

While not explicitly making the link to firm growth, there are a number of empirical studies highlighting the multi-dimensional nature of innovation. Cesaratto and Mangano (1993), Hollenstein (1996, 2003), de Jong and Marsili (2006), Jensen et al. (2007) and Leiponen and Drejer (2007) showed that besides the traditional science-based innovations, there is a variety of market-oriented and process, production, supplier-driven paths to innovation. Frenz and Lambert (2009) recognized the so-called wider innovating mode by taking into account evidence on organizational and marketing changes. Srholec and Verspagen (2012) identified what they dubbed research, user, external and production ingredients of innovation strategies.

As innovation comes in many forms also the required knowledge is likely to differ. Accordingly, we argue that a broad understanding of innovation-relevant knowledge is needed to establish a solid conceptual and empirical link to firm growth. The knowledge base approach, introduced by Laestadius (1998) and Asheim and Gertler (2005) is well suited for this purpose. It rests on the argument that innovation outputs are ultimately linked to underlying knowledge dynamics, including the type of knowledge used in innovation processes, the routines to generate new knowledge, and the actors involved in innovation processes. The knowledge base approach distinguishes between an analytical, a synthetic, and a symbolic knowledge base (Asheim, 2007, Asheim et al. 2007).

The analytical knowledge base largely draws on the development and application of basic science such as natural laws (Moodysson et al. 2008). Analytical knowledge requires employees with a high level of academic and scientific training. This also implies that learning takes place in dispersed scientific communities, that the resulting knowledge is usually codified, and that localization and geographical distance are of minor importance because the knowledge is constant across different geographical contexts (Martin and Moodysson 2013).

A synthetic knowledge base is mainly about solving concrete problems associated with specific applications. Frequently, this process involves interactive learning between users and producers, and collaborators. That is why the synthetic knowledge base is usually tacit and more tied to space (Asheim and Hansen 2009). The focus on concrete problem solving requires well-trained technicians, often with background from university or engineering colleges, who have developed a high level of skill and craftsmanship through on-the-job training and learning by doing.

The symbolic knowledge base rests on creating meaning, desire and aesthetic values such as design and brands (Asheim et al. 2007). New knowledge is generated in creative processes typically in specifically assembled project teams. Symbolic knowledge tends to be highly tacit and embedded in the context in which it was created (Martin and Moodysson 2011). It usually requires a deep understanding of the culture, norms, habits, values and everyday practices of specific social groups making it difficult to transfer this type of knowledge to other contexts and space. Nevertheless, university training in specific fields such as arts and design are crucial for symbolic knowledge bases.

2.3 The hypotheses

Pina and Tether (2016) show empirically that the drivers of innovation and types of innovation activities of knowledge intensive business services differ notably by knowledge bases. Based on a large-scale survey, Aslesen and Freel (2015) find that the firm-internal organisation of innovation processes as well as channels and geography of knowledge sources depend on the dominant knowledge bases of industries. Herstad et al. (2014) provide evidence that the engagement in global innovation networks is influenced by the type of knowledge base firms hold.

All three types of knowledge bases have been shown to sustain innovation processes and competitive advantage of firms. Asheim and Grillitsch (2015), for instance, show that the maritime industry in a semi-peripheral region in Norway generated world market leaders by drawing largely on a synthetic knowledge base. Similarly, the increasing role of symbolic knowledge related to design and aesthetic innovation processes (Creusen and Schoormans 2005, Krippendorff 2006, Eisenman 2013) have been proven to considerably contribute to firm performance (Bloch 1995, Gemser and Leenders 2001, Hertenstein et al. 2005, Martin and Moodysson 2011). Following the knowledge base approach, therefore, all three knowledge bases are expected to be relevant, which leads to our baseline hypothesis:

> *H1: The presence of analytical, synthetic as well as symbolic knowledge bases increases firm growth.*

Several authors have argued that in particular the combinations of different types of innovation explain differences in firm performance (Gera and Gu 2004, Damanpour and Aravind 2012, Le Bas et al. 2015). Brown and Duguit (1991) note that technological and non-technological innovations are usually co-produced, which results from the fact that the latter follow in the wake of the former (Brown 2002). Likewise, Schubert (2010) shows that non-technological innovations can have a profound effect on the success of product and process innovations. Tavassoli and Karlsson (2015) find evidence that firms introducing both technological and non-technological innovations have a higher labour productivity. While these works make an empirical case for the existence of complementarities between different kinds of innovation strategies, the focus on innovation outputs makes it difficult to derive conceptual justifications for the complementarities, as the actual learning processes leading to innovation are ignored.

In this respect, the knowledge base literature is more specific. Moodysson et al. (2008) argue for a complementarity between the analytical and synthetic knowledge bases in the life science cluster in Scania. As synthetic knowledge is strongly based on experiential knowledge, the complementarity arises because analytical knowledge can help designing experimental settings that are a priori promising, thus avoiding unguided trial-and-error learning. Based on a case study of agro-food and biotechnology in Swedish and Canadian clusters, Coenen et al. (2006) argue that although biotechnology is more focused on analytical and agro-food more on synthetic knowledge, in both sectors there are strong signs that both knowledge bases are combined. Accordingly, Martin and Moodysson (2011) find that new media companies typically need to mobilise analytical, synthetic and symbolic knowledge bases in sequence during an innovation project.

Strambach and Klement (2012) introduce the distinction between cumulative knowledge dynamics, which is learning on the base of previous experience within a knowledge base, and combinatorial knowledge dynamics, which refers to the combination of initially separated knowledge bases. Based on evidence of 62 case studies in 22 European regions, they argue that in particular radical innovation processes increasingly require the latter. Manniche (2012), using the same empirical evidence, points out that the different knowledge bases can be well identified but that they are often combined in innovation processes within firms. Tödtling and Grillitsch (2015) show in a study on the ICT sector in Austrian regions, that firms are indeed more likely to generate products new to the market if they combine different types of knowledge through collaboration or recruitment from diverse types of partners and geographical scales. In a large-scale quantitative study using Swedish registry data Grillitsch et al. (forthcoming) find that in particular the combination of analytical, synthetic and symbolic knowledge boosts innovation performance of firms. From this follows the second hypothesis:

H2: The growth-enhancing effect of combinations of various knowledge bases is stronger than that of the isolated knowledge bases.

Despite the general expectation of a positive relationship between knowledge and growth, there is considerable evidence of heterogeneity. Coad and Rao (2008) show that the leverage of innovation input as measured by R&D is stronger for fast-growing firms due to the high risks and costs of innovation processes. They argue that firms devoting large resources to R&D but failing to obtain valuable results display negative or at least weaker association between innovation input and growth. The same argument can be extended to knowledge bases. In fact, this effect may even be stronger for knowledge base combinations because of their presumably high growth impact. Conversely, incremental innovations that are typically associated with moderate growth potentials, such as improving existing products or processes, are a typical result of cumulative knowledge dynamics within a knowledge base. Based on these expectations we derive our third hypothesis:

H3: The growth-enhancing effect of the knowledge bases, and in particular knowledge base combinations, is strongest for firms in the upper part of the growth distribution.

So far we have primarily focused on the benefits of investing into knowledge bases. However, there is an extensive literature reminding about the fact that innovation entails considerable costs with unknown outcome (Mansfield et al. 1977, Bloom and van Reenen 2002, Coad and Rao 2008). The costs do not only contain expenditures in the form of the direct resources devoted to innovation (diMasi et al. 2003), but also relate to difficulties in knowledge integration (Grant 1996, Grimpe and Kaiser 2010), costs for protecting knowledge assets either by formal and informal protection mechanisms (Besson 2008, Schubert 2011), creating complementary assets (Teece 1986) and costs for overcoming institutional tensions, which may occur when key employees are vested in established technologies (Behagel et al. 2011, Schubert and Andersson 2015). Furthermore, innovations often show a high level of associated risk (Mansfield et al. 1977, Eliasson 1991, Kerr et al. 2014). In particular, when innovation comes in the form of winner-takes-all races for dominant designs (Abernathy and Utterback 1978, Utterback and Suarez 1993) or for key patents (Reignganum 1983, Fudenberg et al. 1983), the risks can be threatening even on the level of the organization. Thus firms need to make careful investment decisions about which innovation projects to follow. If firms rank available innovation projects and then choose the most promising first, firms will sooner or later meet the marginal investment project after which costs exceed the (risk-adjusted) expected outcomes. Thus, both the high costs and high risks imply that innovation and growth should not be monotonously associated. As regards knowledge bases, an important consideration is also that if combinations of different knowledge bases are indeed most conducive for innovation and firm growth, there must be decreasing returns of investing in one specific knowledge base only. Thus, as regards knowledge bases, we expect the following pattern to hold:

H4: The relationship between the relative size of knowledge bases and firm growth follows an inverted u-shape pattern.

3. Data & methodology

The empirical study uses a longitudinal dataset compiled by merging structural business statistics, business registry data, and personal data provided by Statistics Sweden (SCB). Business statistics include data on sales, value added, cash-flow, investments, total assets of firms and industry classifications. This data is complemented with business register data, which provides information about the location and legal form of firms. The personal database covers all individuals aged 16 and over who were registered in Sweden on December 31 of each year. The occupational and educational data at the level of the individual is linked to the respective employer. Unlike the existing empirical studies on innovation and firm growth, which are based on rather limited and selective evidence, the dataset by principle covers the population of all Swedish firms from 2004 to 2011.

As customary in the literature, we measure firm growth by the log-difference of turnover. The variables of main interest capture knowledge bases at the level of firms and are constructed based on occupational data (Asheim and Hansen 2009, Martin 2012, Grillitsch et al. forthcoming). Occupations are classified according to the Swedish Standard Classification for Occupations (SSYK), which is in line with the International Standard Classification of Occupations (ISCO-88). An occupation identifies the type of job that an individual is performing and the minimum qualification typically required for that job. Against the backdrop of the conceptual literature, we identified the occupations that could be associated without doubt to one of the three knowledge bases. This entailed a qualified judgement about the type of knowledge used, the knowledge generation process and the knowledge outcomes associated with each occupation. More details on how this has been done and the resulting grouping of occupations can be found in Annex 1.

Table 1 shows the relative frequency of the occurrence of the three knowledge bases as well as the knowledge base combinations in the longitudinal panel of firms over 2004-2011. The dataset comprises 1,034,734 observations of 225,063 firms.[1] Synthetic knowledge is the most common (13.74%), followed by symbolic knowledge (5.72%), while only a small fraction maintains analytical knowledge (1.33%). Almost every fifth observation harbours at least some element of the innovation-related knowledge base (18.42%). In contrast, knowledge base combinations are quite rare (2.11%). The most frequent combination is synthetic and symbolic (1.25%), followed by analytical and synthetic (0.54%), while the combination of analytical and symbolic is extremely scarce (0.07%). All three knowledge bases appear in-house very sporadically (0.25%).

These figures may appear very low in particular in comparison to data obtained from CIS (reference). Here the share of product innovators was 29%, process innovators accounted for 21%, organizational innovators for 23% and marketing for 30% in 2010-2012. In total, the share of firms conducting any kind of innovation was 50%. However, it should be kept in mind that the Swedish CIS contains only information for firms with 10 employees and above, while we include all firms. If we apply this restriction to our sample, we find that 52% of all firms have at least one knowledge base, of which 5% had an analytical knowledge base, 11% a symbolic and 36% had a synthetic. So, even in terms of prevalence the knowledge base concept is slightly broader then most encompassing CIS definition of innovativeness.

[1] The panel dataset is unbalanced with 36.5% of the firms being observed over the whole period and 85,8% of the firms being present in at least two consecutive periods.

Table 1: Relative knowledge base frequencies in % for 2004-2011 (full sample)

	% of observations
Firms having ...	
... analytical knowledge	1.33
... synthetic knowledge	13.74
... symbolic knowledge	5.72
Firms with at least one knowledge base	*18.42*
Analytical only	0.46
Synthetic only	11.70
Symbolic only	4.15
Firms with knowledge combinations	*2.11*
Analytical & synthetic	0.54
Analytical & symbolic	0.07
Synthetic & symbolic	1.25
All three knowledge bases	0.25
Number of Observations	1,034,734
Number of Firms	225,063

As controls, we include a generic human capital variable measuring the share of employees with tertiary education. This variable is included in order to identify the additional explanatory value of the knowledge base typology as compared to the conventional human capital measurements. Further, we control for firm size by including the logarithm of total sales, the firms' ability to finance growth is accounted for by measuring cash flow per total assets and capital investments per total assets are included to control for capital endowment. Furthermore, we account for the fixed effects of firm location in Swedish counties (20 regions), 2-digit NACE-codes (81 categories), and the year of the observation. Descriptive statistics are provided in Annex 2.

As concerns estimation of H1, H2, and H4, our baseline empirical model follows the standard template of the econometric literature on the growth of firms as follows:

$$\log\left(\frac{\text{turnover}_{i,t}}{\text{turnover}_{i,t-1}}\right) = \alpha + kb_{i,t-1}\beta + x_{i,t-1}\gamma + \lambda \text{county}_{i,t} + \delta \text{industry}_i + \varphi z_t + \mu_i + \varepsilon_{i,t} \qquad (2)$$

where i refers to the firm, and t is time. Thus, we represent growth of firms as a function of the main variables of our interest represented by the knowledge base of the firm ($kb_{i,t-1}$) firm characteristics ($x_{i,t-1}$), industry effects (industry_i), county effects ($\text{county}_{i,t}$), temporal shocks (z_t), unobserved individual effects (μ_i), and random errors ($\varepsilon_{i,t}$). Depending on the assumptions on μ_i this model can be estimated by Random Effects (RE) and Fixed Effects (FE). Even though regular Hausman tests indicated the failure of the zero correlation assumption implying that the FE model is the appropriate estimator, we report also the RE results as a reference. In any case, the results of the key variables of interest do not differ much. To further corroborate this robustness, we also present the OLS results.

As an alternative to Eq. (2) we allow for autocorrelation in firm growth rates, by including the lagged dependent variable:

$$\log\left(\frac{\text{turnover}_{i,t}}{\text{turnover}_{i,t-1}}\right) = \alpha + \eta\left(\frac{\text{turnover}_{i,t-1}}{\text{turnover}_{i,t-2}}\right) + kb_{i,t-1}\beta + x_{i,t-1}\gamma + \lambda \text{county}_{i,t} + \delta \text{industry}_i + \varphi z_t + \mu_i + \varepsilon_{i,t} \qquad (3)$$

which implies a dynamic panel data model. Such models can be estimated by a variety of procedures. We rely on the Arrelano-Bover estimator (Arrelano and Bover 1995) because oft its efficiency properties.[2]

In addition, H3 allows the estimated effects to differ across the growth distribution. To accommodate for this generalization we use quantile regression. Although we are unable to control for the fixed effect through this approach, we use a variance estimator clustered over the cross-section observations to account for the time dependence in the panel observations. In that respect, our estimator mimics a random effects quantile regression.

4. Results

Table 2 presents the benchmark results using the pooled OLS, RE, FE and AB estimators in the respective columns. Overall, the knowledge base coefficients are highly statistically significant regardless of the estimator, although their magnitude somewhat differs depending on the underlying assumptions, but the main picture is clear. All three knowledge bases are conducive to firm growth, which corroborates our baseline hypothesis (H1). Moreover, the effect of knowledge base combinations is markedly stronger than of any single knowledge base alone (H2). According to the AB model, for instance, firms with all three knowledge bases are estimated to grow by about 23 percentage points and those with two-way combinations by 16 to 19 percentage points faster than the base category of firms without the innovation-relevant knowledge bases. Indeed, this is a healthy boost given the fact that the average growth rate is around 5% only.

[2] In order to implement these models, it is important to make decisions on how lagged instruments are included. This can be assessed by a test on the AR-1 and AR-2 components. In particular, if the model is correctly specified the AR-1 component is significant while the AR-2 component is not. A high number of different specifications, including with and without restrictions on the time lags, with and without including instruments in the levels and with and without collapsing instruments, have been tested (Roodman 2009). The best specification involves the inclusion of the IVs in levels, no restrictions on the included lags, and no collapsing of instruments. Nonetheless, even though other specifications tend to fail the AR-1 and AR-2 tests, the key results remain qualitatively similar.

Table 2: Regressions of knowledge base combinations on firm growth

	(1) OLS	(2) RE	(3) FE	(4) AB
KB: analytic only	0.0472***	0.0660***	0.0374***	0.1065***
	(0.0066)	(0.0070)	(0.0075)	(0.0112)
KB: synthetic only	0.0639***	0.1020***	0.0547***	0.0966***
	(0.0016)	(0.0019)	(0.0022)	(0.0031)
KB: symbolic only	0.0110***	0.0340***	0.0442***	0.0779***
	(0.0024)	(0.0030)	(0.0035)	(0.0049)
KB: analytic & synthetic	0.1501***	0.2160***	0.1172***	0.1871***
	(0.0057)	(0.0069)	(0.0072)	(0.0091)
KB: analytic & symbolic	0.0912***	0.1377***	0.0909***	0.1562***
	(0.0152)	(0.0155)	(0.0150)	(0.0188)
KB: synthetic & symbolic	0.1346***	0.2131***	0.1195***	0.1756***
	(0.0037)	(0.0047)	(0.0051)	(0.0063)
KB: all three KB	0.2096***	0.3708***	0.1717***	0.2344***
	(0.0064)	(0.0101)	(0.0107)	(0.0129)
Log turnover	-0.0407***	-0.1444***	-0.5935***	-0.0597***
	(0.0004)	(0.0005)	(0.0010)	(0.0004)
Cash-flow per total assets	0.0122	0.0228	0.0389**	-0.0049
	(0.0351)	(0.0167)	(0.0164)	(0.0215)
Capital investments per total assets	5.4501***	5.9319	3.7222	3.8000
	(1.3366)	(5.3552)	(4.9427)	(5.7886)
Share of employees w. tertiary education	0.0063***	-0.0012	-0.0071**	-0.0413***
	(0.0018)	(0.0023)	(0.0034)	(0.0016)
Growth (t-1)	..	..	..	-0.0956***
				(0.0013)
Constant	0.7027***	2.2736***	9.1718***	0.9469***
	(0.0062)	(0.0089)	(0.0157)	(0.0056)
County dummies	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	..	..
Year dummies	Yes	Yes	Yes	Yes
Observations	1034734	1034734	1034734	798689
Firms	225063	225063	225063	190978
R2 / R2-within	0.029	0.226	0.337	..
F / chi2	219***	99037***	11407***	809868***
AB AR1 test	..	..	..	-288.98***
AB AR2 test	..	..	..	-0.87

Note: Standard errors in parentheses; standard errors of OLS regressions are clustered at the level of the firm; ***, **, * indicate significance at the 10%, 5%, and 1% levels; R2 is reported for OLS regressions, R2-within for random effects (RE) and fixed effects (FE) regressions, F-statistics are reported for OLS and FE regressions; Wald Chi2 for RE and Arellano–Bond regressions.

Table 3 presents the effects estimated at different quantiles of firm growth rates. The results confirm that the knowledge bases and combinations thereof are particularly relevant for high growth firms (H3). The estimated effects of knowledge bases on firm growth approximately double for the 75th quantile as compared to the median. If the 99th quantile of the fastest growing firms is considered, the jump is very large. All else equal, if all three knowledge bases are present these top performers are expected to record as much as 73 percentage points higher growth as compared to the base category. Interestingly, however, for the fastest growing firms, symbolic knowledge alone does not seem to contribute much, which further underlines the need for combinations. While the magnitude of the estimated coefficients is potentially inflated, because this estimator does not control for unobserved heterogeneity and growth autocorrelation, the previous results indicated that the bias is not likely to drive the main conclusions.

Table 3: Quantile regressions of knowledge base combinations on firm growth

	(1) Q(0.25)	(2) Q(0.5)	(3) Q(0.75)	(4) Q(0.99)
KB: analytic only	0.0128***	0.0197***	0.0430***	0.2073***
	(0.0049)	(0.0045)	(0.0056)	(0.0277)
KB: synthetic only	0.0138***	0.0212***	0.0478***	0.2693***
	(0.0014)	(0.0009)	(0.0015)	(0.0142)
KB: symbolic only	-0.0055**	0.0015	0.0087***	0.0186
	(0.0022)	(0.0015)	(0.0022)	(0.0175)
KB: analytic & synthetic	0.0524***	0.0541***	0.0936***	0.4867***
	(0.0048)	(0.0034)	(0.0051)	(0.0590)
KB: analytic & symbolic	0.0452***	0.0244***	0.0415***	0.3203***
	(0.0090)	(0.0051)	(0.0078)	(0.0881)
KB: synthetic & symbolic	0.0521***	0.0395***	0.0716***	0.4468***
	(0.0027)	(0.0020)	(0.0038)	(0.0302)
KB: all three KB	0.0646***	0.0455***	0.0910***	0.7334***
	(0.0036)	(0.0039)	(0.0045)	(0.0641)
Log turnover	0.0017***	-0.0075***	-0.0310***	-0.2076***
	(0.0003)	(0.0002)	(0.0003)	(0.0020)
Cash-flow per total assets	0.0106*	-0.0209	0.0253***	0.0152***
	(0.0060)	(0.1454)	(0.0021)	(0.0035)
Capital investments per total assets	6.4885***	4.9537***	3.9297***	-15.7389***
	(0.1468)	(0.4868)	(0.2771)	(0.4235)
Share of employees w. tertiary education	-0.0177***	0.0068***	0.0396***	0.1932***
	(0.0015)	(0.0010)	(0.0018)	(0.0155)
Constant	-0.0757***	0.1762***	0.6752***	4.1770***
	(0.0046)	(0.0032)	(0.0052)	(0.0382)
County dummies	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Observations	1034734	1034734	1034734	1034734
Firms	225063	225063	225063	225063
R^2	0.006	0.020	0.016	0.013

Note: t-statistics in parentheses; standard errors of OLS regressions are clustered at the level of the firm; ***, **, * indicate significance at the 10%, 5%, and 1% levels

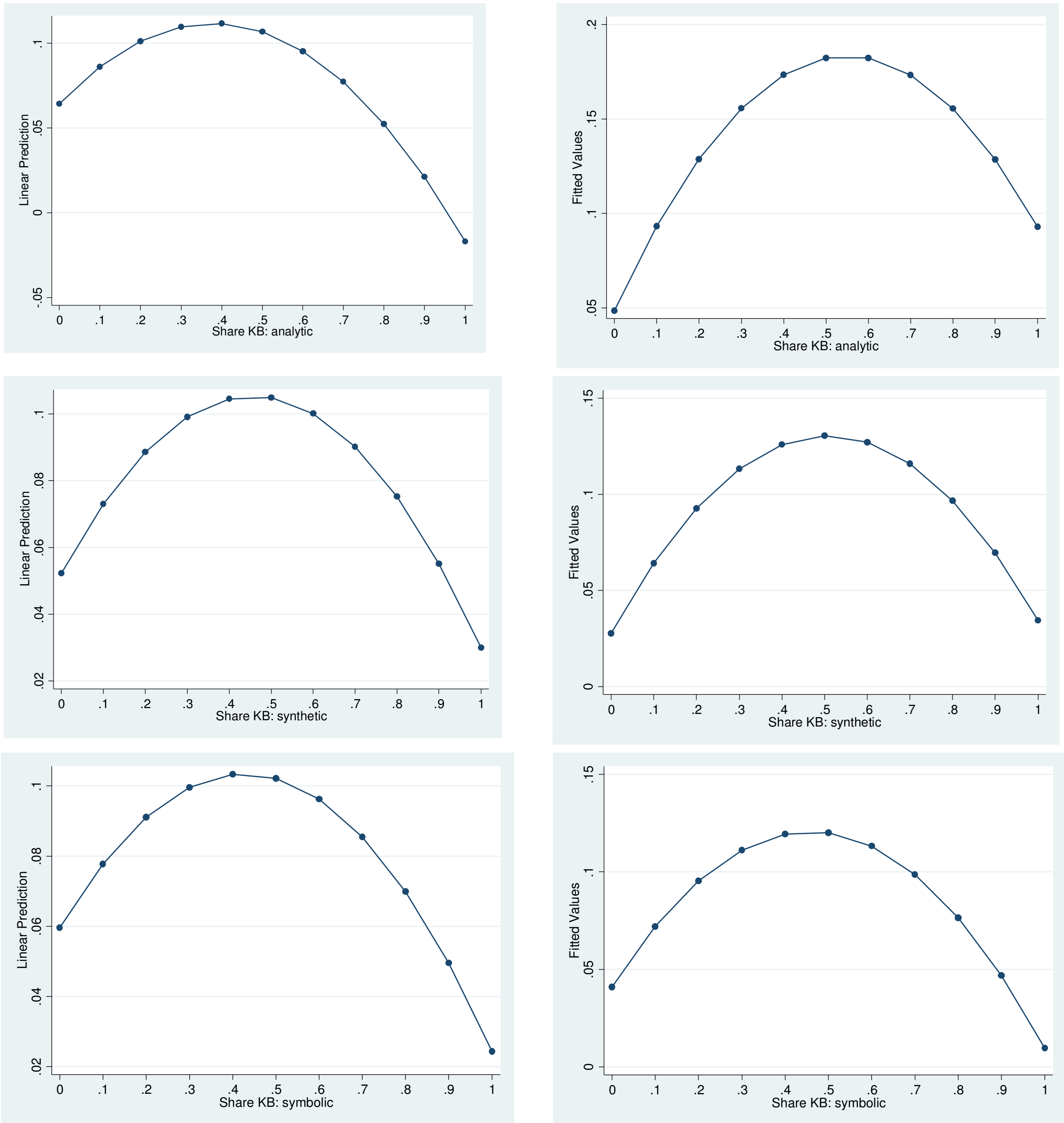
Table 4 provides the analysis of potential curvilinear effects of knowledge base intensities on growth of firms. The first estimate gives the linear effects only, while the second estimate includes squared terms of the respective shares. Only results of the more consistent FE and AB estimators are presented for the sake of brevity. The results confirm that these relationships generally follow an inverse u-shaped curve (H4). The FE estimate indicates that increasing analytical knowledge within the firm contributes to firm growth until a share of approximately 38% is reached. Further increasing analytical knowledge beyond the 38% share, however, leads to declining contribution to firm growth. Similar inversed u-shaped relationships with turning points at 46% and 41% are detected for synthetic and symbolic knowledge bases, respectively The AB estimates indicate slightly higher turning points, but prove robust. Graphical representations of these relationships can be found in Figure 1. Again, these results highlight the benefits from combining knowledge bases, as the fastest growth is not achieved by ever increasing specialisation in one knowledge base but – especially when reaching the respective turning points – by tapping into other types of knowledge.

Table 4: FE and AB regressions of shares of knowledge bases on firm growth

	(1) FE	(2) FE	(3) AB	(4) AB
Share KB: analytic	-0.0233	0.2509***	0.1466***	0.4901***
	(0.0149)	(0.0401)	(0.0282)	(0.0577)
Share KB: synthetic	0.0052	0.2324***	0.0581***	0.4025***
	(0.0047)	(0.0123)	(0.0081)	(0.0170)
Share KB: symbolic	-0.0103	0.2054***	0.0080	0.3444***
	(0.0070)	(0.0182)	(0.0111)	(0.0252)
Share KB: (analytic)square	..	-0.3321***	..	-0.4462***
		(0.0443)		(0.0669)
Share KB: (synthetic)square	..	-0.2547***	..	-0.3969***
		(0.0127)		(0.0183)
Share KB: (symbolic)square	..	-0.2407***	..	-0.3789***
		(0.0187)		(0.0263)
Log turnover	-0.5911***	-0.5926***	-0.0502***	-0.0544***
	(0.0010)	(0.0010)	(0.0003)	(0.0003)
Cash-flow per total assets	0.0397**	0.0395**	-0.0046	-0.0051
	(0.0164)	(0.0164)	(0.0216)	(0.0216)
Capital investments per total assets	3.7041	3.6982	3.6907	3.6909
	(4.9460)	(4.9440)	(5.8039)	(5.7982)
Share of employees w. tertiary education	-0.0042	-0.0052	-0.0197***	-0.0273***
	(0.0034)	(0.0034)	(0.0019)	(0.0019)
Growth (t-1)	..	..	-0.0970***	-0.0964***
			(0.0013)	(0.0013)
Constant	9.1443***	9.1649***	0.8150***	0.8763***
	(0.0157)	(0.0157)	(0.0048)	(0.0053)
County dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Observations	1034734	1034734	798689	798689
Firms	225063	225063	190978	190978
R2-within	0.336	0.336	..	..
F / chi2	12781***	11714***	59135***	59921***
AB AR1 test	..	..	-289.39***	-288.97***
AB AR2 test	..	..	-0.94	-0.88
Analytical: Turning point	..	38%	..	55%
Synthetic: Turning point	..	46%	..	50%
Symbolic: Turning point	..	41%	..	44%

Note: Standard errors in parentheses; ***, **, * indicate significance at the 10%, 5%, and 1% levels; F-statistics are reported for FE regressions; Wald Chi2 for Arellano–Bond regressions.

Figure 1: Estimated relationship between the share of the knowledge bases and firm growth (FE left panels; AB right panels)



6. Conclusions

This paper re-examined the long-standing question of how knowledge, innovation and growth are linked to each other at the firm level. Despite strong theoretical support of a positive relationship, the empirical evidence has not provided robust evidence. This may have to do with the most commonly used innovation indicators, which tend to relate to limited forms of innovation. In contrast, we approach this question from the root by focussing on the types of knowledge, and combinations thereof, which are relevant and important to generate multiple forms of innovation. This has allowed us to establish highly robust and positive relationships with firm growth.

In order to capture innovation-relevant knowledge, we draw on the differentiated knowledge base approach. The knowledge base approach is explicitly based on the notion that there are different modes of innovation. In this respect, the knowledge base approach very directly takes into account what firms actually do, when they innovate, instead of hiding these activities behind abstract figures such as R&D expenditures or patents. This also allows for a much broader understanding of when certain types of innovation processes may be beneficial. The differentiated knowledge base approach distinguishes between analytical, synthetic, and symbolic knowledge. Traditional innovation indicators mainly refer to analytical and to some degree to synthetic knowledge. Thus, the relevant types of knowledge driving different forms of innovation are captured in a much broader way than previous studies on firm growth have done.

Based on an empirical operationalization of this approach we find that there is a very robust relationship between the knowledge bases and firm growth across a wide range of estimation approaches, controlling also for unobserved heterogeneity and autocorrelation in growth. In addition to the analytical and synthetic knowledge base, our results show that the introduction of symbolic knowledge is important for explaining firm growth. This is in line with the literature on design and aesthetic innovation processes (Creusen and Schoormans 2005, Krippendorff 2006, Eisenman 2013) for which several authors have shown that they considerably contribute to firm performance (Bloch 1995, Gemser and Leenders 2001, Hertenstein et al. 2005). Beyond this, the results show that the combination of two or more knowledge bases has by far the strongest effect on firm growth.

Moreover, our findings resonate well with studies showing that the most innovative firms combine different types of innovation and knowledge (Jensen et al. 2007; Tödtling and Grillitsch, 2015; Grillitsch et al. forthcoming). Jensen et al. (2007), for instance, found that firms are most innovative if they combine science and technology (STI) driven innovations with innovations based on learning through doing, using, and interacting (DUI). This approach can also work in the strategic innovation literature that conceives innovation to be based on knowledge recombination (Fleming 2001, Nerkar and Rosenkopf 2003, Yayavaram and Ahuja 2008, Neuhäusler et al. 2015), because the knowledge base concept allows for a very direct measurement of knowledge combinations. Since most of the latter literature relied on patent data in order to measure knowledge, the analyses needed to be restricted to patent-intensive sectors. This shortcoming does not apply to the knowledge base approach, which can principally be used for all sectors and firms.

Evidence of a positive link between the knowledge bases and firm growth shall however not hide potential pitfalls. First, firms at the upper part of the growth distribution appear to experience a stronger link between the existence of the knowledge bases and growth. Coad and Rao (2008) predicted this effect because of the interaction of costs and risks of innovation projects. But it could also result from strategic differences. High-growth firms grow faster because of unique products with considerable consumer value. Sustaining this competitive advantage very often requires innovation activities, which makes high-growth firms more reliant on innovation. On the other hand, innovation may be less crucial for firms operating in stable environments with relatively settled market shares. Second, we also find evidence that it does not pay off to simply increase the relative importance of a specific knowledge base without limits. We found turning points above which a further specialisation in a specific knowledge base becomes detrimental.

The chosen approach of focussing on the types of knowledge that are relevant for generating innovation and promoting firm growth has yielded robust and strong empirical results. Accordingly, it may be promising for future studies to investigate the underlying sources of innovation and growth, thereby complementing the large number of existing studies using the traditional innovation indicators. The typology of the differentiated knowledge base approach has proven to be a good starting point for such research efforts. However, we are aware that the three types of knowledge may not be the only ones that are relevant for generating innovation. For instance, the knowledge base approach does not consider knowledge or firm capabilities to integrate different types of knowledge or to manage innovation processes, or complementary assets to turn innovation into growth. It would therefore be useful to deepen work on singling out which types of knowledge drive innovation and growth. Furthermore, it would be interesting to better understand the interplay between the firm-internal and firm-external sources of firm growth. This relates to one limitation of this paper, which is that we could not capture external sources of knowledge. This is the price that we were prepared to pay for using population wide register data of high reliability.

Acknowledgements

Financial support from VINNOVA Core Funding of Centers for Innovation Systems Research project 2010–01370 on “Transformation and Growth in Innovation Systems: Innovation Policy for Global Competitiveness of SMEs and R&I Milieus” is gratefully acknowledged. All the usual caveats apply.

References

Abernathy, W. J., and J. M. Utterback. 1978. “Patterns of Industrial Innovation.” Technology Review 80 (7): 40–47.

Arellano, M., and O. Bover. 1995. Another look at the instrumental variable estimation of error-components models. Journal of Econometrics 68: 29–51.

Asheim, B. T., Gertler, M. S. (2005). The geography of innovation: regional innovation systems. In: Fagerberg, J., Mowery, D. C. & Nelson, R. R. (Eds) The Oxford handbook of innovation. Oxford University Press, Oxford, pp. 291-317.

Asheim, B. T. 2007. Differentiated knowledge bases and varieties of regional innovation systems. Innovation: The European Journal of Social Science Research, 20(3), 223-241.

Asheim, B. T., Coenen, L., Vang, J., 2007. Face-to-face, buzz, and knowledge bases: sociospatial implications for learning, innovation, and innovation policy. Environment and Planning C, 25**,** 655-670.

Asheim, B. T., Grillitsch, M., 2015. Smart specialisation: Sources for new path development in a peripheral manufacturing region. Papers in Innovation Studies Nr. 2015/11, CIRCLE, Lund University.

Asheim, B. T., Hansen, H. K. 2009. Knowledge Bases, Talents, and Contexts: On the Usefulness of the Creative Class Approach in Sweden. Economic Geography, 85(4), 425-442.

Aslesen, H. W., Freel, M. 2012. Industrial knowledge bases as drivers of open innovation? Industry and Innovation 19, 563–84.

Behagel, L., E. Caroli, and M. Roger. 2011. “Age Biased Technological Change, Training and Employment Prospects of Older Workers.” IZA Discussion Papers No. 5544.

Bessen, J. (2008). The value of U.S. patents by owner and patent characteristics. Research Policy, 37, 932–945.

Bloom, Nick and Van Reenen, John (2002) Patents, real options and firm performance *Economic Journal*, 112 (478). 97-116

Bloch, P. H. 1995. Seeking the ideal form: Product design and consumer response. Journal of Marketing, 59(3): 16-29.

Bond, S. 2002. Dynamic panel data models: A guide to micro data methods and practice. Working Paper CWP09/02, Cemmap, Institute for Fiscal Studies. Available at http://cemmap.ifs.org.uk/wps/cwp0209.pdf.

Brown, J. S. (2002). Research that reinvents the corporation. Harvard Business Review, 80, 105.

Brown, J. S., & Duguid, P. (1991). Organizational learning and communities of practice: Toward a unified view of working, learning, and innovation. Organizational Science, 2, 40–57.

Coad, A., Rao, R., 2008. Innovation and firm growth in high-tech sectors: A quantile regression approach. Research Policy, 37, 633-648.

Coenen, L., Moodysson, J., Ryan, C. D., Asheim, B.T., Phillips, P., 2006. Comparing a Pharmaceutical and an Agro-Food Bioregion: On the Importance of Knowledge Bases for Socio-spatial Patterns of Innovation. Industry and Innovation, 13(4), 393-414.

Creusen, M. E. H., & Schoormans, J. P. L. 2005. The different roles of product appearance in consumer choice. Journal of Product Innovation Management, 22: 63-81.

Damanpour, F., Aravind, D. (2012): Manegerial Innovation: Conceptions, Processes, and Antecedents, Management and Organization Review, 8, 423-454

de Jong, J.P.J., Marsili, O., 2006. The fruit flies of innovations: A taxonomy of innovative small firms. Research Policy, 35, 213-229.

Del Monte, A., Papagni, E., 2003. R&D and the growth of firms: empirical analysis of a panel of Italian firms. Research Policy, 32, 1003-1014.

DiMasi, J.A., Hansen, R.W., Grabowski, H.G. (2003): The price of innovation: new estimates of drug development costs., 22, 151-185

Eisenman, M. (2013): Understanding Aesthetic Innovation in the Context of Technological Evolution, Academy of Management Review, 38. 332-351.

Eliasson, G. (1991): Modeling the experimentally organized economy. Journal of Economic Behavior and Organization, 16, 153-182

Ennen, E., Richter, A., 2010. The Whole is More Than the Sum of Its Parts – Or Is It? A Review of the Empirical Literature on Complementarities in Organizations. Journal of Management, 36, 207-233.
Frenz, M., Lambert, R. (2009). Exploring Non-technological and Mixed Modes of Innovation Across Countries, in OECD (Eds) Innovation in Firms: A Microeconomic Perspective, OECD, Paris, pp. 69-110.
Geroski, P. A., Machin, S. J., Walters, C. F., 1997. Corporate growth and profitability. Journal of Industrial Economics, 45, 171-189.
Fudenberg, D, Gilbert, R., Stiglitz, J., Tirole, J. (1983): Preemption, Leapfrogging, and Compeition in Patent Races, European Economic Review, 22, 3-21.
Gemser, G., & Leenders, M. A. A. M. 2001. How integrating industrial design in the product development process impacts on company performance. Journal of Product Innovation Management, 18: 28-38.
Gera, S. & W. Gu (2004), The Effect of Organizational Innovation and Information Technol-ogy on Firm Performance, International Productivity Monitor 9, 37-51
Geroski, P.A., 1999. The Growth of Firms in Theory and in Practice. CEPR Discussion Paper 2092.
Geroski, P.A., 2005. Understanding the implications of Empirical Work on Corporate Growth Rates. Managerial and Decision Economics, 26, 129-138.
Grant, R.M., 1996. Toward a knowledge-based theory of the firm. Strateg. Manag. J. 17, 109–122.
Gibrat, R., 1931. Les Inégalités Économiques. Paris: Librairie du Recueil Sirey.
Grimpe, C., Kaiser, U., 2010. Balancing internal and external knowledge acquisition: the gains and pains from R&D outsourcing. J. Manag. Stud. 47, 1483–1509.
Grillitsch, M., Martin, R., Srholec, M., forthcoming. Knowledge base combinations and innovation performance in Swedish regions. Economic Geography.
Herstad, S. J., Aslesen, H. W., Ebersberger, B., 2014. On industrial knowledge bases, commercial opportunities and global innovation network linkages. Research Policy 43, 495–504.
Hertenstein, J. H., Platt, M. B., & Veryzer, R. W. 2005. The impact of industrial design effectiveness on corporate financial performance. Journal of Product Innovation Management, 22(1): 3-21.
Hollenstein, H., 1996. A composite indicator of a firm's innovativeness. An empirical analysis based on survey data for Swiss manufacturing. Research Policy, 25, 633-645.
Hollenstein, H., 2003. Innovation modes in the Swiss service sector: a cluster analysis based on firm-level data. Research Policy, 32, 845-863.
Jensen, M.B., Johnson, B., Lorenz, E., Lundvall, B.Å., 2007. Forms of knowledge and modes of innovation. Research Policy, 36, 680-693.
Kerr, W.R., Nanda, R., Rhodes-Kropf, M. (2014): Entrepreneurship as Experimentation, Journal of Economic Perspectives, 28, 25-48
Krippendorff, K. 2006. The Semantic Turn: A New Foundation for Design. Boca Raton, Fl: Taylor and Francis.
Laestadius, S. (1998). Technology level, knowledge formation, and industrial competence in paper manufacturing. In: Eliasson, G., Green, C., Mccann, C. R. (Eds) Microfoundations of Economic Growth: A Schumpeterian Perspective. University of Michigan Press, Ann Arbour, pp. 212-226.
Le Bas, C., C. Mothe & T.U. Nguyen-Thi (2015), The Differential Impacts of Organizational Innovation Practices on Technological Innovation Persistence, European Journal of In-novation Management 18, 110-127
Lee, C-Y., 2010. A theory of firm growth: Learning capability, knowledge threshold, and patterns of growth. Research Policy, 39, 278-289.
Leiponen, A., Drejer, I., 2007. What exactly are technological regimes? Intra-industry heterogeneity in the organization of innovation activities. Research Policy, 36, 1221-1238.
Manniche, J., 2012. Combinatorial Knowledge Dynamics: On the Usefulness of the Differentiated Knowledge Bases Model. European Planning Studies, 20, 1823-1841.
Mansfield, E., Rapoport, J., Romeo, A., Villani, E.,Wagner, S., Husic, F., 1977. The Production and Application of New Industrial Technology. Norton, New York.
Martin, R., 2012. Measuring Knowledge Bases in Swedish Regions. European Planning Studies, 20(9), 1569-1582.

Martin, R., Moodysson, J., 2011. Innovation in symbolic industries: The geography and organization of knowledge sourcing. European Planning Studies, 19**,** 1183-1203.
Martin, R., Moodysson, J. 2013. Comparing knowledge bases: on the geography and organization of knowledge sourcing in the regional innovation system of Scania, Sweden. European Urban and Regional Studies, 20(2), 170-187.
Moodysson, J., Coenen, L., Asheim, B. T., 2008. Explaining spatial patterns of innovation: analytical and synthetic modes of knowledge creation in the Medicon Valley life-science cluster Environment and Planning A, 40**,** 1040-1056.
Pina, K., Tether, B. S., 2016. Towards understanding variety in knowledge intensive business services by distinguishing their knowledge bases. Research Policy, 45(2), 401-413.
Reignganum, J. (1983): Uncertain innovation and the persistence of monopoly, American Economic review, 73, 741-748
Roodman, (2009): How to do xtabond2: An introduction to difference and system GMM in Stata, TheStata Journal, 9, 86-136.
Schubert, T. (2010): Marketing and Organisational Innovations in Entrepreneurial Innovation Processes and their Relation to Market Structure and Firm Characteristics, Review of Industrial Organization, 36, 189-212
Schubert, T. (2011): Assessing the value of patent portfolios: an international country comparison, Scientometrics, 88, 787-804.
Schubert, T., Andersson, A. (2015): Old is gold? The effects of employee age on innovation and the moderating effects of employment turnover, Economics of Innovation and New Technology, 24, 95-113.
Srholec, M., Verspagen, B., 2012. The Voyage of the Beagle into innovation: Explorations on heterogeneity, selection and sectors. Industrial and Corporate Change, 21, 1221-1253.
Strambach, S., Klement, B., 2012. Cumulative and Combinatorial Micro-dynamics of Knowledge: The Role of Space and Place in Knowledge Integration. European Planning Studies, 20**,** 1843-1866.
Sutton, J., 1997. Gibrat's Legacy. Journal of Economic Literature, vol. 35 (1), 40-59.
Tavassoli, Sam & Karlsson, Charlie, 2015. "Persistence of various types of innovation analyzed and explained," Research Policy, Elsevier, vol. 44(10), pages 1887-1901.
Teece, David J., 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. Research Policy, 15, 285-305.
Tödtling, F., Grillitsch, M., 2015. Does Combinatorial Knowledge Lead to a Better Innovation Performance of Firms? European Planning Studies, 23**,** 1741-1758.
Utterback, J. M. and F. F. Suarez (1993). 'Innovation, competition, and industry structure', Research Policy, 22 (1), p 1.

Annex 1: Identification of Knowledge Bases

Following Grillitsch et al. (forthcoming) the knowledge bases are identified using detailed occupational data. The Swedish classification of occupations (SSYK 96) is a national adaptation of the International Standard Classification of Occupations (ISCO-88). Occupations are grouped in a hierarchical framework based on

- The kind of work performed defined as "a set of tasks or duties designed to be executed by one person",
- The skill level defined as "the degree of complexity of constituent tasks", and
- The skill specialization defined as "the field of knowledge required for competent performance of the constituent tasks". (SCB 1998, 17).

The SSYK 96 and ISCO-88 define ten major groups, each of which comprise occupations that require a certain skill level as shown below:

Major Groups		**Skill Level**
1	Legislators, senior officials and managers	-
2	Professionals	4:e
3	Technicians and associated professionals	3:e
4	Clerks	2:a
5	Service workers and shop and market sales workers	2:a
6	Skilled agricultural and fishery workers	2:a
7	Craft and related trade workers	2:a
8	Plant and machine operators and assemblers	2:a
9	Elementary occupations	1:a
0	Armed Forces	-

Skill levels:

4:e	At least three to four years of education starting typically at ages seventeen or eighteen that leads to an academic degree
3:e	Maximum three years of education starting typically at ages seventeen or eighteen not leading to an academic degree
2:a	Completion of upper secondary school/high school
1:a	Requires no or little education

Only major groups 2 and 3 are used for the identification of innovation-relevant knowledge bases for the following reasons:

- Major group 1 consists of individuals performing managing tasks. Managing tasks are general and require different levels of skills, which makes difficult to capture a specific knowledge base.
- Major groups 2 and 3 characterize individuals with a high level of skills and tasks that relate to the concept of knowledge bases.
- Major groups 4 to 9 capture individuals with lower skill levels performing largely routine tasks, being less relevant for the innovation performance of firms.
- Major group 0, i.e. individuals working for armed forces, is not relevant for measuring knowledge bases in firms.

Each major group is divided in a hierarchical framework into submajor groups, minor groups, and unit groups . The assignment of occupations to knowledge bases is done at the most detailed level four-digit level. For each unit group, the SCB (1998) provides a description of the work performed and knowledge required for performing the job, including videos and interviews provided by the Swedish Public Employment Service (Arbetsförmedlingen 2014), on the base of which it is possible to credibly identify the relevant occupations. If the available information did not allow us to clearly identify the knowledge type, we excluded the respective occupation from the analysis; the only exception was the too large to omit occupation "2131 Computer System Designers, Analysts and Programmers.", in the case of which individuals with PhD education were assigned to analytical and the others to synthetic knowledge bases, respectively.

Table 1:*Occupation Groups with Analytical, Synthetic and Symbolic Knowledge Base*

Occupations group (SSYK 96)	
Analytical occupations	
2111	Physicists and astronomers
2112	Meteorologists
2113	Chemists
2114	Geologists and geophysicists
2121	Mathematicians and related professionals
2122	Statisticians
2131	Computer systems designers, analysts and programmers with PhD degree*
2139	Computing professionals not elsewhere classified
2211	Biologists, botanists, zoologists and related professionals
2212	Pharmacologists, pathologists and related professionals
2213	Agronomists and related professionals
2310	College, university and higher education teaching professionals
Synthetic occupations	
2131	Computer systems designers, analysts, and programmers without PhD degree*
2142	Civil engineers
2143	Electrical engineers
2144	Electronics and telecommunications engineers
2145	Mechanical engineers
2146	Chemical engineers
2147	Mining engineers, metallurgists, and related professionals
2148	Cartographers and surveyors
2149	Architects, engineers, and related professionals not elsewhere classified
3111	Chemical and physical science technicians
3112	Civil engineering technicians
3113	Electrical engineering technicians
3114	Electronics and telecommunications engineering technicians
3115	Mechanical engineering technicians
3116	Chemical engineering technicians
3117	Mining and metallurgical technicians
3118	Draughtspersons
3119	Physical and engineering science technicians not elsewhere classified
Symbolic occupations	
2141	Architects, town and traffic planners
2431	Archivists and curators
2451	Authors, journalists, and other writers
2452	Sculptors, painters, and related artists
2453	Composers, musicians, and singers
2454	Choreographers and dancers
2455	Film, stage, and related actors and directors
2456	Designer
3131	Photographers and image and sound recording equipment operators
3471	Decorators and commercial designers
3472	Radio, television, and other announcers
3473	Street, night-club and related musicians, singers, and dancers
3474	Clowns, magicians, acrobats, and related associate professionals
3476	Stage managers, prop masters, etc.

Annex 2: Descriptive statistics

	Variable	Observations	Mean	Std. Dev	Min	Max	1	2	3	4	5	6	7	8	9	10	11
1	Growth	1,034,734	0.049	0.403	-2.746	2.786	1.000										
2	Analytical (yes/no)	1,034,734	0.013	0.114	0.000	1.000	0.007	1.000									
3	Synthetic (yes/no)	1,034,734	0.137	0.344	0.000	1.000	0.013	0.154	1.000								
4	Symbolic (yes/no)	1,034,734	0.057	0.232	0.000	1.000	-0.006	0.094	0.089	1.000							
5	Analytical (share)	1,034,734	0.003	0.044	0.000	1.000	0.003	0.598	0.024	0.004	1.000						
6	Synthetic (share)	1,034,734	0.059	0.198	0.000	1.000	0.005	0.081	0.745	0.003	0.011	1.000					
7	Symbolic (share)	1,034,734	0.027	0.143	0.000	1.000	-0.014	0.002	-0.022	0.777	-0.007	-0.032	1.000				
8	Log turnover	1,034,734	15.409	1.448	11.513	25.412	-0.104	0.150	0.285	0.096	-0.015	-0.010	-0.086	1.000			
9	Cash-flow per total assets	1,034,734	0.000	0.024	-9.770	13.203	0.001	0.001	0.003	-0.001	0.001	0.001	0.000	0.002	1.000		
10	Capital investments per total assets	1,034,734	0.000	0.000	0.000	0.063	0.001	0.000	-0.001	0.000	0.000	0.000	0.000	0.000	0.004	1.000	
11	Share of employees w. Tertiary education	1,034,734	0.191	0.311	0.000	1.000	0.003	0.150	0.163	0.143	0.135	0.210	0.136	-0.065	0.001	0.000	1.00