Research Policy 48 (2019) 234–247

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Knowledge base combinations and firm growth

Markus Grillitsch[a,b], Torben Schubert[b,c], Martin Srholec[b,d,*]

[a] Department of Human Geography, Lund University, P.O. Box 117, SE-221 00 Lund, Sweden
[b] Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE), Lund University, P.O. Box 117, SE-221 00 Lund, Sweden
[c] Fraunhofer Institute for Systems and Innovation Research, Karlsruhe, Breslauer Str. 48, 76135 Karlsruhe, Germany
[d] CERGE-EI, a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Czech Academy of Sciences, Politických vězňů 7, P.O. Box 882, Prague 1, 111 21, Czech Republic

ARTICLE INFO

JEL classification:
C33
D22
O12
O32
O33

Keywords:
Knowledge
Innovation
Firm growth
Micro data
Sweden

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 capture firm innovation activities comprehensively. We contribute to overcoming this limitation by looking in the knowledge processes that drive variegated forms of innovation and aim thereby to establish a solid relationship with firm growth in more detail. Our arguments draw on the differentiated knowledge base approach, distinguishing between analytical, synthetic, and symbolic knowledge. We measure the three types of knowledge bases with detailed longitudinal linked-employer-employee micro-data from Sweden. Econometric findings based on a very large sample of small and medium-sized firms indicate significantly positive effects of the three knowledge types, and in particular combinations thereof, on firm growth. In addition, we show that not only high-growth but also slow-growth firms benefit immensely from the use of combinatory knowledge bases. We find evidence on a curvilinear relation between knowledge bases and growth of firms. Beyond certain thresholds increasing the knowledge bases further results in decreasing firm growth. Our results remain robust in a wide range of specifications and econometric models.

1. Introduction

Despite great efforts to understand the drivers of firm growth results have often remained contradictory and ambiguous. Early studies on this topic have been consistent with the classical Gibrat's law (1931), which states that firm growth is essentially random (Sutton, 1997). More recent evidence, however, 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 fast-growing firms. Likewise Lee (2010), Demirel and Mazzucato, (2012), Deschryvere (2014) and Triguero et al. (2014) followed suit by showing that, although innovativeness is not related to the growth of most firms, there are subpopulations of firms for which the positive link holds.

While there may be positive effects for some firms, this 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. Hence, we propose to consider broader sources of knowledge and their combinations, which drive different forms of innovation and thereby firm growth.

The argument on the importance of different knowledge types for innovation can be traced back to Schumpeter (1911) and has echoed as an important topic since (e.g. Fleming, 2001; Jensen et al., 2007). In this paper, we use the knowledge base approach, which makes a distinction between three types of knowledge: i) analytical knowledge represents the traditional science-based modes of innovation; ii) synthetic knowledge is more tacit, experiential and applied to concrete problem solving; while iii) symbolic knowledge creates meaning, aesthetic value, brands, and design (Asheim and Coenen, 2005; Asheim, 2007; Asheim et al., 2007). A major advantage of this approach is that it explicitly establishes the connection between the combinations of different knowledge types, learning processes and innovation outputs of

* Corresponding author at: CERGE-EI, a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Czech Academy of Sciences, Politických vězňů 7, P.O. Box 882, Prague 1, 111 21, Czech Republic.
E-mail address: martin.srholec@cerge-ei.cz (M. Srholec).

https://doi.org/10.1016/j.respol.2018.08.009
Received 21 April 2016; Received in revised form 13 August 2018; Accepted 15 August 2018
Available online 05 September 2018

firms, thus goes far beyond the traditional focus on analytical knowledge that takes only partly into account synthetic knowledge and even less symbolic knowledge.

Linked to the knowledge combination argument is the literature on the effects of knowledge diversity on firm performance (e.g. Subramanian et al., 2016; Mohammadi et al., 2017; Østergaard et al., 2011; Faems and Subramanian, 2013). While these studies have contributed to showing the great importance of having access to broad knowledge bases, we emphasize that the diversity argument differs from the combination argument. Diversity increases as the frequency distribution of the knowledge bases becomes more even, while the combinations argument proposes that the three knowledge bases are complementary and often required in combination to reap their full potential. However, this does not necessarily mean that they need to be evenly distributed. For example, adding a couple of designers with symbolic knowledge to a large pool of engineers with synthetic knowledge might already do the trick. In an analogy, cooking recipes usually do not recommend using the same amount of meat, veggies and spices, but only a pinch of spices is enough to round off the meal.

Using detailed and extensive micro-data 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 for small and medium-sized firms (SMEs) show that the relationship between these innovation-relevant knowledge bases and firm growth is strong and robust and that this holds not just for high growth firms. Furthermore, the capability of the 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 growth, which underlines the need for combinations.

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 (and other factors). From this follows that growth in absolute terms is proportionate to absolute size, because of which the Gibrat's law is sometimes referred to as rule of proportionate growth. Empirical tests of Gibrat's law can be designed in two ways. First, because Gibrat's law gives rise to a log-normal size distribution (Eeckhout, 2004), it is possible to test whether observed distributions are indeed log-normal. Second, using data on the level of the individual firms it is possible to test whether the observed growth rates depend on size or other factors, which, if Gibrat's law holds, they should not.

The empirical evidence shows a mixed picture (Sutton, 1997 and Caves, 1998). 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. On the contrary, studies using data on a broader spectrum of firms revealed that there are systematic patterns in the growth of firms. Lotti et al. (2003) and Calvo (2006) argued that smaller firms need to grow faster than larger ones in order to survive, which is in direct contrast to Gibrat's law. To reconcile the contradictory findings, Lotti et al. (2009) provided evidence that Gibrat's law fails when both firms surviving in the future and those that do not are considered but may still hold in the long-run, when all inefficient firms have been driven out of the market. The argument is that once all firms in a cohort have reached their minimum efficient size and once the market forces have eliminated major differences in efficiency, Gibrat's law might still hold for the surviving firms. While this argument is convincing for making growth and size independent, it may not be relevant for other variables. In particular, while market forces tend to eliminate firm heterogeneity, innovation activities tend to recreate it.

Evolutionary economists have long 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, 1986). Second, innovation implies organizational learning (Phillips, 1971), which will strengthen dynamic capabilities (Teece et al., 1997) and generate unique knowledge (Grant, 1996), which is hard to imitate (Barnes, 1991). In this respect, innovation creates growth-inducing asymmetries both on the level of product and service characteristics as well as in the abilities of firms to create future innovations. Moreover, 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). Thus, while Lotti et al. (2009) made a convincing point that Gibrat's law may hold in the long-run with respect to size, the asymmetry creating effect of innovation suggests that the arguments in favor of Gibrat's law are much less compelling.

Despite the 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; Calvo, 2006). Others found a non-significant or even negative effect (e.g. Almus and Nerlinger, 1999; Freel and Robson, 2004; Lööf and Heshmati, 2006; Corsino and Gabriele, 2010). More recent studies showed that the positive link is highly conditional on other firm-level characteristics (McKelvie et al., 2017), including patenting (Demirel and Mazzucato, 2012), the persistence of innovation (Deschryvere, 2014; Triguero et al., 2014), or the use of internal vs. external R&D (Segarra and Teruel, 2014), and whether firms introduce product or process innovation (Santi and Santoleri, 2017). There is also evidence that these factors not only change over time but also condition each other dynamically (Bogliacino et al., 2017). Hence, the results suggested that the relationship depends on detailed characteristics of the firms' innovation behavior. By relying on data for R&D, patent and technological innovation counts, however, this literature has largely ignored the variegated forms of innovation and underlying knowledge processes.

2.2. 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. Hollenstein (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 (Asheim and Coenen, 2005; Asheim, 2007) is well suited for this purpose. It rests on the argument that innovation outputs ultimately relate 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 (Herstad et al., 2014; Aslesen and Freel, 2012; Pina and Tether, 2016). 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, problem solving 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 places. Nevertheless, university training in specific fields such as arts and design are crucial for symbolic knowledge bases.

Although there are strong theoretical grounds for the knowledge base approach, and substantial empirical support for the relationship between knowledge bases and innovation activities of firms (for a recent account see Asheim et al., 2017), there are no studies investigating the link between knowledge base combinations and firm growth.

2.3. The hypotheses

All three types of knowledge bases sustain innovation processes and in turn also competitive advantage of firms. Grillitsch and Asheim (2016), 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) considerably contribute to firm performance (Bloch, 1995; Gemser and Leenders, 2001; Hertenstein et al., 2005). Following the knowledge base approach, therefore, we expect that all three knowledge bases are relevant in their own right, which leads to our baseline hypothesis:

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

Several authors 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; Yang et al., 2017). Brown and Duguid (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 during an innovation project sequentially.

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 from 62 case studies in 22 European regions, they argue that in particular radical innovation processes increasingly require the latter. Manniche (2012) points out that the different knowledge bases, although individually identifiable, 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. (2017) 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 impact of knowledge on growth, this relationship is likely to be heterogeneous. Although resources devoted to knowledge generation and innovation may increase the probability of superior performance, they may also have a sting in the tail. If innovation efforts are not successful, the incurred costs will outweigh the benefits for the firm. Thus, the returns to innovation distribution is highly skewed. Many scholars have argued that for example high-growth firms differ fundamentally from more ordinary firms (Coad et al., 2016a,b). Supporting this argument, Coad and Rao (2008) show that for the fastest growing firms the impact of innovation as measured by R&D and patents is strong, while for the slowest growing firms innovation even has a negative effect. Firms only achieve modest growth on average that may have little to do with innovation; however, their results indicate that highly innovative firms can succeed spectacularly, if they make a break through, while on the flipside firms whose innovation efforts fail to obtain valuable results tend to perform worse than those that make no attempt to innovate. 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 tend to be 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 (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 (Schubert, 2011), creating complementary assets (Teece, 1986) and costs for overcoming institutional tensions, which may occur when key employees are vested in established technologies (Schubert and Andersson, 2015). Furthermore, innovations often show a high level of associated risk (Eliasson, 1991; Kerr et al., 2014). In particular, when innovation comes in the form of winner-takes-all races for dominant designs (Utterback and Suarez, 1993) or for key patents (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 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 micro dataset provided by Statistics Sweden (SCB) that is compiled by merging firm-level information from structural business statistics and business registry with individual-level job occupation and education data. 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 individual-level 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, the dependent variable of firm growth is measured by the log-difference of turnover. The explanatory variables of main interest capture knowledge bases at the level of firms and are constructed based on the occupational data (Asheim and Hansen, 2009; Grillitsch et al., 2017). Each employee of a firm is registered in the data with one occupation, which allowed us to identify whether the employee holds analytical, synthetic or symbolic knowledge.[1] 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 included a qualified judgement about the type of knowledge used, the knowledge generation process and the knowledge outcomes associated with each occupation. Each individual is registered with only one occupation in the data, hence assigned to one knowledge base. More details on the classification procedure and the resulting grouping of occupations can be found in Annex 1.[2]

Table 1 shows the occurrence of these innovation-related knowledge bases (KB) in the longitudinal panel of firms over 2004–2011. Using the three categories implies eight mutually exclusive combinatory outcomes ranging from neither of the knowledge bases, only one of them, two of them simultaneously and all three at once being present in the firm. The dataset comprises 1,034,734 observations of 225,063 firms in an unbalanced panel.[3] The figures are presented for the whole data set and separately for SMEs with less than 250 employees and large firms with 250 or more employees.

Overall, almost every fifth observation has at least one employee with occupation classified into these knowledge bases (18.4%). Synthetic knowledge is the most common (13.7%), followed by symbolic knowledge (5.7%) and analytical knowledge base (1.3%). Knowledge base combinations are quite rare (2.2%), of which the most frequent are synthetic and symbolic (1.3%), followed by analytical and synthetic (0.5%), while the combination of analytical and symbolic is extremely rare (0.1%). All three knowledge bases are also combined quite sporadically (0.3%). As can be expected, observations without any of the innovation-related knowledge bases are far more prevalent among SMEs (82.0%) than large firms (8.3%). Consequently, the base category of firms with zeros across the board refers to a typical SME but a rather particular kind of a large firm, which is important to keep in mind in the regression analysis.

As controls, we include a generic quality of human capital variable based on 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 their capital endowment is taken into account by capital investments per total assets. 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:

$$log\left(\frac{turnover_{i,t}}{turnover_{i,t-1}}\right) = \alpha + KB_{i,t-1}\beta + x_{i,t-1}\gamma + county_{i,t}\,\lambda + industry_i\delta + \phi z_t + \mu_i + \varepsilon_{i,t} \tag{2}$$

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 ($KB_{i,t-1}$), other firm characteristics ($x_{i,t-1}$), county effects ($county_{i,t}$), industry effects ($industry_i$), 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, amongst other approaches, by Pooled Ordinary Least Squares (POLS) and Fixed Effects (FE). Even though regular Hausman tests indicated

[1] 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).

[2] As response to a reviewer comment, a robustness check was conducted with a more narrow definition of symbolic knowledge excluding the occupational groups 2454, 3473, and 3474. The results of the robustness check are fully compatible with the results presented below. The results are available upon request from the authors.

[3] 36.5% of the firms are observed over the whole period and 85.8% of the firms are present in at least two consecutive periods.

Table 1
Relative frequency of knowledge base (KB) combinations (in %).2004–2011.

	Number of observations			% of total		
	SMEs	Large firms	All	SMEs	Large firms	All
None of the three KBs	843588	500	844088	82.0	8.3	81.6
Analytical only	4752	34	4786	0.5	0.6	0.5
Synthetic only	119136	1963	121099	11.6	32.5	11.7
Symbolic only	42686	214	42900	4.2	3.5	4.2
Analytical & synthetic	5017	545	5562	0.5	9.0	0.5
Analytical & symbolic	741	25	766	0.1	0.4	0.1
Synthetic & symbolic	11411	1520	12931	1.1	25.2	1.3
All three KBs	1362	1240	2602	0.1	20.5	0.3
Total	1028693	6041	1034734	100.0	100.0	100.0

the failure of the zero correlation assumption implying that the FE model is the appropriate estimator, we report also the POLS results as a reference. In any case, the results of the key variables of interest do not differ much by the estimator.

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

$$log\left(\frac{turnover_{i,t}}{turnover_{i,t-1}}\right) = \alpha + \eta\left(\frac{turnover_{i,t-1}}{turnover_{i,t-2}}\right) + KB_{i,t-1}\beta + x_{i,t-1}\gamma + county_{i,t}\lambda + industry_i\delta + \phi z_t + \mu_i + \varepsilon_{i,t} \quad (3)$$

which implies a dynamic panel data model. Such models can be estimated by a variety of procedures, which rely on the construction of time-based instrumental variables from lagged observations. We rely on the Arellano-Bover (AB) estimator (Arellano and Bover, 1995) because of its efficiency properties. Like regular IV-based models, instrument validity is an issue in dynamic panel data models.[4] The inspection of the autoregressive structure of the error terms does not indicate problems, while the Sargan overidentification test rejects the null of exogenous instruments. One way to assess the results based on AB models is to check the size of the coefficients. Roodman (2009) highlights that reasonable estimates of the effect of the lagged dependent variable should lie between the POLS and FE estimates, which is the case in all our regressions. Thus, although the specification tests are ambiguous, the results appear to lie within a reasonable range.[5]

In addition, H3 allows the estimated effects to differ across the growth distribution. To accommodate for this generalization we use quantile regression (Koenker and Hallock, 2001). Quantile regressions go beyond the analysis of the effects for an average firms (POLS, FE, and AB) and allow assessing the effect of knowledge base combinations for high growth firms (upper quantiles) or low growth firms (lower quantiles). Quantile regressions have been applied to investigate the relationship between innovation and R&D on firm growth (e.g. Coad and Rao, 2008; Mazzucato and Parris, 2015; Coad et al., 2016a,b). 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

The regression results are reported using the sample of SMEs, excluding large firms, for four main reasons. First, the full sample is dominated by SMEs (more than 99% of total observations), which drive the results. Second, as already vindicated above (Table 1), the vast majority of large firms maintain at least some of the innovation-related knowledge base occupations and in turn those few that do not represent a poor comparison group for deriving the inferences. Third, large firms rely less on knowledge rooted in individuals and encode more of their knowledge into depersonalized organizational routines (Clercq et al., 2012), which we do not observe. Last but not least, then not surprisingly, the coefficients of knowledge base combinations are far less precisely estimated for large firms than SMEs.[6] From now on in this section, hence, "firms" refer specifically to the SMEs.

Table 2 presents the benchmark results using the POLS, FE and AB estimators in the respective columns. Although their magnitude somewhat differs depending on the underlying assumptions, the knowledge base coefficients are highly statistically significant regardless of the estimator. All three knowledge bases are conducive to firm growth, which corroborates our baseline hypothesis (H1). Moreover, the effect of the 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 25 percentage points and those with two-way combinations by 16–19 percentage points faster than the base category of firms without the innovation-relevant knowledge bases. Indeed, this is a substantial boost given the fact that the average growth rate is around 5% only.

Table 3 presents the effects estimated at different quantiles of firm growth rates. Beside results for the 10%, 25%, 50%, 75% and 90% quantiles, which are customarily reported (e.g. Coad and Rao, 2008), the enormous size of the sample in hand also allows us to obtain results for the extreme of 1% and 99% quantiles, for which there is sufficient data to derive reliable estimates (1% corresponds to 10,287 observations). For the sake of facilitating the comparison, Fig. 1 visually depicts the estimated coefficients of the variables for knowledge base 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. [7]

On one hand, 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 75% quantile and triple for the 90% quantile as compared to the median. If the 99% quantile of the fastest growing firms is considered, the jump becomes even more pronounced. All else equal, if all three knowledge bases are present these top growth performers are expected to record as much as 53 percentage points higher growth as compared to the base category, whereas otherwise same firms growing at the median rate are estimated to benefit only from 5 percentage points growth premium. Hence, we observe about tenfold increase. Interestingly, however, for the fastest growing firms, symbolic knowledge alone does not seem to make much difference, which further

[4] Two approaches have been used to assert validity. The first more informal rests on the result that in valid models the AR(1) term of the residuals is zero while the AR(2) should be significantly different from zero. That can be easily tested by an autoregressive model on the error terms. The second test is based on the fact that dynamic panel data models are overidentified, often strongly.

[5] We also performed a number of robustness checks such as collapsing the list of instruments but did not discover noteworthy changes in the main results.

[6] In the sample of large firms, the knowledge base coefficients do not come out statistically significant at the conventional levels, except only of a few cases in the basic POLS estimates, and magnitude of the estimated effects is noticeably lower, especially for the combinations. Results of estimates for large firms are available upon request from the authors.

[7] Although it is recommended to use bootstrapped standard errors in quantile regressions because the asymptotic variance of the estimators depends on the density standard error (Hahn, 1995), we rely on analytical standard errors. The first reason is that regular bootstrapping methods (e.g. based on drawing pairs) do not provide an efficiency gain. Both analytical and bootstrapping methods converge at the same rate, unless more complex bootstrapping methods (e.g. studentized bootstraps) are used. Second, because of the high number of observations our quantile regressions took several hours to compute. Implementing a bootstrap with conventional numbers of replications (500-2000) would therefore led to a prohibitive computation time.

Table 2
Regressions of knowledge base combinations on firm growth (SMEs).

	(1) POLS	(2) FE	(3) POLS	(4) FE	(5) AB
KB: analytical only	0.0459***	0.0367***	0.0358***	0.0279***	0.1078***
	(0.0066)	(0.0075)	(0.0068)	(0.0084)	(0.0113)
KB: synthetic only	0.0628***	0.0545***	0.0416***	0.0476***	0.0975***
	(0.0016)	(0.0023)	(0.0016)	(0.0025)	(0.0032)
KB: symbolic only	0.0096***	0.0442***	0.0074***	0.0353***	0.0785***
	(0.0024)	(0.0035)	(0.0025)	(0.0039)	(0.0049)
KB: analytical & synthetic	0.1453***	0.1177***	0.1078***	0.0893***	0.1940***
	(0.0061)	(0.0075)	(0.0063)	(0.0082)	(0.0095)
KB: analytical & symbolic	0.0886***	0.0940***	0.0735***	0.0664***	0.1620***
	(0.0156)	(0.0154)	(0.0155)	(0.0162)	(0.0194)
KB: synthetic & symbolic	0.1280***	0.1220***	0.0873***	0.1056***	0.1802***
	(0.0039)	(0.0053)	(0.0039)	(0.0058)	(0.0066)
KB: all three	0.1892***	0.1902***	0.1240***	0.1649***	0.2464***
	(0.0095)	(0.0130)	(0.0097)	(0.0136)	(0.0156)
Log turnover	−0.0421***	−0.5955***	−0.0172***	−0.6082***	−0.0634***
	(0.0004)	(0.0010)	(0.0004)	(0.0014)	(0.0004)
Cash flow per total assets	0.0124	0.0346**	−0.0130	−0.0138	−0.0049
	(0.0351)	(0.0171)	(0.0421)	(0.0204)	(0.0215)
Capital investments per total assets	5.4323***	3.7004	4.3187***	0.8407	3.7998
	(1.3412)	(4.9487)	(0.6876)	(5.0514)	(5.7901)
Share of employees w. tertiary education	0.0067***	−0.0060*	0.0061***	−0.0041	−0.0430***
	(0.0018)	(0.0034)	(0.0018)	(0.0041)	(0.0016)
Growth (t-1)	..	..	−0.1060***	−0.0224***	−0.0959***
			(0.0020)	(0.0012)	(0.0013)
Constant	0.7243***	9.1802***	0.3731***	9.4886***	1.0016***
	(0.0063)	(0.0158)	(0.0060)	(0.0237)	(0.0055)
County dummies	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	..	Yes	..	..
Year dummies	Yes	Yes	Yes	Yes	Yes
Observations	1028693	1028693	793470	793470	793470
Firms	224389	224389	190247	190247	190247
R2 / R2-within	0.029	0.337	0.034	0.308	
F / chi2	220***	11,374***	158***	7466***	64,273***
AB AR1 test	..	..	..	..	−288.21***
AB AR2 test	..	..	..	..	−0.87
Instruments	..	..	..	..	231
Sargan test	..	..	..	..	3408***

Note: Standard errors in parentheses; standard errors of POLS 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 fixed effects (FE) regressions, F-statistics are reported for OLS and FE regressions; Wald Chi2 for Arellano–Bond regressions.

underlines the need for combinations.

On the other hand, a similar pattern also appears at the bottom of the growth distribution of firms leading to a symmetrical U-shape, except only of the rare combination of analytical and symbolic knowledge. The results therefore indicate that not only high-growth but also low-growth firms benefit highly from knowledge bases and their combinations. This result is surprising and has not been documented in the literature so far. A tentative explanation lies in the role of satisficing rather than optimizing behaviour (Nelson and Winter 1982). Satisficing behaviour means that firms compare their current performance level against a pre-determined satisficing reference point (Schubert et al., 2018). If they fall short of this reference point, they start searching for alternatives. Looking for new business opportunities typically implies that firms start questioning established routines (Cope, 2003), become susceptible to knowledge held by key partners or key employees (Clercq et al., 2012) and thus strive for new knowledge bases.

Table 4 provides the analysis of potential curvilinear effects of knowledge base intensities on growth of firms. The first estimates gives the linear effects only (Columns 1 and 3), while the second estimate includes squared terms of the respective shares (Columns 2 and 4). Results of both FE and AB estimators are presented for comparison, but the main conclusions are fairly robust to the model specification. Fig. 2 gives graphical representations of the estimated relationships.

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 this threshold, however, leads to declining contribution to firm growth. Similarly, turning points at 46% and 43% are detected for synthetic and symbolic knowledge bases, respectively. The AB estimates indicate slightly higher turning points, but the same shape of the relationships. 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 by tapping into other types of knowledge, especially when reaching the turning points.

The robustness of the curvilinear relationship is tested by splitting the sample at the respective turning points for analytical, synthetic and symbolic knowledge. As illustrated by the graphs, the relationships should be positive for firms with knowledge base shares below the turning point and negative for firms with knowledge base shares above the turning point. Table 5 shows that this pattern holds firmly for all three knowledge bases. Furthermore, we report the number of observations and firms for the split samples, which confirm that there is sufficient data to estimate the relationships above the turning points[8].

As regards the control variables, firm size has a significant negative effect on firm growth, which contradicts Gibrat's law (1931) but is in

[8] Furthermore, we run the models with cubic terms, which also resulted in inverse U-shaped relationships. The results can be obtained upon request from the authors.

Table 3
Quantile regressions of knowledge base combinations on firm growth (SMEs).

	(1) Q(0.01)	(2) Q(0.1)	(3) Q(0.25)	(4) Q(0.5)	(5) Q(0.75)	(6) Q(0.9)	(7) Q(0.99)
KB: analytical only	0.1742***	0.0135	0.0123**	0.0191***	0.0417***	0.0658***	0.2017***
	(0.0455)	(0.0172)	(0.0050)	(0.0046)	(0.0057)	(0.0117)	(0.0283)
KB: synthetic only	0.1583***	0.0242***	0.0133***	0.0210***	0.0471***	0.0991***	0.2646***
	(0.0178)	(0.0031)	(0.0014)	(0.0009)	(0.0015)	(0.0030)	(0.0133)
KB: symbolic only	0.1453***	0.0121**	−0.0056**	0.0008	0.0068***	0.0085	−0.0172
	(0.0239)	(0.0052)	(0.0023)	(0.0016)	(0.0021)	(0.0055)	(0.0195)
KB: analytical & synthetic	0.3231***	0.0951***	0.0530***	0.0543***	0.0904***	0.1636***	0.4239***
	(0.0594)	(0.0118)	(0.0050)	(0.0037)	(0.0063)	(0.0161)	(0.0310)
KB: analytical & symbolic	0.0369	0.1190***	0.0462***	0.0226***	0.0314***	0.0573*	0.2850***
	(0.1267)	(0.0318)	(0.0096)	(0.0053)	(0.0092)	(0.0295)	(0.0875)
KB: synthetic & symbolic	0.4230***	0.1122***	0.0538***	0.0402***	0.0628***	0.1238***	0.4003***
	(0.0643)	(0.0078)	(0.0029)	(0.0022)	(0.0037)	(0.0073)	(0.0359)
KB: all three	0.6539***	0.1700***	0.0755***	0.0493***	0.0768***	0.1750***	0.5319***
	(0.1170)	(0.0129)	(0.0046)	(0.0068)	(0.0077)	(0.0172)	(0.0489)
Log turnover	−0.0933***	0.0008	0.0013***	−0.0078***	−0.0321***	−0.0801***	−0.2178***
	(0.0041)	(0.0008)	(0.0003)	(0.0002)	(0.0003)	(0.0006)	(0.0022)
Cash flow per total assets	0.0935***	0.0505***	0.0111*	−0.0206	0.0257***	0.0339***	0.0189***
	(0.0064)	(0.0016)	(0.0061)	(0.1351)	(0.0022)	(0.0124)	(0.0036)
Capital investments per total assets	19.1701***	9.0302***	6.4805***	4.9508***	3.8805***	−1.3406***	−16.1064***
	(0.5666)	(0.1996)	(0.1475)	(0.4547)	(0.2786)	(0.1200)	(0.4373)
Share of employees w. tertiary education	−0.1624***	−0.0583***	−0.0176***	0.0066***	0.0397***	0.0893***	0.1977***
	(0.0182)	(0.0035)	(0.0015)	(0.0010)	(0.0018)	(0.0042)	(0.0181)
Constant	0.5769***	−0.2081***	−0.0698***	0.1802***	0.6909***	1.6021***	4.3291***
	(0.0655)	(0.0123)	(0.0048)	(0.0033)	(0.0053)	(0.0110)	(0.0419)
County dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1028693	1028693	1028693	1028693	1028693	1028693	1028693
Firms	224389	224389	224389	224389	224389	224389	224389
R^2	0.010	0.004	0.006	0.020	0.017	0.015	0.014

Note: Standard errors in parentheses; standard errors of quantile regressions are clustered at the level of the firm; ***, **, * indicate significance at the 10%, 5%, and 1% levels.

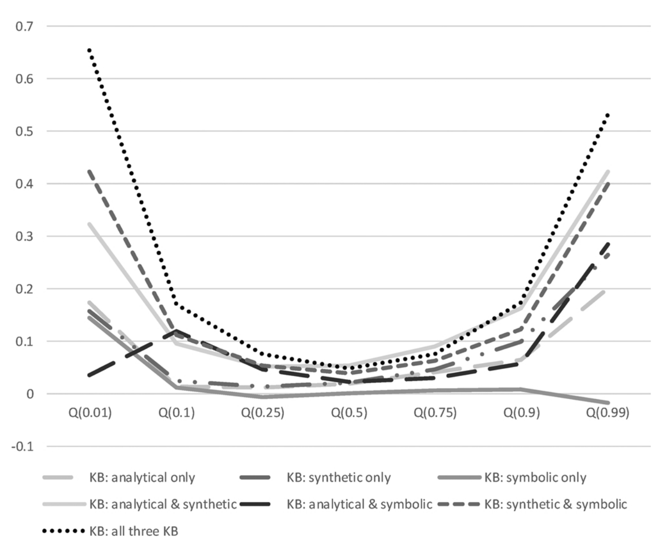

Fig. 1. Expected effect of knowledge base combinations on firm growth at different quantiles of the firm growth distributions (SMEs).

line with the arguments of Lotti et al. (2003) and Calvo (2006). Interestingly, the effect of the share of employees with tertiary education varies by model and is even significantly negative in some cases. The inclusion of this variable allows separating a specific effect of human capital as captured by knowledge bases from a generic effect of the quality of human capital as measured by education attainment. The results indicate that knowledge bases have a more robust and stronger effect on firm growth than generic human capital and that the effect of the latter depends heavily on firm characteristics. More specifically, generic human capital has a negative effect for low growth firms but a strong positive effect for high growth firms (Table 3), which is similar to the pattern that Coad and Rao (2008) observed for the relationship between innovativeness and firm growth. Financial resources measured as cash flow per total assets tend to affect firm growth positively as the coefficient is either significant positive or insignificant. Somewhat surprising are the results for capital investments per total assets, as the observed effects tend to be positive except for the fastest growing firms (Table 3, Columns 6 and 7). One possible explanation is that in mature (and low-growth) industries, competition often rests on process innovations that are achieved for instance by investments in machinery (Cooke, 1995). Conversely, competition in growing industries is often driven by product innovations, which rest less on capital investments but largely on knowledge bases. The results could even indicate that high-growth firms that shift their focus from knowledge base-driven, more radical product innovations towards capital-driven, less radical process innovations may lose competitive advantage.

5. Conclusions

This paper re-examined the long-standing question of how knowledge, innovation and growth relate to each other at the firm level. Despite strong theoretical support of a positive relationship, empirical research on this topic has not provided robust evidence. This may have to do with the most commonly used innovation indicators, which tend to be rather limited. In contrast, we approach this question from the root by focussing on broad types of knowledge, and combinations thereof, which are relevant for generating various forms of innovation. For this purpose, we draw on the differentiated knowledge base approach, which directly looks into what knowledge firms actually use, when they innovate, instead of relying on abstract figures such as R&D expenditures or patents. Thus, the relevant types of knowledge driving different forms of innovation are captured in a much broader way than

Table 4
FE and AB regressions of shares of knowledge bases on firm growth (SMEs).

	(1) FE	(2) FE	(3) AB	(4) AB
Share KB: analytical	−0.0232	0.2558***	0.1481***	0.4908***
	(0.0149)	(0.0402)	(0.0281)	(0.0577)
Share KB: synthetic	0.0053	0.2356***	0.0560***	0.4069***
	(0.0047)	(0.0123)	(0.0081)	(0.0170)
Share KB: symbolic	−0.0103	0.2076***	0.0071	0.3532***
	(0.0070)	(0.0182)	(0.0111)	(0.0253)
Share KB: analytical square	..	−0.3378***	..	−0.4448***
		(0.0444)		(0.0669)
Share KB: synthetic square	..	−0.2581***	..	−0.4013***
		(0.0128)		(0.0183)
Share KB: symbolic square	..	−0.2431***	..	−0.3858***
		(0.0187)		(0.0263)
Log turnover	−0.5931***	−0.5947***	−0.0550***	−0.0592***
	(0.0010)	(0.0010)	(0.0003)	(0.0003)
Cash flow per total assets	0.0347**	0.0346**	−0.0044	−0.0049
	(0.0171)	(0.0171)	(0.0216)	(0.0216)
Capital investments per total assets	3.6815	3.6752	3.6882	3.6929
	(4.9521)	(4.9500)	(5.8024)	(5.7967)
Share of employees w. tertiary education	−0.0031	−0.0041	−0.0211***	−0.0291***
	(0.0034)	(0.0034)	(0.0019)	(0.0019)
Growth (t-1)	..	..	−0.0973***	−0.0967***
			(0.0013)	(0.0013)
Constant	9.1523***	9.1735***	0.8868***	0.9473***
	(0.0158)	(0.0158)	(0.0051)	(0.0056)
County dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Observations	1028693	1028693	793470	793470
Firms	224389	224389	190247	190247
R^2-within	0.336	0.337		
F / chi2	12744***	11,681***	62,621***	6342***
AB AR1 test	..	..	−288.80***	−288.38***
AB AR2 test	..	..	−0.89	−0.95
Instruments	..	..	127	205
Sargan test	..	..	3266***	3449***
Analytical: Turning point	..	38%	..	55%
Synthetic: Turning point	..	46%	..	51%
Symbolic: Turning point	..	43%	..	46%

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.

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 of SMEs 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 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; Eisenman, 2013) which showed that they considerably contribute to firm performance (Bloch, 1995; Gemser and Leenders, 2001; Hertenstein et al., 2005). Even more importantly, the results confirm that combinations of knowledge bases have by far the strongest effect on firm growth.

The findings resonate well with studies showing that the most dynamic firms combine different types of innovation and knowledge (Jensen et al., 2007; Tödtling and Grillitsch, 2015; Grillitsch et al., 2017). Jensen et al. (2007), for instance, found that firms are most innovative if they combine science and technology driven innovations with innovations based on learning through doing, using, and interacting. This also resonates with the strategic innovation literature that underscores the need of knowledge recombination for innovation (Fleming, 2001; Rosenkopf and Nerkar, 2001; Neuhäusler et al., 2016). Since most of this literature relied on patent data in order to measure knowledge, the analyses had 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.

Moreover, not surprisingly, firms at the upper part of the growth distribution appear to benefit from a stronger link between the knowledge bases and growth. Coad and Rao (2008) predicted this effect because of the interaction of costs and risks of innovation projects. However, it could also result from strategic differences. High-growth firms perform better because of unique products with considerable consumer value. Sustaining this competitive advantage usually requires innovation activities. Contrary to our expectations, we found that also low-growth firms benefit strongly from knowledge bases. We interpret this result by satisficing behaviour in the sense that underperforming firms are under stronger pressure to pursue new business opportunities. Nevertheless, further analysis is required to underpin this

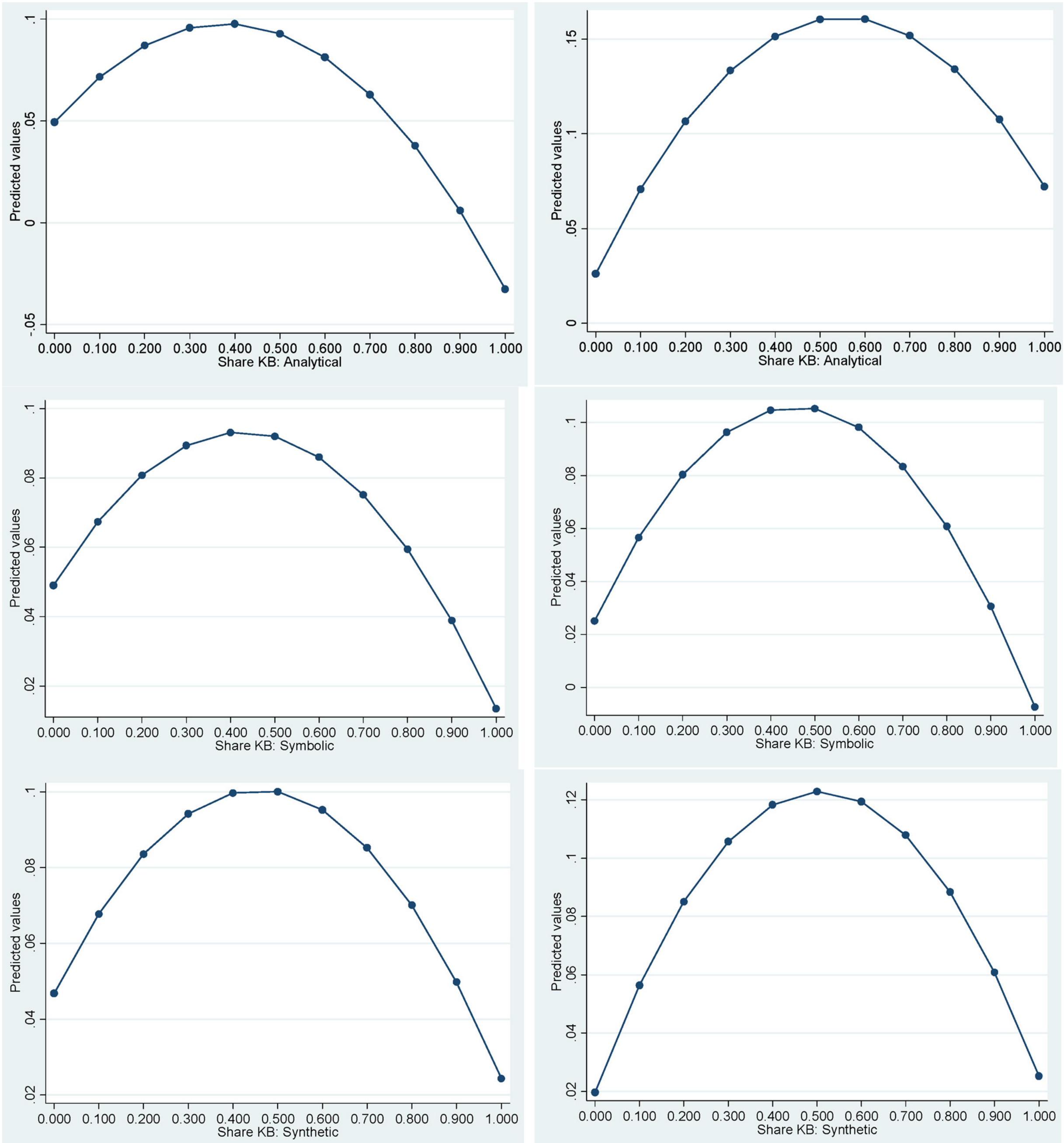

Fig. 2. Estimated relationship between the share of the knowledge bases and firm growth (FE left panels; AB right panels) (SMEs).

interpretation, for instance, by taking into account different types of low growth firms and their exit rates. Finally, we find evidence that it does not pay off simply to increase the relative importance of a specific knowledge base without limits but that there are turning points above which a further specialisation in a specific knowledge base becomes detrimental.

Less surprising but still important is that the results are inconclusive for large firms. This calls for further research to verify empirically whether, to what extent and through which processes knowledge base combinations matter for large firms. A key question in this regard is if knowledge base combinations materialise in large firms also through combinations of knowledge held by individual employees or if learning and innovation processes are embedded to a larger extent in organisational routines, which are not captured by our approach. For instance, large firms may have strong symbolic knowledge in the marketing department, advanced synthetic knowledge in production facilities, and analytical knowledge in the research and development unit. However, the barriers for integrating the types of knowledge between organisational units of large firms are arguably much higher than if individuals with different types of knowledge work directly together in a small firm.

It would be useful to deepen work on singling out which types of knowledge drive innovation and growth. The typology of the differentiated knowledge base approach has proven to be a good starting point in this respect; however, 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 the firm capabilities to integrate different types of knowledge and complementary assets to turn innovation into growth. It would be also interesting to investigate further how firms combine firm-internal and firm-external knowledge bases, which we could not do because the registry-type of data in hand did not provide evidence on the latter. Finally, it would be fruitful to

Table 5
Robustness checks of curvilinear relationships (SMEs).

	(1) FE Below turning point	(2) FE Above turning point	(3) AB Below turning point	(4) AB Above turning point
Share KB: analytical	0.2303***	−0.2205***	0.1383***	−0.4649*
	(0.0388)	(0.0722)	(0.0317)	(0.2633)
Observations	1025767	2926	792237	1233
Firms	223979	1183	190037	526
Share KB: synthetic	0.1952***	−0.1114***	0.1891***	−0.2762***
	(0.0123)	(0.0140)	(0.0126)	(0.0701)
Observations	968510	60183	759310	34160
Firms	214096	18584	183779	11569
Share KB: symbolic	0.1758***	−0.1796***	0.0951***	−0.0900
	(0.0192)	(0.0192)	(0.0193)	(0.0559)
Observations	998846	29847	771340	22130
Firms	219234	9173	185872	7270

Note: Standard errors in parentheses; ***, **, * indicate significance at the 10%, 5%, and 1% levels; control variables are included in the models but not reported to save space.

examine in future research whether the relationship between knowledge bases and firm's growth is mediated by third factors, such as size, age or sectoral differences.

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" and the Czech Science Foundation (GAČR) project (17-09265S) on "Frontiers of empirical research on public financing of business R&D and innovation" is gratefully acknowledged. We are grateful for the comments and suggestions by the editor and two anonymous reviewers. All the usual caveats apply.

Annex 1 Classification of occupations by knowledge bases

Following Grillitsch et al. (2017) 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:

Table A1
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.

- 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 4–9 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, 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 A1 lists the occupations by knowledge base.

Annex 2 Descriptive statistics (SMEs)

	Variables	Observations	Mean	Std. Dev	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12
1	Growth	1028693	0.050	0.404	−2.746	2.786	1.000											
2	KB: Analytical (yes/no)	1028693	0.012	0.107	0.000	1.000	0.008	1.000										
3	KB: Synthetic (yes/no)	1028693	0.133	0.340	0.000	1.000	0.013	0.129	1.000									
4	KB: Symbolic (yes/no)	1028693	0.055	0.227	0.000	1.000	−0.006	0.058	0.067	1.000								
5	KB: Analytical (share)	1028693	0.003	0.044	0.000	1.000	0.003	0.639	0.024	0.003	1.000							
6	KB: Synthetic (share)	1028693	0.059	0.198	0.000	1.000	0.005	0.071	0.753	−0.002	0.010	1.000						
7	KB: Symbolic (share)	1028693	0.027	0.143	0.000	1.000	−0.014	0.002	−0.021	0.796	−0.007	−0.031	1.000					
8	Turnover (mio. SEK)	1028693	19.60	160.2	0.100	35879	−0.012	0.066	0.118	0.058	0.004	0.010	−0.011	1.000				
9	Employees	1028693	8.000	18.00	1.000	249.0	0.009	0.157	0.299	0.118	0.003	0.031	−0.026	0.327	1.000			
10	Cashflow per total assets	1028693	0.000	0.024	−9.770	13.20	0.001	0.001	0.003	0.000	0.001	0.001	0.000	0.000	0.001	1.000		
11	Capital investments per total assets	1028693	0.000	0.000	0.000	0.063	0.001	0.000	−0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.004	1.000	
12	Share of employees w. tertiary education	1028693	0.190	0.312	0.000	1.000	0.003	0.152	0.164	0.143	0.134	0.209	0.136	0.015	0.001	0.001	0.000	1.000

References

Almus, M., Nerlinger, E.A., 1999. Growth of new technology-based firms: which factors matter? Small Bus. Econ. 13 (2), 141–154.
Arbetsförmedlingen [Swedish Public Employment Service], 2014. Yrken [Occupations] Ä-O. (downloaded on 1 June 2014 from http://www.arbetsformedlingen.se/For-arbetssokande/Yrke-och-framtid/Yrken-A-O.html).
Arellano, M., Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. J. Econom. 68 (1), 29–51.
Asheim, B.T., 2007. Differentiated knowledge bases and varieties of regional innovation systems. Innov. Eur. J. Soc. Sci. Res. 20 (3), 223–241.
Asheim, B.T., Coenen, L., 2005. Knowledge bases and regional innovation systems: Comparing Nordic clusters. Res. Policy 34 (8), 1173–1190.
Asheim, B.T., Hansen, H.K., 2009AsheimA. Knowledge bases, talents, and contexts: on the usefulness of the creative class approach in Sweden. Econ. Geogr. 85 (4), 425–442.
Asheim, B.T., Coenen, L., Vang, J., 2007. Face-to-face, buzz, and knowledge bases: sociospatial implications for learning, innovation, and innovation policy. Environ. Plann. C 25 (5), 655–670.
Asheim, B.T., Grillitsch, M., Trippl, M., 2017. Introduction: combinatorial knowledge bases, regional innovation, and development dynamics. Econ. Geogr. 93 (5), 429–435.
Aslesen, H.W., Freel, M., 2012. Industrial knowledge bases as drivers of open innovation? Ind. Innov. 19 (7), 563–584.
Audretsch, D.B., Coad, A., Segarra, A., 2014. Firm growth and innovation. Small Bus. Econ. 43 (4), 743–749.
Barnes, M., 1991. Innovation—why project management is essential to successful businesses. Int. J. Proj. Manag. 9 (4), 207–209.
Bloch, P.H., 1995. Seeking the ideal form: product design and consumer response. J. Mark. 59 (3), 16–29.
Bloom, N., Van Reenen, J., 2002. Patents, real options and firm performance. Econ. J. 112 (478), 97–116.
Bogliacino, F., Lucchese, M., Nascia, L., Pianta, M., 2017. Modeling the virtuous circle of innovation. A test on Italian firms. Ind. Corp. Chang. 26 (3), 467–484.
Brown, J.S., 2002. Research that reinvents the corporation. Harv. Bus. Rev. 80 (8), 105–114.
Brown, J.S., Duguid, P., 1991. Organizational learning and communities of practice: toward a unified view of working, learning, and innovation. Organizational Science 2 (1), 40–57.
Calvo, J.L., 2006. Testing Gibrat's law for small, Young and innovating firms. Small Bus. Econ. 26 (2), 117–123.
Caves, R.E., 1998. Industrial organization and new findings on the turnover and mobility of firms. J. Econ. Lit. 36 (4), 1947–1982.
Clercq, D., de, Sapienza, H.J., Yavuz, R.I., Zhou, L., 2012. Learning and knowledge in early internationalization research: past accomplishments and future directions. J. Bus. Ventur. 27 (1), 143–165.
Coad, A., 2007. Testing the principle of 'growth of the fitter': the relationship between profits and firm growth. Struct. Change Econ. Dyn. 18 (3), 370–386.
Coad, A., 2009. The Growth of Firms: A Survey of Theories and Empirical Evidence. Edward Elgar, Cheltenham.
Coad, A., Rao, R., 2008. Innovation and firm growth in high-tech sectors: a quantile regression approach. Res. Policy 37 (4), 633–648.
Coad, A., Cowling, M., Siepel, J., 2016a. Growth processes of high-growth firms as a four-dimensional chicken and egg. Ind. Corp. Chang. 26 (4), 537–554.
Coad, A., Segarra, A., Teruel, M., 2016b. Innovation and firm growth: does firm age play a role? Res. Policy 45 (2), 387–400.
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. Ind. Innov. 13 (4), 393–414.
Cooke, P. (Ed.), 1995. The Rise of the Rustbelt. UCL Press, London.
Cope, J., 2003. Entrepreneurial learning and critical reflection: discontinuous events as triggers for 'higher-level' learning. Manag. Learn. 34 (4), 429–450.
Corsino, M., Gabriele, R., 2010. Product innovation and firm growth: evidence from the integrated circuit industry. Ind. Corp. Chang. 20 (1), 29–56.
Creusen, M.E.H., Schoormans, J.P.L., 2005. The different roles of product appearance in consumer choice. J. Prod. Innov. Manage. 22 (1), 63–81.
Damanpour, F., Aravind, D., 2012. Manegerial innovation: conceptions, processes, and antecedents. Manag. Organ. Rev. 8 (2), 423–454.
Dasgupta, P., 1986. The theory of technological competition. In: In: Stiglitz, J.E., Mathewson, G.F. (Eds.), New Developments in the Analysis of Market Structure. International Economic Association Series, vol 77. Palgrave Macmillan, London, pp. 519–549.
de Jong, J.P.J., Marsili, O., 2006. The fruit flies of innovations: a taxonomy of innovative small firms. Res. Policy 35 (2), 213–229.
Del Monte, A., Papagni, E., 2003. R&D and the growth of firms: empirical analysis of a panel of Italian firms. Res. Policy 32 (6), 1003–1014.
Demirel, P., Mazzucato, M., 2012. Innovation and firm growth: Is R&D worth it? Ind. Innov. 19 (1), 45–62.
Deschryvere, M., 2014. R&D, firm growth and the role of innovation persistence: an analysis of Finnish SMEs and large firms. Small Bus. Econ. 43 (4), 767–785.
DiMasi, J.A., Hansen, R.W., Grabowski, H.G., 2003. The price of innovation: new estimates of drug development costs. Journal of health economics 22 (2) 151–185.
Dosi, G., 1988. Sources, procedures, and microeconomic effects of innovation. J. Econ. Lit. 26 (3), 1120–1171.
Eeckhout, J., 2004. Gibrat's law for (all) cities. Am. Econ. Rev. 94 (5), 1429–1451.
Eisenman, M., 2013. Understanding aesthetic innovation in the context of technological evolution. Acad. Manag. Rev. 38 (3), 332–351.
Eliasson, G., 1991. Modeling the experimentally organized economy. J. Econ. Behav. Organ. 16 (1-2), 153–182.
Faems, D., Subramanian, A.M., 2013. R&D manpower and technological performance: the impact of demographic and task-related diversity. Res. Policy 42 (9), 1624–1633.
Fleming, L., 2001. Recombinant uncertainty in technological search. Manage. Sci. 47 (1), 117–132.
Freel, M.S., Robson, P.J., 2004. Small firm innovation, growth and performance: evidence from Scotland and Northern England. Int. Small Bus. J. 22 (6), 561–575.
Frenz, M., Lambert, R., 2009. Exploring Non-technological and mixed modes of innovation across countries. In: OECD (Ed.), Innovation in Firms: A Microeconomic Perspective. OECD, Paris, pp. 69–110.
Fudenberg, D., Gilbert, R., Stiglitz, J., Tirole, J., 1983. Preemption, leapfrogging, and compeition in patent races. Eur. Econ. Rev. 22 (1), 3–21.
Gemser, G., Leenders, M.A.A.M., 2001. How integrating industrial design in the product development process impacts on company performance. J. Prod. Innov. Manage. 18 (1), 28–38.
Gera, S., Gu, W., 2004. The effect of organizational innovation and information technology on firm performance. Int. Prod. Monit. 9 (3), 37–51.
Geroski, P., Machin, S., 1992. Do innovating firms outperform non-innovators? Bus. Strategy Rev. 3 (2), 79–90.
Gibrat, R., 1931. Les Inégalités Économiques. Librairie du Recueil Sirey, Paris.
Grant, R.M., 1996. Toward a knowledge-based theory of the firm. Strateg. Manag. J. 17 (S2), 109–122.
Grillitsch, M., Asheim, B.T., 2016. Cluster policy: renewal through the integration of institutional variety. In: Fornahl, D., Hassink, R., Menzel, M.-P. (Eds.), Broadening Our Knowledge on Cluster Evolution. Routledge, Abingdon, pp. 76–94.
Grillitsch, M., Martin, R., Srholec, M., 2017. Knowledge base combinations and innovation performance in Swedish regions. Econ. Geogr. 93 (5), 458–479.
Grimpe, C., Kaiser, U., 2010. Balancing internal and external knowledge acquisition: the gains and pains from R&D outsourcing. J. Manag. Stud. 47 (8), 1483–1509.
Hahn, J., 1995. Bootstrapping quantile regression estimators. Econ. Theory 11 (1), 105–121.
Herstad, S.J., Aslesen, H.W., Ebersberger, B., 2014. On industrial knowledge bases, commercial opportunities and global innovation network linkages. Res. Policy 43 (3), 495–504.
Hertenstein, J.H., Platt, M.B., Veryzer, R.W., 2005. The impact of industrial design effectiveness on corporate financial performance. J. Prod. Innov. Manage. 22 (1), 3–21.
Hollenstein, H., 2003. Innovation modes in the Swiss service sector: a cluster analysis based on firm-level data. Res. Policy 32 (5), 845–863.
Jensen, M.B., Johnson, B., Lorenz, E., Lundvall, B.Å., 2007. Forms of knowledge and modes of innovation. Res. Policy 36 (5), 680–693.
Kerr, W.R., Nanda, R., Rhodes-Kropf, M., 2014. Entrepreneurship as experimentation. J. Econ. Perspect. 28 (3), 25–48.
Koenker, R., Hallock, K., 2001. Quantile regression: an introduction. J. Econ. Perspect. 15 (4), 43–56.
Le Bas, C., Mothe, C., Nguyen-Thi, T.U., 2015. The differential impacts of organizational innovation practices on technological innovation persistence. Eur. J. Innov. Manag. 18 (1), 110–127.
Lee, C.-Y., 2010. A theory of firm growth: learning capability, knowledge threshold, and patterns of growth. Res. Policy 39 (2), 278–289.
Leiponen, A., Drejer, I., 2007. What exactly are technological regimes? Intra-industry heterogeneity in the organization of innovation activities. Res. Policy 36 (8), 1221–1238.
Lööf, H., Heshmati, A., 2006. On the relationship between innovation and performance: a sensitivity analysis. Econ. Innov. New Technol. 15 (4-5), 317–344.
Lotti, F., Santarelli, E., Vivarelli, M., 2003. Does Gibrat's Law hold among young, small firms? J. Evol. Econ. 13 (3), 213–235.
Lotti, F., Santarelli, E., Vivarelli, M., 2009. Defending Gibrat's Law as a long-run regularity. Small Bus. Econ. 32 (1), 31–44.
Manniche, J., 2012. Combinatorial knowledge dynamics: on the usefulness of the differentiated knowledge bases model. Eur. Plan. Stud. 20 (11), 1823–1841.
Martin, R., Moodysson, J., 2011. Innovation in symbolic industries: the geography and organization of knowledge sourcing. Eur. Plan. Stud. 19 (7), 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. Eur. Urban Reg. Stud. 20 (2), 170–187.
Mazzucato, M., Parris, S., 2015. High-growth firms in changing competitive environments: the US pharmaceutical industry (1963 to 2002). Small Bus. Econ. 44 (1), 145–170.
McKelvie, A., Brattström, A., Wennberg, K., 2017. How young firms achieve growth: reconciling the roles of growth motivation and innovative activities. Small Bus. Econ. 49 (2), 273–293.
Mohammadi, A., Broström, A., Franzoni, C., 2017. Workforce composition and innovation: how diversity in employees' ethnic and educational backgrounds facilitates firm-level innovativeness. J. Prod. Innov. Manage. 34 (4), 406–426.
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. Environ. Plan. A 40 (5), 1040–1056.
Neuhäusler, P., Schubert, T., Frietsch, R., Blind, K., 2016. Managing portfolio risk in strategic technology management: evidence from a panel data-set of the world's largest R&D performers. Econ. Innov. New Technol. 25 (7), 651–667.
Østergaard, C.R., Timmermans, B., Kristinsson, K., 2011. Does a different view create something new? The effect of employee diversity on innovation. Res. Policy 40 (3), 500–509.
Phillips, A., 1971. Technology and Market Structure: a Study of the Aircraft Industry.

Heath Lexington Books, Lexington, Mass.
Pina, K., Tether, B.S., 2016. Towards understanding variety in knowledge intensive business services by distinguishing their knowledge bases. Res. Policy 45 (2), 401–413.
Roodman, D., 2009. How to do xtabond2: an introduction to difference and system GMM in Stata. Stata J. 9 (1), 86–136.
Rosenkopf, L., Nerkar, A., 2001. Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry. Strateg. Manag. J. 22 (4), 287–306.
Santi, C., Santoleri, P., 2017. Exploring the link between innovation and growth in Chilean firms. Small Bus. Econ. 49 (2), 445–467.
SCB, 1998. SSYK 96 standard för svensk yrkesklassificering 1996 [Standard for Swedish classification of occupations]. MIS Meddelanden i samordningsfrågor för Sveriges officiella statistik [Reports on statistical coordination for the official statistics of Sweden], Statistics Sweden 1998/3.
Schubert, T., 2010. Marketing and organisational innovations in entrepreneurial innovation processes and their relation to market structure and firm characteristics. Rev. Ind. Organ. 36 (2), 189–212.
Schubert, T., 2011. Assessing the value of patent portfolios: an international country comparison. Scientometrics 88 (3), 787–804.
Schubert, T., Andersson, A., 2015. Old is gold? The effects of employee age on innovation and the moderating effects of employment turnover. Econ. Innov. New Technol. 24 (1-2), 95–113.
Schubert, T., Baier, E., Rammer, C., 2018. Firm capabilities, technological dynamism and the internationalisation of innovation: a behavioural approach. J. Int. Bus. Stud. 49 (1), 70–95.
Schumpeter, J.A., 1911. Theorie der wirtschaftlichen Entwicklung. Duncker & Humbolt, Leipzig.
Segarra, A., Teruel, M., 2014. High-growth firms and innovation: an empirical analysis for Spanish firms. Small Bus. Econ. 43 (4), 805–821.
Srholec, M., Verspagen, B., 2012. The Voyage of the Beagle into innovation: explorations on heterogeneity, selection and sectors. Ind. Corp. Chang. 21 (5), 1221–1253.
Strambach, S., Klement, B., 2012. Cumulative and combinatorial micro-dynamics of knowledge: the role of space and place in knowledge integration. Eur. Plan. Stud. 20 (11), 1843–1866.
Subramanian, A.M., Choi, Y.R., Lee, S.-H., Hang, C.-C., 2016. Linking technological and educational level diversities to innovation performance. J. Technol. Transf. 41 (2), 182–204.
Sutton, J., 1997. Gibrat's legacy. J. Econ. Lit. 35 (1), 40–59.
Tavassoli, S., Karlsson, C., 2015. Persistence of various types of innovation analyzed and explained. Res. Policy 44 (10), 1887–1901.
Teece, D.J., 1986. Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. Res. Policy 15 (6), 285–305.
Teece, D.J., Pisano, G., Shuen, A., 1997. Dynamic capabilities and strategic management. Strateg. Manag. J. 18 (7), 509–533.
Tödtling, F., Grillitsch, M., 2015. Does Combinatorial Knowledge Lead to a Better Innovation Performance of Firms? Eur. Plan. Stud. 23 (9), 1741–1758.
Triguero, A., Córcoles, D., Cuerva, M.C., 2014. Persistence of innovation and firm's growth: evidence from a panel of SME and large Spanish manufacturing firms. Small Bus. Econ. 43 (4), 787–804.
Utterback, J.M., Suarez, F.F., 1993. Innovation, competition, and industry structure. Res. Policy 22 (1), 1–21.
Yang, C., Bossink, B., Peverelli, P., 2017. High-tech start-up firm survival originating from a combined use of internal resources. Small Bus. Econ. 49 (4), 799–824.
Yasuda, T., 2005. Firm growth, size, age and behavior in Japanese manufacturing. Small Bus. Econ. 24 (1), 1–15.