Available online at www.sciencedirect.com
SCIENCE DIRECT
Research Policy 34 (2005) 33–45
research policy
www.elsevier.com/locate/econbase

# How do firms' knowledge bases affect intra-industry heterogeneity?☆ An analysis of the Spanish pharmaceutical industry

Pablo D'Este*

*SPRU—Science and Technology Policy Research, Freeman Centre, University of Sussex, Falmer, Brighton BN1 9QE, UK*

Received 1 January 2004; accepted 26 October 2004
Available online 23 December 2004

**Abstract**

This paper aims to contribute to the analysis of within industry inter-firm variety. The paper develops two themes: (i) the analysis of intra-industry heterogeneity, and (ii) the extent to which higher performance is associated with the capacity of firms to expand their knowledge base. The main contribution of the paper is empirical, based on a data source consisting of information on documents published in international scientific journals by Spanish pharmaceutical firms. The empirical results support the argument that the firm's knowledge base is a main driver of persistent heterogeneity within industries. We find systematic variety in terms of how firms articulate their research activities, and positive correlation between firms' knowledge diversification and performance.


*Keywords:* Firm heterogeneity; Knowledge diversification; Bibliometric analysis; Spanish pharmaceutical industry

## 1. Introduction

This paper aims to contribute to the analysis of within industry inter-firm variety. Building upon the knowledge-based theory of the firm (Nelson and Winter, 1982; Penrose, 1959; Fransman, 1994), this paper develops two themes. First, the analysis of intra-industry heterogeneity: why do firms that operate in the same industry differ, and why are such differences persistent? Second, the paper investigates the extent to which higher performance is associated with the capacity of firms to expand their knowledge base (rather than with their initial conditions).

Based on a data source consisting of information on documents published in scientific international journals by Spanish pharmaceutical firms, we examine the

☆ This paper was prepared for the November 2003 Conference in honour of Keith Pavitt "What do we know about innovation", organised by SPRU and CENTRIM, where the paper was awarded the prize for Best Paper by an upcoming researcher in the theme: "The industrial dynamics of innovation and competition", sponsored by ICC.
* Tel.: +44 1273 877964; fax: +44 1273 685865.
*E-mail address:* P.D-Este-Cukierman@sussex.ac.uk.

0048-7333/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.respol.2004.10.007

profile of firms' knowledge base. The results of the empirical analysis support the argument that the firm's knowledge base is a main driver of persistent heterogeneity within industries. On the one hand, because of the systematic variety in terms of how firms articulate and organise their research activities; and on the other hand, because of the positive correlation between the firms' knowledge diversification and performance.

The paper is structured as follows. Section 2 reviews the literature and lays out the questions to be addressed. Section 3 describes the data sources. Section 4 identifies the firms we propose to analyse and describe their research performance. Section 5 provides empirical evidence on the breadth and profile of firms' knowledge base. Section 6 discusses the relationship between knowledge breadth and performance. Finally, Section 7 presents the conclusions.

## 2. Literature background

While a number of theories have been developed to address the origins and dynamic of competitive advantage, there is no generally accepted perspective (see Cockburn et al., 2000, for a review of these theories). By focusing on the ability of the firm to develop new capabilities, the knowledge-based theory of the firm provides a distinctive and fruitful framework to contribute to an empirically grounded theory of the origins of competitive advantage (Dosi et al., 2000). The purpose of this paper is to study the factors that shape the scope and direction of firm's technological capability accumulation and to examine the extent to which such factors are relevant drivers of persistent intra-industry heterogeneity—in terms of both behaviour and performance.

### *2.1. Variety and stability of firms' knowledge bases*

Several authors (Patel and Pavitt, 1997; Granstrand et al., 1997) have pointed out that, even when large firms are increasingly technologically diversified, their technological competencies profiles are similar within industries while differing significantly between industries. At the same time, empirical evidence suggests that firms within the same industry display distinct strategic paths (Noda and Collis, 2001; McGee and Thomas, 1986). These two streams of empirical evidence have led to the argument that there is not a one-to-one match between technological competence and product specialisation and that the dynamics of technological diversification are different from the dynamics of downstream diversification (Gambardella and Torrisi, 1998).

This paper argues that these two dynamics are closely interconnected: while firms need to accumulate a similar set of technological competencies to compete in a certain industry, firms are likely to 'use' such competencies in different ways. Or, in other words, the firm's knowledge base consists of something more than the distribution of patents (or publications) across technological fields; it also embraces how firms deploy such competencies to deliver new products. As Pavitt (1998) and Nelson (1998) argue, there are two complementary elements in firm-specific knowledge: bodies of understanding (based on competencies in specific technological fields) and bodies of practice (which refers to the organisational knowledge that links the bodies of understanding with the firm's downstream, product specialisation)—see also Dibiaggio and Nesta (2003). Building upon the distinction between bodies of understanding and bodies of practice, this paper argues that intra-industry firm diversity should be analysed appropriately not only by examining the firms 'technological competencies or the firms' product diversification strategies, but also by paying attention to the interface between the two. In this paper, we study this 'interface' by focusing on the firms' research activities examined according to the products or processes that such activities are expected (by the firms) to originate and develop. From hereon, we refer to this 'interface' as downstream-profiled research activities.

Moreover, as has been stressed by many authors (Stiglitz, 1987; Pavitt, 1987; Cantwell, 1989; Antonelli, 1999), the process of firms' learning is characterised by being a local process of knowledge acquisition: firms do not appear to scan all possible choices, but rather 'they follow a specific course acting almost instinctively to capitalise on their past experience' (Holbrook et al., 2000, p. 1030). In this sense, the evolution of firm's knowledge base profiles over time can be characterised as being cumulative and incremental and, thus, can be predicted to display a fairly stable pattern over time.

In short, building upon the above discussion, we hypothesise that firms display stable patterns of specialisation in downstream-profiled research activities (that is, research activities evaluated according to the products or processes that they are expected to enable), and that such stable patterns of specialisation vary widely across firms.

### 2.2. Knowledge strategies and performance

The business strategy literature has highlighted the importance of intra-industry firm variety (versus inter-industry variety) to account for long-term, non-transient dispersion in profit rates (Rumelt, 1991). Consistent with this argument, several authors (Henderson and Cockburn, 1994; Powell et al., 1996; Bierly and Chakrabarti, 1996) have shown that superior performance is associated with the firm's capacity to create and accumulate knowledge—from the stock of patents in specific disciplines to research network particularities and other forms of organisational architectures devoted to learning processes. To strengthen such capacity, the management of knowledge becomes a crucial dimension for developing valuable, rare, difficult-to-imitate and non-substitutable resources: in other words, a source of sustainable competitive advantage.

A degree of consensus has emerged among scholars as to what constitutes the crucial conflicting (but potentially reinforcing) forces that knowledge management needs to reconcile in order to create an active and fruitful learning organisation (Argyris and Schon, 1978; Leonard-Barton, 1995). First, as argued by Cohen and Levinthal (1990), March (1991) and Levinthal and March (1993), firms are likely to be rewarded in the long term, the more they invest in in-house research activities and the more exploratory the nature of the search, since such efforts contribute to strengthening the firm's capacity to take advantage of knowledge generated outside its boundaries, and counterbalance the myopic features of experiential learning. However, such strategies generally involve higher levels of risk and costs than strategies guided by short-term optimisation of resources.

Second, firms need not only to be competitive specialised players—in order to achieve world-class command within a certain discipline—but they also need to broaden their areas of expertise. Broadening the knowledge base helps the firm, on the one hand, to be flexible in the face of technological change and, on the other hand, to impose causal ambiguity on competitors by creating knowledge through the combination of different (but familiar to the firm) technologies (Leonard-Barton, 1995; Brusoni et al., 2001). However, as Pavitt (1998, p. 440) points out, it is diversity downstream (in the product and process configurations that can be generated from a given technological knowledge) rather than just technological diversity that drives competition amongst innovating firms.

Building upon the above discussion, we would expect that better performance is strongly associated with the firm's capacity to expand the scope of its areas of expertise in research activities. We also examine whether the firms' diversity in technological competencies and firms' diversity in downstream-profiled research activities impact differently on performance.

## 3. Description of data sources and variables

Our empirical analysis is designed to throw light on the characteristics of firms' knowledge bases and on their implications for a firm's competitive position. To do this, we focus on the publication profiles of a set of active research players in an R&D intensive industry: the Spanish pharmaceutical industry. In using publication data, this study follows the path of those of Koenig (1982), Narin and Rozek (1988), Gambardella (1995), McMillan and Hamilton (2000) and Cockburn et al. (2000) among others, which have drawn on publication counts as an important indicator of research activity in the pharmaceutical industry.

The main objective in constructing the sample of pharmaceutical firms was to include every Spanish domestic firm active in research in order to have a sample of firms with a similar strategy. In other words, the aim was to work with an innovative group of firms seeking to accumulate capabilities oriented towards the generation and development of new products. In this sense, we considered every domestic firm that had published at least one document in the period 1981–2000. The publication data were gathered from the Institute for Scientific Information's (ISI) Science Citation Index (SCI) and the Web of Science. We collected data on every document pub-

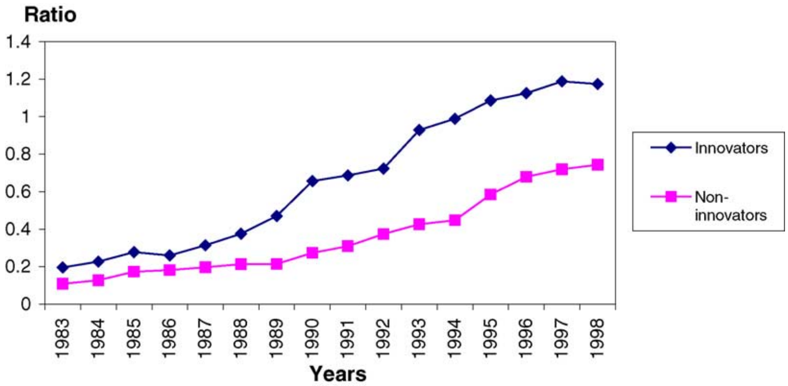


Fig. 1. Publication/100 employees (5-year moving averages).

lished in the journals included in the ISI SCI in this period for which at least one author's address matched that of a Spanish domestic pharmaceutical firm.

This yielded a total of 1210 published documents and a sample of 32 pharmaceutical firms (the list of firms is included in Appendix A)—accounting for approximately 28% of the total Spanish pharmaceutical market in 1999 in terms of sales. Most of these firms were founded before 1955, and to a large extent they have continued to be domestic-owned since then; however, 11 out of the 32 firms had undergone changes in their ownership structure due to partial or total acquisition by multinational corporations (MNCs). Since most of these acquisitions only occurred in the 1980s and 1990s, and since author affiliations in all cases included domestic addresses, we included all 32 firms in our research.[1] Finally, sales, employment, value of total assets and profits data come from Dun & Bradstreet and Sistema de Análisis de Balances Ibéricos (SABI) publications.

## 4. Research performance of Spanish pharmaceutical firms

Spanish domestic firms published 1210 documents between 1981 and 2000, 1032 of which were citable (all type of documents excluding Meeting Abstracts). As Fig. 1 shows, firms have steadily increased research intensity (as measured by publications/100 employees) over the period 1981–2000. Sequeira (1998) provides further empirical evidence on technological capabilities (regarding manufacturing and patenting activities) accumulated by Spanish pharmaceutical firms over this period.

However, while a relatively large number of firms have been active in research, as shown by the publications data, not all firms that publish have been equally successful in terms of technological performance (i.e. achieving *international* patents from active ingredients developed in-house). Indeed, only five firms can claim to have been successful in introducing new in-house chemical entities for which international patents have been granted. It should be noted that a large proportion of these in-house developed, internationally patented products have been economic successes, accounting for over 10% of total sales within a few years of market introduction (Galdon, 1996; D'Este, 2003).

We will refer to this group of five firms as 'innovative firms', as opposed to the other 27 firms that publish, but that have not been successful in developing in-house active ingredients, which we refer to as 'non (successful) innovative' firms. This designation does not mean that the latter group of firms is not undertaking innovative activities; on the contrary, this group has

[1] Moreover, from a random sample of 50 single authored articles (published since 1998), we confirmed that, with the exception of three cases, the authors were effectively affiliated to the firm (and not to a university or other publicly funded research centre). Therefore, we are confident about our assumption that authors are effectively affiliated to the firm. We are grateful to Dr. F. Jiménez-Sáez (INGENIO, Universidad Politecnica de Valencia, Spain) for his assistance in identifying authors' affiliations.

obtained a large number of international patents over the period 1981–2000, but for new methods, processes or formulations rather than new active ingredients.

## 5. Breadth, variety and persistence of firms' knowledge bases

As discussed in Section 2, in order to achieve world-class command within a certain discipline, firms need not only to be competitive specialised players, but they also need to broaden their knowledge base (Leonard-Barton, 1995; Patel and Pavitt, 1997; Brusoni et al., 2001). In Section 5.1, we investigate whether firms have been broadening their knowledge base by studying two different dimensions of knowledge expertise. On the one hand, we look at the scope of the scientific bodies of knowledge that firms have managed to integrate, as captured by the number of scientific disciplines in which firms have published: what can be interpreted as the firm's scientific competencies (see, for similar interpretation, Narin and Rozek, 1988). On the other hand, we investigate whether firms have expanded their knowledge expertise in terms of downstream-oriented research activities. To do so, we look at the scope of firms' publications across therapeutic areas. Since firms that are active in research in the pharmaceutical industry are oriented towards the discovery and development of products that claim to be *therapeutically* novel, firms need to accumulate a certain level of experience in order to be effective in researching in a number of therapeutic areas (Henderson and Cockburn, 1994). Thus, we wonder whether there is a tendency for firms to be relatively specialised around a narrow set of therapeutic fields. Finally, in Section 5.2, we examine firms' diversification profiles in order to analyse both distinctiveness and stability of the knowledge base.

### *5.1. Breadth of scientific competencies and of downstream-profiled research activities*

This section looks at whether firms have managed to integrate knowledge across a wide variety of scientific fields or have been accumulating scientific knowledge in only a narrow set of scientific fields. The CHI classification of journals in scientific fields is used in order to calculate the extent of diversification across scientific fields of firms' published documents. This shows that 99% of the documents published by the Spanish pharmaceutical firms belong to three aggregate scientific categories: clinical medicine (including scientific sub-fields such as pharmacology, pharmacy, cancer, cardiovascular system, gastroenterology, etc.), biomedical research (including sub-fields such as biochemistry and molecular biology, microbiology, biomedical engineering, etc.) and chemistry (i.e. organic chemistry, analytical chemistry, physical chemistry, etc.). This distribution profile of publications is essentially the same as that described by Narin and Rozek (1988) for the US pharmaceutical industry in 1976, where 86% of firms' publications were classified in these three aggregated scientific categories

In order to compare the scientific knowledge diversification of innovators and non-innovators in a meaningful way, we need to establish a minimum number of publications (some firms have too few to infer anything in terms of diversification). To do this, we compare the group of innovative firms with a sub-sample of 14 firms from the non-innovative group, which published at least 10 documents during 1981–2000. These two groups differ significantly in terms of size (number of employees): the group of non-innovative firms having an average size of 420 employees and the innovative firms 1000 employees.

Table 1 summarises the degree of diversification of documents published across scientific sub-fields by Spanish pharmaceutical firms in the period 1981–2000. It shows that the two groups of firms have very similar levels of diversification, either as measured by the Herfindahl Index (1 − Herfindahl Index, to have a measure of diversification instead of concentration) or as measured by the percentage of publications accounted for by the largest scientific sub-field. Moreover, even when the number of scientific sub-fields in which the firm has published at least one document is two times higher in the innovative group, the publication ratios of the two groups do not significantly differ when the number of publications is normalised by the number of employees.

According to Table 1, we would reject the hypothesis that innovative and non-innovative firms display different diversification levels of scientific competencies. There is no sign that innovative firms present a pattern of publications more evenly distributed across scientific fields, nor that they are actively publishing

Table 1
Diversification across scientific sub-fields (1981–2000)

| | 1 − Herfindahl Index | C1 (% of publications accounted for by the largest sub-field) | No. of sub-fields with at least one publication | No. of sub-fields/100 employees |
|---|---|---|---|---|
| Innovators | 0.64 | 57 | 13 | 1.7 |
| Non-innovators[a] | 0.66 | 48 | 6 | 2.9 |
| *t*-test[b] | Not significant | Not significant | Significant | Not significant |

[a] The group of non-innovators includes only the subset of 14 (largest publishing) firms.
[b] *t*-test for equality of means (two-tailed, 5% significance level). Non-parametric tests (i.e. Mann–Whitney test), testing differences in average ranks for the two groups, yielded similar results.

Table 2
Diversification across therapeutic categories (1981–2000)

| | 1 − Herfindahl Index | C1 (% of publications accounted for by largest therapeutic area) | No. of therapeutic areas with at least one publication | No. of therapeutic areas/100 employees |
|---|---|---|---|---|
| Innovators | 0.77 | 33 | 8.6 | 1.8 |
| Non-innovators[a] | 0.29 | 79 | 3.4 | 1.1 |
| *t*-test[b] | Significant | Significant | Significant | Not significant |

[a] The group of non-innovators includes only the subset of 14 (largest publishing) firms.
[b] *t*-test for equality of means (two-tailed, 5% significance level). Non-parametric tests (i.e. Mann–Whitney test), testing differences in average ranks for the two groups, yielded similar results.

across a larger number of scientific fields (relative to their size).

We also investigate whether firms have been able to accumulate knowledge across a wide variety of therapeutic scientific fields or have accumulated knowledge in only a narrow set of therapeutic areas. To do this, we classified the publications in terms of therapeutic fields (across the 15 therapeutic fields at the one digit level of the Anatomical Therapeutic Classification (ATC)).[2] The classification was carried out on the basis of the information contained in the keywords and abstracts provided by the ISI SCI database. Of the 1032 citable documents, 12.8% were not classifiable into a therapeutic category because no clear indication could be obtained from the keywords or abstracts.

Table 2 provides figures on the diversification of firms' publications across therapeutic categories comparing, as we did in Table 1, the firms in the innovative group with the 14 largest firms from the non-innovative group. The patterns this time are significantly different. Innovative firms display a much higher degree of publication diversification across therapeutic fields. Table 2 shows that the innovative firms display a broader spectrum of research expertise, while non-innovators remain narrowly focused (in several cases one therapeutic area accounts for all the documents published by the firm) throughout the period considered (1981–2000).

To sum up, analyses of breadth of the firms' knowledge base in terms of scientific knowledge background and in terms of research pipeline produced very different results. While both groups display a similar degree of diversification across scientific fields, they significantly differ in the degree of diversification of their research across therapeutic areas. Innovative firms have achieved a much broader knowledge base in terms of research across therapeutic fields than non-innovative firms.

### 5.2. *Variety and persistence of the firms' knowledge base*

In this section we focus not on the extent of diversification, but on the diversification profile in order to analyse both distinctiveness and stability of the knowledge base. It has been shown that firms' scientific knowledge bases are similarly diversified between

[2] The 15 therapeutic areas are: alimentary tract and metabolism (A), blood and blood forming organs (B), cardiovascular system (C), dermatologicals (D), genito-urinary system (G), systemic hormonal preparations (H), general anti-infectives (J), cytostatics (L), musculo-skeletal system (M), central nervous system (N), parasitology (P), respiratory system (R), sensory organs (S), diagnostic agents (T) and various (V).

innovative and non-innovative firms. We now want to test whether the composition of scientific fields in which firms are accumulating knowledge is similar or different across firms. In order to answer this question, we correlated each firm's scientific knowledge profile (i.e. publication shares across scientific sub-fields) against those of all other firms. We consider all scientific sub-fields in which firms publish at least once over the period 1981–2000: this produces a total of 38 scientific sub-fields.[3] We find that the scientific knowledge profiles of these firms are remarkably similar, as indicated by the fact that 53% of the cross-firm correlations (90 out of 171) were positive and significant (see Table 3).[4]

Table 3 shows that active research firms accumulate knowledge in similar scientific sub-fields: firms have a similar knowledge base composition in terms of the profile of scientific sub-fields in which they publish. In other words, this evidence points to the fact that pharmaceutical firms need to accumulate a similar scientific knowledge background in order to become active research players, a knowledge background that can be interpreted as an entry barrier to competition in innovation.

A completely different picture emerges when we address the firms' research portfolio in terms of the distribution of publications across therapeutic areas. We correlate each firm's publication shares across therapeutic areas against the other firms. We consider all therapeutic areas (at the one digit level of the ATC) in which at least one document has been published by any of these firms in the period 1981–2000, which is all 15 categories. As Table 4 shows, each firm displays a distinct research portfolio. More than 86% of total

Table 3
Correlations of firms' publication shares across 38 scientific sub-fields (period 1981–2000)

| | Esteve | Ferrer | Uriach | Faes | Abello | Andromaco | Alter | Antibioticos | Cusi | Lacer | Lasa | Menarini | Vinas | Vita | ASAC | Leti | Grifols | Salvat |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Almirall-Pr. | 0.961** | 0.964** | 0.944** | 0.857** | 0.029 | 0.219 | 0.908** | 0.501** | 0.916** | 0.145 | 0.698** | 0.945** | 0.743** | 0.933** | 0.576** | 0.045 | −0.041 | 0.188 |
| Esteve | | 0.978** | 0.966** | 0.921** | −0.015 | 0.218 | 0.952** | 0.499** | 0.891** | 0.155 | 0.724** | 0.993** | 0.755** | 0.940** | 0.580** | −0.012 | −0.068 | 0.187 |
| Ferrer | | | 0.977** | 0.883** | −0.039 | 0.234 | 0.944** | 0.514** | 0.911** | 0.200 | 0.746** | 0.980** | 0.762** | 0.965** | 0.590** | −0.021 | −0.056 | 0.165 |
| Uriach | | | | 0.868** | −0.032 | 0.251 | 0.950** | 0.604** | 0.914** | 0.214 | 0.745** | 0.977** | 0.765** | 0.980** | 0.604** | −0.003 | −0.062 | 0.219 |
| Faes | | | | | −0.035 | 0.254 | 0.889** | 0.407** | 0.771** | 0.242 | 0.756** | 0.924** | 0.638** | 0.855** | 0.491** | −0.066 | −0.081 | 0.131 |
| Abello | | | | | | 0.212 | −0.036 | −0.061 | −0.045 | −0.065 | −0.050 | −0.022 | −0.016 | −0.041 | −0.004 | 0.931** | −0.074 | −0.016 |
| Andromaco | | | | | | | 0.220 | 0.041 | 0.179 | 0.547** | 0.566** | 0.213 | 0.270 | 0.303* | 0.186 | 0.413** | −0.055 | −0.055 |
| Alter | | | | | | | | 0.533** | 0.873** | 0.325* | 0.724** | 0.964** | 0.712** | 0.920** | 0.570** | −0.037 | 0.131 | 0.277* |
| Antibioticos | | | | | | | | | 0.548** | 0.112 | 0.378** | 0.530** | 0.379** | 0.606** | 0.349* | −0.056 | −0.112 | 0.385** |
| Cusi | | | | | | | | | | 0.161 | 0.650** | 0.895** | 0.716** | 0.904** | 0.575** | −0.019 | −0.080 | 0.269 |
| Lacer | | | | | | | | | | | 0.677** | 0.180 | 0.077 | 0.277* | 0.084 | −0.077 | 0.360* | 0.127 |
| Lasa | | | | | | | | | | | | 0.734** | 0.539** | 0.798** | 0.416** | −0.050 | −0.090 | 0.080 |
| Menarini | | | | | | | | | | | | | 0.748** | 0.946** | 0.588** | −0.019 | −0.079 | 0.235 |
| Vinas | | | | | | | | | | | | | | 0.771** | 0.477** | 0.043 | −0.104 | 0.310* |
| Vita | | | | | | | | | | | | | | | 0.598** | −0.010 | −0.065 | 0.143 |
| ASAC | | | | | | | | | | | | | | | | 0.044 | −0.085 | 0.159 |
| Leti | | | | | | | | | | | | | | | | | −0.099 | −0.032 |
| Grifols | | | | | | | | | | | | | | | | | | −0.101 |

* $p < 0.05$ (one-tailed), Pearson correlation coefficients.
** $p < 0.01$ (one-tailed), Pearson correlation coefficients.

[3] The 38 scientific sub-fields are: general and internal medicine; allergy; cancer; cardiovascular system; dermatology; endocrinology; gastroenterology; geriatrics; hematology; immunology; obstetrics and gynecology; neurology and neurosurgery; ophthalmology; arthritis and rheumatology; pathology; pharmacology; pharmacy; respiratory system; nephrology; veterinary; hygiene and public health; miscellaneous clinical medicine; physiology; embryology; genetics and hereditary; biochemistry and molecular biology; cell biology, cytology, histology; microbiology; virology; biomedical engineering; microscopy; general biomedical research; organic chemistry; analytical chemistry; physical chemistry; general chemistry; biology; and others.

[4] When Spearman rank-order correlation coefficients were computed, 48% of correlations were positive and significant: these results are not displayed in the paper but are available on request.

Table 4
Correlations of firms' publication shares across 15 therapeutic categories (one digit ATC; period 1981–2000)

| | Esteve | Ferrer | Uriach | Faes | Abello | Andromaco | Alter | Antibioticos | Cusi | Lacer | Lasa | Menarini | Vinas | Vita | ASAC | Leti | Grifols | Salvat |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Almirall-Pr. | 0.536$^{*}$ | 0.479$^{*}$ | −0.058 | 0.793$^{**}$ | −0.188 | −0.203 | 0.347 | −0.165 | −0.218 | −0.058 | 0.458$^{*}$ | 0.129 | 0.372 | 0.363 | 0.439 | −0.196 | −0.239 | 0.361 |
| Esteve | | 0.289 | −0.101 | 0.177 | −0.217 | −0.299 | 0.067 | 0.360 | −0.206 | −0.215 | 0.590$^{*}$ | 0.258 | −0.078 | −0.108 | 0.566$^{*}$ | −0.218 | −0.056 | −0.099 |
| Ferrer | | | 0.142 | 0.579$^{*}$ | −0.164 | −0.199 | −0.012 | −0.151 | −0.157 | −0.140 | 0.465$^{*}$ | −0.153 | 0.555$^{*}$ | 0.551$^{*}$ | 0.320 | −0.157 | −0.153 | 0.559$^{*}$ |
| Uriach | | | | −0.139 | −0.162 | −0.174 | 0.305 | −0.054 | −0.171 | 0.840$^{**}$ | 0.236 | −0.153 | −0.174 | −0.164 | 0.483$^{*}$ | −0.164 | 0.785$^{**}$ | −0.178 |
| Faes | | | | | −0.153 | −0.093 | 0.191 | −0.148 | −0.173 | −0.089 | 0.179 | 0.010 | 0.842$^{**}$ | 0.846$^{**}$ | 0.104 | −0.156 | −0.210 | 0.839$^{**}$ |
| Abello | | | | | | 0.379 | −0.062 | −0.076 | −0.086 | −0.092 | −0.122 | −0.089 | −0.092 | −0.077 | 0.039 | 0.999$^{**}$ | −0.099 | −0.097 |
| Andromaco | | | | | | | −0.072 | 0.022 | −0.133 | −0.077 | 0.019 | −0.139 | −0.071 | −0.048 | −0.089 | 0.383 | −0.065 | −0.015 |
| Alter | | | | | | | | −0.016 | −0.101 | 0.028 | −0.104 | −0.01 | −0.100 | −0.091 | 0.400 | −0.091 | −0.015 | −0.107 |
| Antibioticos | | | | | | | | | −0.093 | −0.103 | −0.117 | −0.096 | −0.100 | −0.084 | 0.050 | −0.084 | 0.323 | −0.100 |
| Cusi | | | | | | | | | | −0.098 | −0.126 | −0.094 | −0.095 | −0.080 | −0.173 | −0.080 | −0.107 | −0.100 |
| Lacer | | | | | | | | | | | 0.233 | −0.095 | −0.105 | −0.088 | 0.338 | −0.088 | 0.879$^{**}$ | −0.111 |
| Lasa | | | | | | | | | | | | −0.133 | −0.041 | −0.019 | 0.802$^{**}$ | −0.113 | 0.204 | −0.036 |
| Menarini | | | | | | | | | | | | | 0.056 | −0.084 | −0.165 | −0.084 | −0.113 | 0.033 |
| Vinas | | | | | | | | | | | | | | 0.989$^{**}$ | −0.186 | −0.085 | −0.115 | −0.998$^{**}$ |
| Vita | | | | | | | | | | | | | | | −0.156 | −0.071 | −0.096 | 0.990$^{**}$ |
| ASAC | | | | | | | | | | | | | | | | 0.031 | 0.373 | −0.196 |
| Leti | | | | | | | | | | | | | | | | | −0.096 | −0.090 |
| Grifols | | | | | | | | | | | | | | | | | | −0.121 |

* $p < 0.05$ (one-tailed), Pearson correlation coefficients.
** $p < 0.01$ (one-tailed), Pearson correlation coefficients.

possible correlations between firms were not significantly different from zero.[5] Therefore, the knowledge base varies to a large extent across firms in terms of their knowledge accumulation in research across therapeutic areas. In other words, variety across firms is largely confirmed by the firms' research portfolios and thus supports the proposition that the articulation of (scientific) knowledge is a firm-specific process. Firms deploy a common (though industry specific) set of scientific competencies along a variety of firm-specific research paths oriented to the generation and development of new products.

Finally, we investigate whether the firm's knowledge base is stable over time in order to assess the extent to which learning processes build upon prior experience and the extent to which the localised nature of learning imposes restrictions on shifts towards new, previously unexplored areas. Given that most of the non-innovative firms published in the 1980s but not in the 1980s, and given that they present a very narrow pattern of research diversification across therapeutic fields, we focus on the group of innovative firms in order to study firm's knowledge base stability. We test whether stability in firms' knowledge base is a significant phenomenon using the results derived from the correlations between the distributions of publications across therapeutic areas in different periods of time for each of the five innovative firms. The results, shown in Table 5, reject in all cases that the distribution of publications over time (for each firm) is random, highlighting that firms' research across therapeutic areas does not change abruptly over time, which supports the proposition that cumulativeness is a major influence on firm's research choices.

## 6. Knowledge base diversification and performance

We conducted some regressions to examine the relationship between company performance and knowledge base diversification. The sample is composed of 19 Spanish pharmaceutical firms (the 5 innovative and 14 non-innovative that have published more than 10 documents), during the period 1995–2000. Three

[5] The same percentage is obtained for Spearman rank-order correlation coefficients. These results are available on request.

Table 5
Analysis of persistence in research across 15 therapeutic categories: Spearman rank-order correlation coefficients

| | 1991–1995/1996–2000 | 1986–1990/1996–2000 | 1981–1990/1991–2000 |
|---|---|---|---|
| Almirall-Prodesfarma | 0.66*** | 0.629*** | 0.746*** |
| Ferrer | 0.83*** | 0.681*** | 0.681*** |
| Esteve | 0.433* | 0.474** | 0.492** |
| Uriach | 0.593*** | 0.461** | 0.583** |
| Faes | 0.589*** | – | – |

*Notes*: Second column compares the periods 1991–1995 and 1996–2000; third column compares the periods 1986–90 and 1996–2000; fourth column compares the periods 1981–1990 and 1991–2000. For each firm, the total number of publications was as follows: Almirall-Prodesfarma, 185; Ferrer, 130; Esteve, 89; Uriach, 70; Faes, 28. Faes has most of its publications in the 1980s and thus only the sub-periods 1991–1995 and 1996–2000 could be compared.

* $p < 0.1$ (one-tailed).

** $p < 0.05$ (one-tailed).

*** $p < 0.01$ (one-tailed).

measures of performance were applied: (1) returns on sales (ROS); (2) returns on assets (ROA); and (3) the log of sales per employee (ln(SALES/$L$)). We regressed each of these measures on three sets of variables.

First, a set of variables accounting for the extent to which the companies expanded their knowledge base, both in terms of technological (scientific) competencies and in terms of downstream-profiled research activities, during the period 1995–2000. To measure the first set of variables, we have computed a '1 − Herfindahl Index' in order to have a measure of diversification on a year-by-year basis, over the firm's distribution of publications. Thus, we define $\text{THERAPSCOPE}_{it}$ as the '1 − Herfindahl Index' computed on each firm's distribution of publications across therapeutic categories for every year during 1995–2000. $\text{SCIENCESCOPE}_{it}$ is the '1 − Herfindahl Index' of the distribution of publications across scientific fields computed for every year and company.[6]

We also include a set of variables to assess the extent to which the firm displayed a highly diversified (or narrow) knowledge base at the beginning of the period considered. We computed the degree of knowledge diversification (1 − Herfindahl Index) for each firm, on the basis of its distribution of publications over the period 1981–1994. This was done for both the distribution of publications across therapeutic areas ($\text{THSCOPE95}_i$) and across scientific fields ($\text{SCISCOPE95}_i$). We also explored another way to capture the firm's 'accumulated' knowledge breadth: an alternative regression includes $\ln(\text{STOCKPUB})_i$, which accounts for the total amount of papers published by the firm between 1981 and 1994, and two interaction variables that assess the 'combined' effect of the firm's stock of publications with $\text{THSCOPE95}_i$ (called $\text{INTERACT1}_i$) and with $\text{SCISCOPE95}_i$ (called $\text{INTERACT2}_i$).

The second set of variables attempts to control for the firm's conditions pre-1995. In particular, we consider the firm's size in 1995 (in terms of the log of number of employees ($\ln(\text{EMP95})_i$), and the age of the firm ($\text{AGE}_i$) to assess whether firm size and earlier entry have a positive impact on performance. A positive impact of these two variables could be expected because of the greater internal financial resources available to larger firms and as a result of first-mover advantages gained by early entrants.

And thirdly, the last set of variables accounts for firms' investment efforts in physical assets and number of employees. We want to assess whether the firm's commitment to investments in physical assets is strongly associated with performance, compared with the firm's efforts towards intangible assets (such as those that are accounted for by $\text{THERAPSCOPE}_{it}$ and $\text{SCIENCESCOPE}_{it}$). We take the stock of capital (as measured by the value of annual total assets) per employee, as capturing the firm's capital intensity

[6] Note that for each single firm (and year) we have used data on firms' publications for the four preceding years, and then calculated a 5-year average. Therefore, the distribution of publications corresponding to the year 1995 includes both data on publications in that year and also data on publications in the four preceding years. This is done in order to make the best use of the information available and to obtain more meaningful trends for these variables.

Table 6
Relationship between knowledge base diversification and performance

| | Measures of performance | | | | | |
|---|---|---|---|---|---|---|
| | Returns on sales | | Returns on assets | | ln(SALES/EMPLOYEES) | |
| | Specification 1 | Specification 2 | Specification 1 | Specification 2 | Specification 1 | Specification 2 |
| CONST | −0.266** (0.112) | −0.422*** (0.11) | 0.079 (0.132) | −0.094 (0.128) | 7.394*** (0.653) | 7.542*** (0.653) |
| THERAPSCOPE | 0.071*** (0.025) | 0.082*** (0.024) | 0.102*** (0.03) | 0.102*** (0.028) | 0.430*** (0.148) | 0.376** (0.142) |
| SCIENCESCOPE | 0.051** (0.025) | 0.049** (0.024) | 0.015 (0.029) | 0.016 (0.028) | −0.270* (0.145) | −0.239* (0.142) |
| ln(EMP95) | 0.058 (0.05) | 0.090* (0.048) | −0.007 (0.059) | 0.029 (0.056) | 0.192 (0.292) | 0.209 (0.283) |
| AGE | −0.0006** (0.0003) | −0.0001 (0.0003) | −0.0004 (0.0003) | 0.0002 (0.0003) | 0.002 (0.002) | 0.0014 (0.002) |
| THSCOPE95 | 0.095** (0.043) | – | 0.059 (0.05) | – | −0.392 (0.248) | – |
| SCISCOPE95 | 0.005 (0.024) | – | 0.042 (0.028) | – | 0.129 (0.139) | – |
| LN(K/L) | 0.056*** (0.011) | 0.065*** (0.011) | 0.026** (0.013) | 0.037*** (0.013) | 0.296*** (0.062) | 0.292*** (0.065) |
| LN(EMP) | −0.118** (0.047) | −0.138*** (0.045) | −0.052 (0.056) | −0.077 (0.053) | −0.043 (0.275) | −0.079 (0.270) |
| ln(STOCKPUB) | – | −0.039*** (0.013) | – | −0.048*** (0.015) | – | 0.042 (0.075) |
| INTERACT1 | – | 0.025** (0.012) | – | 0.021 (0.014) | – | −0.085 (0.072) |
| INTERACT2 | – | 0.037*** (0.013) | – | 0.052*** (0.015) | – | −0.021 (0.078) |
| No. of observations | 114 | 114 | 114 | 114 | 114 | 114 |
| $R^2$ | 0.53 | 0.54 | 0.42 | 0.45 | 0.48 | 0.47 |
| Adjusted $R^2$ | 0.49 | 0.50 | 0.38 | 0.40 | 0.44 | 0.43 |

*Notes*: Standard errors in parenthesis. The 114 observations correspond to our 19 firms over the period 1995–2000.
* $p<0.1$.
** $p<0.05$.
*** $p<0.01$.

deepening ($\ln(K/L)_{it}$), and the firm's number of employees over time ($\ln(\text{EMP})_{it}$).[7]

Thus, using a linear specification, our regressions are of the following form (where the dependent variable is one of our three measures of performance, subscript "$i$" denotes firms and subscript "$t$" denotes time):[8]

$$\begin{aligned}\text{PERFORMANCE}_{it} &= \text{CONST} + \beta_1\,\text{THERAPSCOPE}_{it} \\ &\quad + \beta_2\,\text{SCIENCESCOPE}_{it} + \beta_3(\ln(\text{EMP95}))_i \\ &\quad + \beta_4\,\text{AGE}_i + \beta_5\,\text{THSCOPE95}_i \\ &\quad + \beta_6\,\text{SCISCOPE95}_i + \beta_7\,\ln\left(\frac{k}{L}\right)_{it} \\ &\quad + \beta_8\,\ln(\text{EMP})_{it} + \varepsilon_{it}.\end{aligned}$$

Table 6 presents the OLS estimates. One of the interesting outcomes displayed in Table 6 is that the estimated coefficient of THERAPSCOPE is positive and significant in all regressions. This shows that there is a positive relationship between all performance measures and the extent to which firms have expanded their downstream-profiled research activities. Also, such positive relationship is not outweighed by the impact of firm size or firm's capital intensity deepening and, thus, the strength of the relationship between THERAPSCOPE and performance has proved to be robust.

The other measure of firm's knowledge breadth—the extent to which the firm has expanded its scientific competencies—is more erratic in terms of its degree of association with performance. However, as shown by the positive impact of the interaction effect, INTERACT2, on the two rates of return measures, the results indicate that firms that have accumulated scientific competencies in a broader set of fields have been rewarded in terms of performance. In other words, it is not just the actual stock of publications that matters; whether this stock of publications embraces more scientific fields is also significant. These results are consistent with those from

[7] With the exception of variables accounting for firms' knowledge diversification, all other variables, including the three performance measures, have been computed using 3-year moving averages. So, for example, the values for firm employees in 1995 is the average of the number of employees for the years 1994–1996.

[8] This is the expression for our first specification. The second specification includes the stock of publications and the two interaction effects (while removing the two variables accounting for the accumulated degree of firm's knowledge base diversification up to 1995).

other studies (e.g. Gambardella and Torrisi, 1998) that show that better performance is associated with companies that have increased their technological diversification.

One possible reason why the two variables accounting for firms' knowledge breadth display a different profile when related to performance may be due to time-lag features. The outcomes from the regressions in Table 6 provide some preliminary evidence to support the argument that the extent to which firms have managed to expand their scientific competencies has a positive, but deferred, impact on performance in comparison with firm's capacity to expand its downstream-profiled research activities (which seem to have a more direct, concurrent relationship with performance). In other words, the opportunities arising from technological (or scientific) diversification may take a long time to materialise and, therefore, the impact of technological diversification on performance can be assessed better by focusing on firms' cumulative efforts.

There is one instance where firm scientific diversification is not significantly associated with performance. This is the case when performance is measured in terms of the log of sales per employees. In this case, the firm's capital intensity deepening may outweigh the impact of variables related to technological diversification. Nevertheless, even in this specification, there is a significant and positive association between the firm's increasing diversification in downward-profiled research activities and performance.

Finally, it is important to note that such variables as initial size or age have very little or no impact on performance. This lends further support to the argument that purposive management (and particularly management of firms' intangible assets) is important in explaining performance: better performance does not emerge straightforwardly from earlier entry or larger size, deliberate efforts towards innovation must be made.

## 7. Summary and conclusions

There have been two themes running through this paper: intra-industry firm variety, and the relationship between the knowledge base and performance of the firm. The empirical results from this study support the following conclusions:

1. To get a better understanding of the interfaces between firms' scientific (or technological) competencies and product diversification, particular attention should be paid to downstream-profiled research activities. As the empirical evidence presented here shows, while firms competing in the same industry tend to accumulate competencies across a similar set of scientific fields, they display significantly different patterns in terms of the downstream (or product-oriented) profile of research activities. Moreover, these differences persist over time: firms do not suddenly change their direction of research, but remain firmly committed to certain research paths, with gradual changes occurring over time. In short, both *variety* between firms and *persistency* in firms' differences within an industry owe much to the distinction between 'products' and 'technologies', and to the firm-specific organisational knowledge that allows a given set of technologies to be deployed in different ways.
2. Better performance is positively associated with the capacity of firms to expand their knowledge breadth, as measured both in terms of diversification of scientific competencies and diversification of downstream-profiled research activities. However, the (positive) association with performance is much more consistent across different performance measures in the case of diversification of downstream-profiled research activities than in the case of diversification of scientific competencies. This result is consistent with the argument in this paper that while scientific (or technological) competencies are essentially the necessary ticket of entry to competition in industry, it is how those scientific (or technological) competencies are organised to produce new products and processes that will potentially generate a distinctive capability and, eventually, may result in sustainable competitive advantage.

In summary, the empirical results support the argument that the firm's knowledge base is a main driver of persistent heterogeneity within industries, on the one hand, because of the systematic variety in terms of how firms articulate and organise their research activities and their background knowledge, and on the other hand, because of the positive correlation between the firms' knowledge diversification and performance.

This study has two limitations that would be promising future avenues of research. First, the time length in this study is probably too short to properly analyse the direction of causality between performance and the breadth of the firm's knowledge base: longer time series would be necessary to consider time-lags and causality. Second, firms' organisational knowledge should be studied in more depth. Investigating more qualitative aspects of the firm's knowledge architecture would add to our understanding of the interfaces between technological capabilities and downstream strategies.

## Acknowledgements

The author is extremely grateful for the comments from Ed Steinmueller, Aldo Geuna, Mike Hobday, Jon Sapsed, Gustavo Crespi, Ammon Salter, two anonymous referees of *Research Policy* and several participants at the November 2003 Conference in honour of Keith Pavitt. While writing this paper, the author benefited from financial support from the "NewKInd" Project (funded by the European Commission), in an earlier version, and from an ESRC Postdoctoral Fellowship.

## Appendix A. List of Spanish firms

Laboratorios Abello SA
Laboratorios ASAC
Aldo Union, SA
Almirall-Prodesfarma
Laboratorios Alter SA
Laboratorios Andromaco SA
Laboratorios Alonga SA
Antibioticos SA
Laboratorios Aristegui
Industrial Farmaceutica Cantabria SA
Cepa
Laboratorios Cinfa SA
Laboratorios Cusi SA
Laboratorios Esteve, SA
Faes
Ferrer Internacional SA
Laboratorios Grifols SA
Juste SA Quimico Farmaceutica
Laboratorios Lacer SA
Lasa
Laboratorios Leti SA
Llorente
Menarini (Puig)
Laboratorios Normon SA
Laboratorios Rubio SA
Seid SA
Laboratorios SALVAT
J. Uriach & Cia
Laboratorios Vinas SA
Laboratorios Vita SA
Laboratorios Dr. Andreu
Laboratorios ELMU

## References

Antonelli, C., 1999. The Microdynamics of Technological Change. Routledge, London.

Argyris, C., Schon, D., 1978. Organizational Learning: A Theory of Action Perspective. Addison-Wesley, Reading, MA.

Bierly, P., Chakrabarti, A., 1996. Generic knowledge strategies in the US pharmaceutical industry. Strategic Management Journal 17, 123–135.

Brusoni, S., Prencipe, A., Pavitt, K., 2001. Knowledge specialization, organizational coupling, and the boundaries of the firm: why do firms know more than they make. Administrative Science Quarterly 46 (4), 597–621.

Cantwell, J., 1989. Technological Innovation and Multinational Corporations. Basil Blackwell, Oxford.

Cockburn, I.M., Henderson, R.M., Stern, S., 2000. Untangling the origins of competitive advantage. Strategic Management Journal 21, 1123–1145.

Cohen, W.M., Levinthal, D.A., 1990. Absorptive capacity: a new perspective on learning and innovation. Administrative Science Quarterly 35, 128–152.

D'Este, P., 2003. Uncovering the origins of firm heterogeneity: evidence from the pharmaceutical industry in Spain. Unpublished D.Phil. thesis, SPRU, University of Sussex.

Dibiaggio, L., Nesta, L., 2003. Technology strategy and knowledge dynamics: the case of biotech. Industry and Innovation 10 (3), 329–347.

Dosi, G., Nelson, R.R., Winter, S.G., 2000. The Nature and Dynamics of Organizational Capabilities. Oxford University Press, Oxford.

Dun & Bradstreet, various years. Duns 50,000 Spanish Leading Companies. Dun & Bradstreet Spain SA, Madrid.

Fransman, M., 1994. Information, knowledge, vision and theories of the firm. Industrial and Corporate Change 3 (3), 713–757.

Galdon, C., 1996. Innovacion y Progreso en la Industria Farmaceutica. Doce Calles, Madrid.

Gambardella, A., 1980. Science and Innovation. The US Pharmaceutical Industry During the 1980s. Cambridge University Press, Cambridge.

Gambardella, A., Torrisi, S., 1998. Does technological convergence imply convergence in markets? Evidence from the electronics industry. Research Policy 27, 445–463.

Granstrand, O., Patel, P., Pavitt, K., 1997. Multi-technology corporations: why they have distributed rather than distinctive core competencies. California Management Review 39, 8–25.

Henderson, R.M., Cockburn, I.M., 1994. Measuring competence? Exploring firm effects in pharmaceutical research. Strategic Management Journal 15, 63–84.

Holbrook, D., Cohen, W.M., Hounshell, D.A., Klepper, S., 2000. The nature, sources, and consequences of firm differences in the early history of the semiconductor industry. Strategic Management Journal 21, 1017–1041.

Koenig, M., 1982. A bibliometric analysis of pharmaceutical research. Research Policy 12, 15–36.

Leonard-Barton, D., 1995. Wellsprings of Knowledge: Building and Sustaining the Sources of Innovation. Harvard Business School Press, Boston, MA.

Levinthal, D.A., March, J.G., 1993. The myopia of learning. Strategic Management Journal 14, 95–112.

March, J., 1991. Exploration and exploitation in organizational learning. Organization Science 2 (1), 71–87.

McGee, J., Thomas, H., 1986. Strategic groups: theory, research and taxonomy. Strategic Management Journal 7, 141–160.

McMillan, G.S., Hamilton, R.D., 2000. Using bibliometrics to measure firm knowledge: an analysis of the US pharmaceutical industry. Technology Analysis & Strategic Management 12 (4), 465–475.

Narin, F., Rozek, R.P., 1988. Bibliometric analysis of US pharmaceutical industry research performance. Research Policy 17, 139–154.

Nelson, R., 1998. Selection criteria and selection process in cultural evolution theories. In: Ziman, J. (Ed.), Technological Innovation as an Evolutionary Process. Cambridge University Press, Cambridge.

Nelson, R., Winter, S., 1982. An Evolutionary Theory of Economic Change. Belknap, Cambridge, MA.

Noda, T., Collis, D.J., 2001. The evolution of intra-industry firm heterogeneity: insights from a process study. Academy of Management Journal 44 (4), 897–925.

Patel, P., Pavitt, K., 1997. The technological competencies of the world's largest firms: complex and path dependent, but not much variety. Research Policy 26, 141–156.

Pavitt, K., 1987. International patterns of technological accumulation. In: Vahne, J.E. (Ed.), Strategies in Global Competition. Croom Helm Ltd., London.

Pavitt, K., 1998. Technologies, products and organization in the innovating firm: what Adam Smith tells us and Joseph Schumpeter doesn't. Industrial and Corporate Change 7 (3), 433–452.

Penrose, E., 1982. The Theory of Growth of the Firm. Wiley, New York.

Powell, W.W., Koput, K.W., Smith-Doerr, L., 1996. Interorganizational collaboration and the locus of innovation networks of learning in biotechnology. Administrative Science Quarterly 41 (1), 116–145.

Rumelt, R.P., 1991. How much does industry matter? Strategic Management Journal 12, 167–185.

Sequeira, K., 1998. The patent system and technological development in late industrialising countries: the case of the Spanish pharmaceutical industry. Unpublished D.Phil. Thesis, SPRU, University of Sussex.

Stiglitz, J.E., 1987. Learning to learn localized learning and technological progress. In: Dasgupta, P., Stoneman, P. (Eds.), Economic Policy and Technical Performance. Cambridge University Press, Cambridge.