Industrial and Corporate Change, 2016, Vol. 25, No. 6, 1019–1038
doi: 10.1093/icc/dtw012
Advance Access Publication Date: 16 March 2016
Original article

Internationalization of corporate R&D activities and innovation performance

Jaana Rahko*

Department of Economics, University of Vaasa, Vaasa, Finland. e-mail: jaana.rahko@uva.fi

*Main author for correspondence.

Abstract

The internationalization of corporate research and development (R&D) activities is a growing phenomenon, but previous empirical studies provide inconclusive evidence of its effects on the innovation performance of firms. This article examines how the innovation performance of European firms changes when they begin to internationalize their R&D activities. Propensity score matching and difference-in-differences methods are applied to control for self-selection and to estimate the causal effect of R&D internationalization. Patent inventor data are used to track the locations of corporate R&D activities. Information on patent applications, patent citations, and technological fields is used to measure innovation output, quality, and diversity, respectively. The results indicate that firms with a greater number of previous innovations are more likely to begin international R&D activities. Moreover, beginning R&D internationalization further increases the innovative output of firms. The results also indicate a weaker increase in the technological diversity of innovation activities. In contrast, the difference in the average quality of innovations in favor of international firms is shown to be due to self-selection.

JEL classification: O32, F23, L25

1. Introduction

The internationalization of corporate research and development (R&D) activities is a prevalent phenomenon that has grown considerably over recent decades (Moncada-Paternò-Castello *et al.*, 2011; European Commission, 2012). International R&D investments are considered to be driven not only by the market-seeking objectives of firms but also by knowledge-seeking motives and improved access to new technological knowledge (Kuemmerle, 1999; Chung and Alcácer, 2002; Le Bas and Sierra, 2002). Consequently, the innovation performance effects of international R&D have begun to receive attention in recent empirical studies. However, the results obtained thus far are inconclusive. Some studies, e.g., Penner-Hahn and Shaver (2005) and Chen *et al.* (2012), report that R&D internationalization has a positive effect on the innovation performance of firms, but contrary findings have also been reported (Singh, 2008).

There are two alternative but not mutually exclusive explanations for why international R&D may be linked to the innovation performance of firms. First, firms self-select into conducting R&D abroad. Thus, firms that engage in overseas R&D could either be more-innovative firms that are able to cover the additional fixed costs of internationalization or less-innovative firms that go abroad to catch up and compensate for their technological weaknesses. Second, internationally distributed R&D activities can improve the innovation performance of firms by providing improved access to local scientists, knowledge spillovers, and universities (Florida, 1997; von Zedtwitz and Gassmann, 2002).[1] Prior empirical studies typically employ panel models that control for bias caused by time-invariant omitted variables (e.g., Chen *et al.*, 2012 and Hsu *et al.*, 2014); however, these methods do not properly account for the endogenous self-selection of firms. Therefore, we cannot interpret prior results as causal. The aim of the present article is to account for the self-selection process and offer more reliable results on the causal effect of the start of international R&D activities on firm innovation performance. This study contributes to the literature by applying propensity score matching and difference-in-differences (DID) methods to control for the endogenous self-selection process. The combination of these methods has been used by De Loecker (2007), Greenaway and Kneller (2008), Hanley and Monreal Pérez (2012), and others to estimate the causal effect of exporting on firm performance. However, to the best of our knowledge, this method has not been applied in the context of R&D internationalization. This combination of methods does not necessarily eliminate all endogeneity problems (Dehejia, 2005; Smith and Todd, 2005). Nevertheless, by considering only firms that begin internationalization and by controlling for the self-selection process, this approach can account for endogeneity better than the previous studies have done. Several robustness tests are conducted to further narrow down remaining endogeneity concerns.

This study analyzes European firms and their international R&D investments. European firms have exhibited a higher level of R&D internationalization than their American or Japanese competitors (European Commission, 2012), and therefore, the innovation performance effects of R&D internationalization should be especially important for European firms. The extant literature, however, has mostly analyzed US and Asian firms. In the present article, the analysis concentrates on medium-sized and large European firms, and their R&D activities are studied by analyzing their worldwide priority patent applications. Following previous studies, patent inventor addresses are used to track the international distribution of corporate R&D activities. Again, consistent with prior studies, we use information on patent applications, patent citations, and technological fields of patents to measure innovation output, quality, and diversity, respectively.

Our findings suggest that more-innovative firms in the lead-up are more likely to begin overseas R&D operations. Thus, we find evidence of positive self-selection. This self-selection explains 35–100% of the observed differences in innovation performance between domestic firms and firms that start international R&D. After controlling for self-selection using propensity score matching, the start of overseas R&D activities has a statistically significant positive effect on innovative output and on the technological diversity of innovation activities. In contrast, the difference in the innovation quality in favor of international firms is shown to be entirely due to self-selection. This result diverges from some earlier studies that find international R&D to have a positive effect on innovation quality. Thus, our results imply that self-selection does matter, and it needs to be accounted for when estimating the effects of international R&D activities.

The remainder of this article is organized as follows. The second section reviews prior literature discussing the internationalization of R&D activities. The third section presents the data and main variables. The fourth section presents the descriptive statistics. The fifth section discusses the empirical model and results. The sixth section analyzes the sensitivity of the results, and the seventh section concludes the article.

2. Previous literature

The literature indicates that since the 1980s, R&D activities have rapidly become more internationally dispersed (OECD, 2008; Picci, 2010). This is especially the case in Europe. One reason for this dispersion is that firms from small European countries have needed to internationalize their R&D activities due to the pressures of international demand and limited resources in their home countries (von Zedtwitz and Gassmann, 2002). Moreover, many European firms source knowledge and offset home country technological weaknesses by establishing R&D units in the United States and other countries (Almeida, 1996; Florida, 1997). Increasingly prevalent international R&D investments can also have a considerable influence on the innovation performance of firms and countries. The majority of empirical studies have focused on the drivers of R&D internationalization and how countries should adjust their policies to attract foreign R&D investment (e.g., Hegde and Hicks, 2008; Athukorala and Kohpaiboon, 2010). However, the effects on firm performance have also begun to receive attention.

1 In contrast, Argyres and Silverman (2004) argue that by centralizing R&D activities, firms can achieve economies of scale and avoid coordination and communication costs.

2.1 Benefits of R&D internationalization

The market-seeking, or existing capabilities exploiting, view of foreign direct investment suggests that R&D internationalization may be used to gain access to new markets and utilize the innovations developed in the home market internationally (Kuemmerle, 1999; von Zedtwitz and Gassmann, 2002). International R&D facilities may be necessary to adapt existing, domestically developed innovations to the conditions and legal regulations of foreign markets. In addition, local R&D units may allow a quicker introduction of new products to local markets (Lewin *et al.*, 2009). Thus, this view implies that more-innovative firms are more likely to establish overseas R&D units to gain access to foreign markets; however, the view does not predict subsequent changes in innovation performance.

Studies suggest that R&D internationalization is also driven by knowledge-seeking motives that aim to improve firm innovation performance. These studies argue that internationally distributed R&D provides firms with access to a wide range of new resources (Chung and Alcácer, 2002; Le Bas and Sierra, 2002). Having a local presence in different countries provides improved access to local scientists, informal knowledge networks and universities, and it may thus improve innovation performance (von Zedtwitz and Gassmann, 2002). Knowledge spillovers from competitors, customers, and other parties are another way firms may benefit from international R&D activities (Granstrand *et al.*, 1993; Kuemmerle, 1999). Spillovers are typically national or even local in scope due to factors such as the tacitness of knowledge bases and specialization of local labor markets (Jaffe *et al.*, 1993; Branstetter, 2001; Breschi and Lissoni, 2001). Thus, to access these spillovers, R&D units must be located near knowledge sources. Therefore, R&D activities benefit from collocation with industry peers more than many other corporate activities (Alcácer, 2006; Audretsch and Feldman, 1996). A local presence may also enable firms to improve their cooperation with customers. As customers can serve as an important source of new ideas and product development (von Hippel, 2005) and these ideas may not be easily transferable, a local presence can be crucial. Moreover, an internationally distributed R&D organization allows a firm to create a diverse knowledge base within the firm, which can facilitate innovation and lead to new ideas and combinations of existing knowledge (Patel and Pavitt, 1997). Thus, the more internationalized corporate R&D activities are, the larger the knowledge pools and potential spillovers that a firm can access and the greater the potential improvement of firm innovation performance.

However, R&D internationalization may involve the imitation of competitors rather than increasing original, in-house knowledge production. If this is the main driver of international R&D investments, there may be a negative self-selection process, whereby initially less-innovative firms engage in international R&D to catch up their competitors.

International R&D may also enable a firm to reduce the costs of R&D by utilizing country-specific cost advantages, such as hiring scientists or buying inputs in a low-cost country (von Zedtwitz and Gassmann, 2002). R&D activities may also be located abroad to exploit country-specific R&D subsidies or patent boxes. Whether this type of strategy affects the overall quantity and quality of firms' innovations remains unclear.

2.2 Costs of R&D internationalization

Internationally distributed R&D activities also generate additional costs which may weaken the innovation performance of firms. R&D activities have a potential for economies of scale, but if a firm's R&D facilities are spread out too wide and thin, the firm cannot achieve such economies (Argyres and Silverman, 2004). R&D activities are also subject to economies of scope, as research projects in different technological fields may support one another (Henderson and Cockburn, 1996). Absent proper coordination, these benefits may be lost in distributed R&D organization and lead to weaker innovation performance.

Overseas R&D units often require a certain degree of autonomy to be able to access local knowledge networks and create innovations (Ghoshal and Bartlett, 1988); however, this creates problems for firm-level coordination. Coordination failures may lead to duplicated research efforts and wasted resources. Coordination problems in distributed organizations may be exacerbated by the communication problems that internationalization may create. Geographic and cultural distances make communication and interunit learning more time-consuming and difficult. This is especially true for R&D units because communicating R&D-related information often includes transferring tacit knowledge that requires face-to-face meetings, which become more infrequent as geographic distance increases (von Zedtwitz and Gassmann, 2002). The risk of intellectual property infringements and knowledge spillovers from the firm may also increase with R&D internationalization (Sanna-Randaccio and Veugelers, 2007; Schmiele, 2013).

2.3 Previous empirical studies

As discussed above, we can identify both significant benefits and costs that stem from international R&D activities. Thus, the overall effect of international R&D activities on firm performance remains an empirical question. Furthermore, there are indications that previous innovation performance and other firm characteristics affect which firms engage in international R&D. This selection needs to be considered in the empirical methodology. Next, we will briefly summarize the findings of previous empirical studies and the methodologies they employ.

Using patent data, Iwasa and Odagiri (2004) and Penner-Hahn and Shaver (2005) study the internationalization of R&D activities in Japanese firms and find that it is associated with increased innovative output, at least for some firms. Chen *et al.* (2012) and Hsu *et al.* (2014) study Taiwanese high-tech firms and the geographic diversity of their overseas R&D investment and find that international R&D activities have a nonlinear but positive effect on the average quality of innovations. R&D offshoring is also shown to be associated with higher probability of innovation (Nieto and Rodriguez, 2011). Finally, R&D internationalization is shown to be linked to improved firm productivity (Todo and Shimizutani, 2008; Belderbos *et al.*, 2014), although not all studies confirm this finding (Fors, 1997). Other studies have analyzed the effects of geographically distributed R&D activities within countries rather than across national borders. For example, Singh (2008) studies the innovation quality effects of geographically distributed R&D using US patent data and finds that patents resulting from distributed R&D are of lower quality. Studies by Argyres and Silverman (2004) and Furman *et al.* (2006) suggest that decentralized and geographically distributed R&D is associated with lower innovation performance. On the contrary, Lahiri (2010) and Leiponen and Helfat (2011) find that geographically distributed R&D has a positive effect on the number of patent citations and on imitative innovation. To sum up, the majority of studies find that international R&D improves the innovation performance of firms, but the results with respect to nationally distributed R&D are somewhat contradictory.

Most of the above-mentioned papers on R&D internationalization use patent data to measure the innovation performance of firms. Patent inventor addresses are also often used to determine the R&D locations. The studies have typically employed panel models with firm random effects (e.g., Lahiri, 2010; Chen *et al.*, 2012; Hsu *et al.*, 2014) or firm fixed effects (e.g., Singh, 2008) to control for unobserved firm-specific heterogeneity. Nevertheless, if there exists unobserved firm-specific time-variant heterogeneity that affects both R&D location decisions and innovation performance, then neither random nor fixed effects model can solve the endogeneity problem. Moreover, simultaneity can also cause endogeneity in this setting. The extant research typically lags independent variables by one period, which is said to mitigate the problem. However, this method does not consider that firms' past innovation performance is likely to affect which firms engage in international R&D. Therefore, the methods employed in previous studies can suffer from endogeneity problems, and the results cannot be considered causal.

3. Data and main variables

3.1 Using patent data to determine the R&D locations

Patent data have been used in numerous firm- and country-level studies to examine the reasons for and effects of R&D internationalization, and the present study follows the same approach. Patent application data are obtained from the European Patent Office (EPO) PATSTAT database (2013). The data are aggregated at corporate group level under the assumption that the parent firm is the ultimate owner of its subsidiaries' patents. The Organization for Economic Co-operation and Development (OECD) HAN database (2013) and manual matching and firm ownership information from Bureau van Dijk's Orbis database are used to match subsidiary patents to parent firms.

Patent data are useful in studying R&D internationalization and innovation performance, as patent information is available for a long period and across nearly all countries. Technology classifications added by independent patent examiners also provide information on the technological field of inventions. Moreover, a patent application can be assigned to a country based on the address of patent's inventor. The addresses of inventors provide an accurate picture of where a firm's inventions are developed, and thus, we use this information to track the locations of corporate R&D activities.[2] If all inventors listed on a firm's patent applications in a given year are located in a single country, we conclude that the firm only engages in domestic R&D activities. If the firm's inventors reside in several countries

2 The inventor address can be misleading if, e.g., the inventor recently moved to another country. Nevertheless, according to Bergek and Bruzelius (2010), inventor information provides a fairly reliable picture of the location of R&D activities.

or in one country that changes from year to year, we conclude that the firm has internationally distributed R&D activities.

The treatment variable in our empirical models is the start of international R&D activities. This is a dummy variable that takes the value 1 if a firm begins to engage in international R&D in a given year and 0 otherwise. When we construct this variable, we require that the firm has not conducted international R&D during the preceding 2 years, and that it continues international R&D activities in the next 2 years after the start. Thus, to reliably measure the beginning of international R&D activities, we want to measure the geographic scope of firm's past innovation activities during the sample period and 5 years before it. Only firms that have applied for at least 10 patents during that time period are included in the sample. Patents that are co-applied by several firms are excluded in the determination of R&D locations because we wish to track the locations of firms' in-house R&D activities; however, in other patent-based innovation variables, these patents are included using fractional counting, i.e., a patent is assumed to be uniformly distributed among co-applicants.

To avoid home country bias in the patent data, the worldwide priority patent filings of each firm are counted. By using priority filings from every national patent office, we cover more inventions than by using EPO or Patent Cooperation Treaty (PCT) patent counts (de Rassenfosse *et al.*, 2013). A problem with this approach is that the PATSTAT database has missing inventor information for many national patent offices. The missing inventor country information can, nevertheless, be retrieved by following the steps suggested by de Rassenfosse *et al.* (2013), which recover the missing information with 97% accuracy.

3.2 Innovation performance variables

Firm innovation performance is analyzed using several variables that capture different aspects of innovation activity. The variables are following: number of patent applications as a measure of innovative output, number of citations as a measure of quality-weighted innovative output, number of citations per patent as a measure of the average quality of innovations, technological diversity index as a measure of the technological diversity of innovations, and technological diversity of citing patents as a measure of the breadth of technological impact. Next, we describe how these variables are constructed.

First, the innovation output measure is log $(Patents_{it} + 1)$. $Patents_{it}$ refers to the number of worldwide priority patent applications filed by firm i in a given year t.[3] Different measures of patent output are used as a robustness check.

After the patent application is published, the application may be referenced by other patent applications when subsequent inventions are based on or related to the earlier invention. The number of citations a patent receives is associated with several aspects of patent quality, such as the economic and social value of the patent, firm's market value, and patent renewal rate (Hall *et al.*, 2005; Harhoff *et al.*, 2003; Trajtenberg, 1990). The value distribution of patents is highly skewed, and hence, prior research has often used patent citations to better capture the economic value of firms' patents (Harhoff *et al.*, 1999; Trajtenberg, 1990). Thus, our second innovation performance measure, the quality weighted innovation output, is log $(Citations_{it} + 1)$. $Citations_{it}$ refers to the number of citations that a firm's patent applications filed in year t receive during our observation period. Third, the average quality of firm's innovations is measured using the following ratio: $Citations_{it}/Patents_{it}$. Again, we test the robustness of our results by using different citation measures.

We count the citations that a patent receives directly and as non-priority applications (i.e., subsequent applications that are filed in a different patent office and cover the same invention). The citation information in the PATSTAT database is imperfect for many national patent offices. Thus, we consider citations made in EPO, US, and PCT patents, which are reliably covered in PATSTAT. A patent can receive citations over decades, which we do not have time to observe. This means that some of patents in our sample have a longer period over which to receive citations than do other patents. To avoid this bias, the empirical approach compares citations to patents that are applied in the same year; and therefore, the truncated citation period treats all compared patents equally.

Fourth, we measure the technological diversity of innovation activities. If international R&D is used to source new technologies, firm's technological diversity may increase. Moreover, diversity can improve firm performance (Miller, 2006). Technological diversity is measured using patent technology codes (IPC) at the three-digit level. Using IPC classes, the technological diversity of firm's innovations is calculated as 1 − the Herfindahl index. Using this

3 We add one to the number of patent applications to retain firms with zero patents in our sample.

index in a patent context is suggested by Trajtenberg *et al.* (1997). However, Hall (2002) notes that the index is biased in the case of few patents, and an adjusted index should be employed. Thus, the bias-adjusted technological diversity index is written as follows:

$$\text{Adj. technological diversity index} = \left(1 - \sum_k \left(\frac{N_k}{N}\right)^2\right)\left(\frac{N}{N-1}\right), \tag{1}$$

where N is the number of IPC codes in a firm's patent applications, and N_k is the number of patent applications assigned to technology class k. The index takes values between 0 and 1, where high values indicate a high degree of technological diversity. If several IPC codes are assigned to a patent, we assume that an identical fraction of the patent is assigned to each class. Missing values (i.e., observations with no patent applications or single technology class) are replaced by zeros.

Finally, we measure the breadth of technological impact, i.e., the diversity of citations received. The number of citations describes the quality of a patent and the magnitude of its impact on later inventions. The technological diversity of citations describes patent's generality or breadth of impact (Henderson *et al.*, 1998; Argyres and Silverman, 2004). If the citations come from few technological fields, the invention is likely to be incremental, whereas citations from many different fields indicate an invention with wide applicability. Using the IPC technology codes of each citing patent, the breadth of impact is calculated using the above-described Herfindahl index. Now, N_k is the number of forward citations from patents assigned to class k, and N is the number of IPC codes in citations. Missing values are again replaced by zeros.

Using patents to measure innovation activities is subject to some well-known limitations. Patents only protect technological inventions, and hence, many other inventions are excluded. Furthermore, many firms choose to use trade secrecy or lead-time and do not patent their inventions. Thus, the propensity to patent varies considerably across industries. Moreover, patents can only be used to measure and analyze new-to-market inventions. The advantages and disadvantages of patent data for our purposes have been discussed in detail, e.g., in Patel and Pavitt (1991) and Le Bas and Sierra (2002).

3.3 Control variables

We use several firm-level control variables that are based on firm balance sheet data. These data are obtained from Bureau van Dijk's Orbis database. We expect that, in addition to the innovation variables, the decision to engage in international R&D activities is affected by similar firm characteristics as the decision to enter export markets. Thus, we refer to the literature on export market participation to select the control variables (Wagner, 2007, 2012). Firm turnover is used to control for firm size, and the growth of turnover is used to control for growth performance. We also control for R&D intensity by including the ratio of R&D investment to turnover. Missing R&D expenditure figures are replaced by zeros, and a dummy variable is created to indicate these observations. The control variables also include dummy variables for years, industries, and countries. Industry classification uses NACE codes at the two-digit level. In categories with few firms, the one-digit codes are used instead.

We include one further important control variable for previous long-run innovativeness of firm. Past patent stock is counted using the number of patent applications as follows:

$$Patentstock_{it} = (1-\delta) \times Patentstock_{i,t-1} + Patents_{it}, \tag{2}$$

where the depreciation rate δ is set to 15% following the prior literature (e.g., Hall *et al.*, 2005). The patent stock includes patents since 1995.

4. Descriptive statistics

Our sample covers over 850 medium-sized and large firms in 23 European countries during the period 2003–2009. The sample includes all independent or stock-listed firms that have consolidated balance sheet data available in the Orbis database, have a turnover of over 10 million Euros, and have applied for at least 10 patents. This means that our sample is restricted to relatively large firms, and the results may not directly apply to small firms. Table 1 presents selected descriptive statistics for the sample firms. The financial variables have been deflated to year 2005 real prices using a gross domestic product deflator. On average, our sample firms have annual sales of 6500 million

Table 1. Descriptive statistics

Variable	Mean	Median	SD
International R&D status	0.572	1	0.495
Patents per year	52.936	7	202.846
Patent stock	260.767	32.790	1024.546
Citations	98.947	5	491.567
Citations/Patents	1.417	0.621	2.438
Technological diversity	0.570	0.681	0.342
Breadth of impact	0.526	0.644	0.350
Turnover	6515.671	824.760	18663.320
Growth of turnover	0.052	0.026	0.187
R&D intensity	0.048	0.013	0.143
R&D missing	0.294	0	0.456
Firm age	59.037	39	57.395

Notes: A total of 3598 observations. Financial variables deflated to year 2005 prices and are expressed in millions of Euros.

Table 2. Descriptive statistics by international R&D status

	Domestic firms		Starters		Firms with international R&D	
Variable	Mean	SD	Mean	SD	Mean	SD
Patents per year	4.845	5.899	12.120	19.787	100.000	281.213
Patent stock	22.854	24.391	50.104	75.310	496.663	1421.613
Citations	4.230	12.025	14.648	28.713	193.012	687.904
Citations/Patents	0.826	1.906	1.466	2.784	1.803	2.551
Technological diversity	0.416	0.391	0.590	0.323	0.669	0.268
Breadth of impact	0.372	0.384	0.515	0.334	0.634	0.285
Turnover	2104.17	7672.94	3831.37	10449.82	10462	24457.64
Growth of turnover	0.045	0.214	0.070	0.205	0.051	0.158
R&D intensity	0.058	0.212	0.027	0.046	0.048	0.101
R&D missing	0.484	0.500	0.305	0.461	0.160	0.367
Firm age	43.935	41.274	73.545	57.471	64.250	64.205
Number of firms	419		121		409	
Number of observations	1207		620		1771	

Notes: Financial variables deflated to year 2005 prices and are expressed in millions of Euros.

Euros, and R&D expenditures represent approximately 4.8% of turnover, while the medians are much lower. The mean number of patent applications that firms file each year is 53, while the median is lower, at 7 patents. On average, each patent receives 1.4 citations during the observation period.

The sample firms can be divided by their R&D internationalization status as follows: domestic firms, firms that begin international R&D activities, and firms that engage in international R&D throughout the observation period. Table 2 represents the main characteristics of the different groups. The table indicates that firms conducting a share of their R&D abroad are larger, more R&D intensive, file more patent applications, and receive more citations per patent than firms with domestic R&D activities. International firms also have a higher degree of technological diversity, and their innovations have a greater breadth of impact than those of domestic firms. The firms that begin to internationalize their R&D have intermediate characteristics and are, in general, more similar to domestic firms than to larger firms with uninterrupted international R&D. Thus, the table suggests that R&D internationalization is associated with higher quantity and quality of innovations. These characteristics are in line with abundant evidence that exporting firms are, on average, larger, more productive, and more innovative than are non-exporting firms (Wagner, 2007, 2012). In the R&D literature, the results of, e.g., Lahiri (2010) and Penner-Hahn and Shaver (2005) also point to the same conclusion.

However, Table 2 is uninformative of whether international R&D improves innovation performance or whether observed differences are due to self-selection. As discussed in the literature review, we can expect both effects to be significant. In the next section, we separate the selection effects and analyze how the start of international R&D activities affects innovation performance. Firms that conduct international R&D throughout the observation period are excluded in the following analysis.

5. Estimation and results

To control for the self-selection process and discover the causal effect of international R&D activities on the innovation performance of European firms, we use propensity score matching with DID estimation. This methodology estimates the causal effect by matching the firms that begin international R&D activities to similar firms that engage only in domestic R&D activities. Matching on propensity scores allows us to control for the endogenous self-selection into international R&D, and DID estimation and matching within years remove time-invariant firm-specific differences and common shocks. The treatment variable in our model is the start of international R&D activities.

Let us denote time periods such that a firm begins overseas R&D in period t. Following Heckman *et al.* (1997), the average effect of starting overseas R&D on innovation performance at time period $t+s$ is defined as follows:

$$E\{y^1_{i,t+s} - y^0_{i,t+s}|start_{it} = 1\} = E\{y^1_{i,t+s}|start_{it} = 1\} - E\{y^0_{i,t+s}|start_{it} = 1\}. \quad (3)$$

In the equation above, y denotes the performance variable of interest, and the superscripts denote international R&D status. The difficulty with this expression is that the last term of the above equation is not observable. This term is the performance that a treated firm would have had, had it not started international R&D. To capture this term, each treated firm is matched to one or more similar firms that do not receive the treatment using propensity score matching (Rosenbaum and Rubin, 1983).

The first step of propensity score matching is to estimate a probit model that captures how beginning to engage in R&D internationalization depends on observable pretreatment characteristics of the firm.[4] The dependent variable takes the value 1 when a firm begins international R&D and 0 otherwise. The explanatory variables are either lagged by one period or constant over time. The probability model explaining the decision to internationalize R&D is represented as follows:

$$\Pr\,(start_{it} = 1) = \Phi(y_{i,t-1}, X_{i,t-1}), \quad (4)$$

where $\Phi(.)$ is the normal cumulative distribution function, $y_{i,t-1}$ denotes the lagged innovation performance measure, and $X_{i,t-1}$ denotes all other lagged explanatory variables.[5] Because the innovation performance of firms is measured using several outcome variables, the probit model is estimated separately for each variable. Not including the respective lagged outcome variable in the propensity score estimation could lead to insufficient covariate balance in the matched sample, and the remaining self-selection could bias the results. Moreover, the coefficients of the lagged outcome variables provide evidence of underlying self-selection process. All the propensity score estimations control for the scale and scope of firm's past innovation activities by including the following variables: number of patent applications in the previous year, past patent stock, and past technological diversity index. Two of these variables are also outcome variables, and hence, we estimate four different probit models. Further explanatory variables are included as discussed in Section 3.3.

The number of observations in the probit estimations is smaller than in Table 2, because observations after the start of international R&D as well as all firms with continuous international R&D are excluded from the estimations. The results of the probit models presented in Table 3 show that firms that are older, have higher growth, and have applied for more patents in the past are more likely to start international R&D activities. Moreover, higher past innovation quality increases the probability of engaging in international R&D (column 3). In other words, there is positive self-selection, and firms with superior innovation performance in the past are more likely to engage in international R&D. This result indicates that at least a part of the difference between the groups reported in Table 2 is

4 Logit estimation yields very similar results as well.

5 We also estimated probit models in which 2- and 3-year lags of innovation performance were included. This had only minor effects on our results.

Table 3. The results of propensity score estimation. Firm's probability to start international R&D

	1.	2.	3.	4.
Constant	−2.819***	−2.767***	−2.856***	−2.831***
	(0.390)	(0.391)	(0.392)	(0.389)
log(Patents+1)	−0.004	−0.085	−0.019	0.057
	(0.107)	(0.119)	(0.107)	(0.111)
log(Patent stock)	0.204**	0.214**	0.225**	0.196**
	(0.100)	(0.100)	(0.100)	(0.100)
Technological diversity	0.133	0.131	0.133	0.198
	(0.154)	(0.155)	(0.155)	(0.157)
log(Citations+1)		0.086		
		(0.054)		
Citations/Patents			0.039**	
			(0.019)	
Breadth of impact				−0.347**
				(0.165)
log(Turnover)	0.066	0.061	0.063	0.069
	(0.042)	(0.042)	(0.042)	(0.042)
Growth of turnover	0.525**	0.518**	0.533**	0.542**
	(0.257)	(0.258)	(0.257)	(0.258)
Growth missing	0.090	0.089	0.094	0.094
	(0.220)	(0.221)	(0.221)	(0.220)
R&D intensity	−0.654	−0.695	−0.674	−0.627
	(0.795)	(0.819)	(0.806)	(0.798)
R&D not reported	−0.234*	−0.224*	−0.222*	−0.239*
	(0.133)	(0.133)	(0.133)	(0.133)
log(Firm age)	0.133**	0.134**	0.133**	0.132**
	(0.055)	(0.055)	(0.055)	(0.055)
Pseudo R squared	0.138	0.141	0.142	0.144
Obs	1496	1496	1496	1496

Notes: All explaining variables are lagged by 1 year. All estimations include country, industry, and year dummies.
*Significant at 10% level, ** significant at 5% level, and *** significant at 1% level.

due to self-selection. However, with respect to breadth of impact (column 4), we find evidence of negative self-selection, indicating that R&D activities that are general and relevant to a broad range of technologies are more likely to remain centralized.

Next, each treated firm is matched with similar untreated firms using the propensity scores. The matching is conducted within years and restricted to the area of common support. We apply several different matching estimators: radius matching, one-to-one nearest neighbor matching and kernel matching. There are over 12 times more control observations in our sample than there are treated observations. This ensures that there are many good matches available for most of the treated observations. However, a few observations with very high propensity scores may be poorly matched, although a common support restriction is applied. Therefore, our preferred model is radius matching, which can use many comparison observations while avoiding bad matches. The other matching estimators are discussed in the robustness analysis. Radius matching requires setting a radius size (i.e., the allowed distance between treated and control observations). We use a radius of 0.01 in our preferred model and discuss alternative radii in the robustness analysis.

To verify that the estimated propensity scores balance the covariates in our model, we calculate standardized bias in the covariates between the treated and matched control firms. The covariate means in treated and matched control group are reported in Appendix. Unfortunately, no clear guidelines exist on what level of remaining bias is acceptable. However, following Rosenbaum and Rubin (1985), the remaining bias should be smaller than 20%. Overall, radius matching significantly reduces bias for most variables. The remaining mean bias after radius matching varies slightly across the different propensity scores; however, it is always between 3.1% and 3.7%. Furthermore, the

Table 4. The results of DID estimation. Radius matching, $r = 0.01$

	ATT	SE	Obs
Log(Patents+1)			
t	0.557***	(0.086)	1480
t+1	0.556***	(0.122)	1480
t+2	0.560***	(0.128)	1480
Log(Citations+1)			
t	0.540***	(0.143)	1434
t+1	0.421***	(0.138)	1406
t+2	0.379**	(0.150)	1397
Citations/Patents			
t	−0.034	(0.367)	1216
t+1	0.051	(0.360)	1060
t+2	0.097	(0.384)	1021
Technological diversity			
t	0.145***	(0.049)	1480
t+1	0.063	(0.052)	1480
t+2	0.079	(0.052)	1480
Breadth of impact			
t	0.093	(0.057)	1481
t+1	0.145**	(0.060)	1481
t+2	0.031	(0.055)	1481

Notes: Bootstrapped standard errors with 200 repetitions.
*Significant at 10% level, ** significant at 5% level, and *** significant at 1% level.

largest remaining biases for individual variables are always below 20%. Comparing the pseudo-R^2 values of the propensity score estimation before and after matching reveals a decline in explanatory power from approximately 0.10 to 0.025. These results indicate that propensity score matching balances the observable covariates between treated and control firms and the balancing property is satisfied.[6]

Next, the innovation performance of the treated and matched control group is compared using the DID methodology which, in combination with matching, improves the quality of nonexperimental evaluation studies (Blundell and Costa Dias, 2000; Smith and Todd, 2005). This estimator uses the change in the outcome variable relative to the pretreatment value and estimates the difference in the changes between the treated and non-treated groups. The treatment effect is estimated for the year of the treatment (t), year after ($t+1$), and 2 years after the treatment ($t+2$). The results of the radius matching are reported in the table below. The table also reports bootstrapped standard errors and the number of observations in the area of common support. Only few observations with very high propensity scores fall outside the area of common support.

Table 4 presents the average treatment effects on treated (ATT). The results indicate that after controlling for self-selection, beginning R&D internationalization increases innovative output whether measured by the number of patents or the number of citations. The ATT on the number of patents and citations is approximately 0.55 in the first year, but the effect decreases somewhat over time with respect to the citations outcome. This implies an over 50% increase in the number of patent applications per year, which seems very large effect indeed. However, note that the median of patent applications in the year before start is only 3, and thus, the implied increase is only two patents per year. The median of citations is 2, and thus, the ATT implies increase of one or two citations.

Moreover, the ATTs reveal that the number of citations per patent does not change significantly, although the ATTs are positive in period $t+1$ and $t+2$. This finding indicates that the difference in the citations-patent ratio that

6 The covariate balance was also tested using alternative matching methods. Kernel matching and larger radius ($r = 0.05$) produced approximately similar balance; however, nearest neighbor matching and smaller radius ($r = 0.005$) performed worse.

we observe in the descriptive statistics in Table 2 is entirely due to self-selection and that beginning overseas R&D activities does not improve the average quality of innovations, at least not within 3 years. This result is in contrast to some earlier studies that do not control for self-selection (e.g., Chen *et al.*, 2012). However, the firms that engage in international R&D throughout the observation period have even higher average quality of innovations than beginning firms. Therefore, we cannot entirely exclude the possibility that innovation quality would improve after the 3-year period that we analyze.

The technological diversity of patents and the breadth of impact also increase after beginning overseas R&D. Thus, international knowledge sourcing helps firms to diversify their innovation activities to new fields of technology; and moreover, their inventions have wider applicability. The increase in the indexes varies between 0.031 and 0.145, and the estimated effects are not statistically significant for every lag. Regarding the technological diversity index, the positive effect is the strongest during the first year, whereas for breadth of impact, it is strongest in the second year. The ATT estimates for these variables are always positive, which differs from the results of Argyres and Silverman (2004) that centralized R&D activities lead to greater breadth of impact. The difference is again explained by self-selection because we found evidence of significant negative self-selection with respect to the breadth of impact variable.

Because the skewness of outcome variables and thus the validity of t-tests is a possible concern, we use Wilcoxon signed-ranks test as an additional check. This does not change the results with respect to number of patents, citations, technological diversity, or breadth of impact. However, with respect to the citations-patent ratio, the test indicates a positive and significant effect in period $t+1$ and $t+2$. Thus, there is some indication of improvement in innovation quality as also implied by the ATT estimates. However, the standard errors of ATT estimates are high, and thus the estimates with respect to innovation quality do not enable precise prediction.

Overall, the ATT estimates indicate that the quantity and diversity of firm's innovations increase when the firm engages in international R&D, but the improvement in quality is not statistically significant. Thus, we not only find positive self-selection but also observe a positive effect after start. The effect with respect to the number of patents and citations is highly significant, whereas the results on technological diversity and breadth of impact are not as strong. The change in innovation performance occurs during the first year of internationalization, and the difference compared to domestically operating firms persists in the later years. As observed in Table 2, firms with a long history of international R&D have clearly superior innovation performance than do firms just beginning to engage in international R&D activities. However, we did not find statistically significant improvement in innovation performance over time. Nevertheless, it is possible that the 3-year period we analyze is too short to capture long-run learning and changes. Firms with long international R&D experience might also be quite different from beginning firms in many other ways, which could explain our findings.

The descriptive statistics in Table 2 show that firms beginning international R&D have over two and half times more patents and patent citations than firms with only domestic R&D. They have almost twice as many citations per patent and are more technologically diversified. If we compare them to the firms that conduct international R&D during the whole sample period, the differences are greater still. To obtain a better grasp of the magnitudes of the selection effect and ATT, we can compare the ATTs to unmatched differences between the treated and control firms in observations within common support. The unmatched differences and mean-comparison test results are reported in Table 5. Compared to the unmatched differences, the ATT estimates are approximately 35% lower for the patent and citation variables and up to 65% lower for the technological diversity and breadth of impact variables. The unmatched differences also show a statistically significant difference in citations per patent. The differences with the ATT estimates are significant, and thus, the selection effect forms roughly one-half of the difference in innovation diversity and a somewhat smaller portion of the difference in innovation output outcomes.

6. Robustness analysis

6.1 Firms that increase their R&D spending

The key assumption in estimating the causal effect in the model above is that the differences between treated and non-treated firms are captured by observable characteristics. We calculate the propensity score using several observable firm characteristics, including past innovation performance, and we also use the DID approach and matching within years. Nevertheless, a change in innovation performance could be driven by unobservable shocks

Table 5. The unmatched difference in outcomes between treated and control groups

	Unmatched difference	SE	Obs
Log(Patents+1)			
t	0.897***	0.083	1480
t+1	0.871***	0.094	1480
t+2	0.890***	0.102	1480
Log(Citations+1)			
t	0.784***	0.132	1434
t+1	0.660***	0.124	1406
t+2	0.526***	0.115	1397
Citations/Patents			
t	0.035	0.122	1216
t+1	0.127*	0.092	1060
t+2	0.186**	0.109	1021
Technological diversity			
t	0.263***	0.027	1480
t+1	0.168***	0.034	1480
t+2	0.201***	0.033	1480
Breadth of impact			
t	0.147***	0.036	1481
t+1	0.235***	0.037	1481
t+2	0.090**	0.036	1481

Notes: *Significant at 10% level, **significant at 5% level, and ***significant at 1% level.

that are correlated with the start of R&D internationalization. For example, it seems possible that beginning international R&D activities is related to a general expansion of R&D activities. Therefore, a potentially more accurate control group is firms that expand their R&D activities domestically. Unfortunately, due to the patchy availability of inventor addresses, the R&D locations within countries cannot be reliably tracked. To assess this concern in another manner, we limit our sample to firms that increase their R&D spending in real terms, which we take as an indication of expanding R&D activities.[7] Within this limited sample, we again estimate the treatment effect of R&D internationalization and also use the increase in R&D investments to compute the propensity scores.[8] Thus, we now match firms that increase their R&D investments and begin international R&D with firms with similar growth in their R&D investments but who keep their R&D activities domestic. This limits our sample considerably because many firms do not report their R&D investments, and only approximately one-half of the firms report increases.

The results of these estimations are presented in Table 6. We used radius matching with both 0.05 and 0.01 radii, where the larger radius is our preferred choice because the significantly smaller sample reduces the number of possible matches and leads to significantly higher standard errors. Therefore, the estimates with $r = 0.01$ are mostly statistically insignificant, although the point estimates are similar to our original estimates in Table 4. Overall, the ATT estimates with respect to patents, citations, and citations per patent are similar, albeit somewhat weaker than in our baseline model. The ATT estimates (with $r = 0.05$) for the number of patents and citations in period *t* are approx. 0.51 and lower for the following periods. The standard errors are clearly higher, which leads to weaker statistical significance. With respect to technological diversity and breadth of impact, the ATTs are similar and even slightly higher than in our baseline model. Therefore, we are confident that our main results are not driven by unobservable firm-specific shocks that induce firms to expand their R&D activities.

7 We excluded firms that more than tripled their R&D investments from year to year. These outliers weakened the estimation of propensity scores and led to weaker balancing of covariates after matching.

8 Probit results are available upon request.

Table 6. The results of DID estimation. Radius matching, $r = 0.05$ and $r = 0.01$. Sample restricted to firms that increase their R&D expenditures

	Radius 0.05			Radius 0.01		
	ATT	SE	Obs	ATT	SE	Obs
Log(Patents + 1)						
t	0.511***	0.145	419	0.577**	0.233	413
$t+1$	0.390*	0.218	419	0.680**	0.309	413
$t+2$	0.290	0.241	419	0.495	0.358	413
Log(Citations + 1)						
t	0.508*	0.260	410	0.522	0.415	400
$t+1$	0.465*	0.270	406	0.545	0.421	396
$t+2$	0.166	0.275	402	0.120	0.448	392
Citations/Patents						
t	−0.097	0.560	366	−0.401	0.875	354
$t+1$	0.128	0.552	336	0.167	0.878	324
$t+2$	0.160	0.565	319	0.166	0.946	307
Technological diversity						
t	0.199**	0.099	419	0.155	0.168	413
$t+1$	0.044	0.108	419	0.045	0.172	413
$t+2$	0.051	0.100	419	0.097	0.154	413
Breadth of impact						
t	0.076	0.094	419	0.096	0.176	405
$t+1$	0.167*	0.087	419	0.142	0.150	405
$t+2$	0.052	0.099	419	0.097	0.148	405

Notes: Bootstrapped standard errors with 200 repetitions.
*Significant at 10% level, ** significant at 5% level, and *** significant at 1% level.

6.2 Sensitivity to selection on unobservables

As discussed above, our results are robust to controlling for the expansion of R&D activities, which is one way to test selection on unobservables. Nevertheless, there may exist other unobservables that induce firms to engage in international R&D and improve innovation performance. The presence of such unobservables cannot be directly tested. However, we can test how large the impact of such unobservables would have to be in determining selection to invalidate our main results. Rosenbaum (2002) and DiPrete and Gangl (2004) discuss a method to identify the bounds for the ATT estimates in the presence of unobservables.

According to Rosenbaum (2002), two matched observations with the same observable characteristics should have an identical probability of receiving treatment, i.e., the odds ratio (Γ) should equal 1. For example, if $\Gamma = 2$, then matched firms with the same observable characteristics are actually two times more likely to receive treatment due to unobservables. At each hypothetical value of Γ, the P-values of Wilcoxon signed rank test can be calculated, and assuming additive treatment effects, the Hodges-Lehmann point estimates can also be counted. We then calculate how large Γ, i.e., the magnitude of unobserved heterogeneity, is needed to make the ATT estimates statistically insignificant at the 10% level. These critical levels of Γ are reported in Table 7 for estimates that were statistically significant in Tables 4 and 6. It should be noted that these present the worst-case scenarios assuming an unobservable that has a strong effect on both treatment assignment and outcome. If an unobservable has a strong effect on the assignment but only a weak effect on the outcome, the ATT would remain statistically significant even at the reported levels of Γ. The critical values do not tell us whether unobservables exist; they only measure how sensitive our estimates are to potential unobservables.

Table 7 shows that the robustness to unobservables varies across outcome variables. With respect to the number of patents, the critical Γs are high, which indicates that the results are robust with respect to unobservable heterogeneity. With respect to citations, technological diversity, and breadth of impact outcomes, the critical values in Wilcoxon test

Table 7. Rosenbaum bounds. Critical Γ with cutoff $P = 0.10$

	Baseline model			Firms that increase R&D investments		
	Wilcoxon sign rank	Hodges-Lehmann	Obs	Wilcoxon sign rank	Hodges-Lehmann	Obs
Log(Patents + 1)						
t	5.18	4.57	1480	3.00	2.54	419
$t+1$	3.19	2.87	1480	1.53	1.34	419
$t+2$	3.52	3.16	1480	–	–	419
Log(Citations + 1)						
t	2.08	1.90	1434	2.18	1.88	410
$t+1$	1.61	1.48	1406	2.12	1.83	406
$t+2$	1.49	1.37	1397	–	–	402
Technological diversity						
t	1.50	1.37	1480	1.56	1.36	419
$t+1$	–	–		–	–	
$t+2$	–	–		–	–	
Breadth of impact						
t	–	–		–	–	
$t+1$	1.81	1.66	1481	1.41	1.21	419
$t+2$	–	–		–	–	

range from 1.49 to 2.08 in the baseline model, which are also relatively good values. This means that the results remain statistically significant even if an unobservable covariate causes the odds ratio of treatment assignment to differ by 50% between treated and control firms. However, in the smaller sample of firms that increase R&D investment, the Γs are lower for several outcome variables. Overall, the innovation quantity outcomes appear less sensitive to unobservable heterogeneity than innovation diversity, which again supports the main finding that the start of international R&D appears to have a stronger effect on innovative output and a weaker effect on innovation diversity.

6.3 Alternative specifications

In our baseline estimation, the treatment effect of R&D internationalization is estimated using radius matching. Next, we assess whether these results are sensitive to the choice of matching estimator. The choice of matching algorithm can be important, and there is typically a trade-off between bias and variance (Caliendo and Kopeinig, 2008). First, we estimate the baseline model using different radii: 0.05 and 0.005. Selecting a larger versus a smaller radius involves a similar trade-off between bias and variance as in the choice of matching estimators. Next, kernel matching and one-to-one nearest neighbor matching are considered as alternative matching estimators. If all matching approaches produce similar results, we can be fairly satisfied with our estimation approach.

Matching is conducted within years using the propensity scores estimated in Section 5. In the kernel matching model, we use an Epanechnikov kernel with a bandwidth of 0.06. The results of these alternative estimators are presented in the Appendix.[9] Summarizing these findings, we note that changing the matching estimator or radius has little effect on the estimated treatment effects. The most notable difference is that the results with respect to technological diversity and breadth of impact appear stronger when either kernel matching or a larger radius is used.

Next, we test different specifications of our innovation performance variables. In the baseline model, the logarithm of the number of patents or citations plus one is used due to frequent zero observations in the data. However, this choice may have an impact on the results. Next, we define the patent and citation variables as simply the

9 Huber *et al.* (2013) suggest that propensity score matching could be improved by using Mahalanobis matching and matching also on covariates that are good predictors of outcome. In our context, such covariates are, e.g., past innovation performance or increase in firm's R&D expenditure. The treatment effects were estimated using this approach; however, this did not materially change the results.

Table 8. The results of DID estimation with alternative innovation performance measures. Radius matching, $r = 0.01$

	ATT	SE	Obs
Log(Patents)			
t	0.556***	0.101	1211
$t+1$	0.476***	0.121	1190
$t+2$	0.508***	0.136	1151
Log(Citations)			
t	0.373	0.335	401
$t+1$	0.839**	0.363	341
$t+2$	−0.008	0.380	275
Log(Corrected Citations)			
t	0.716***	0.140	1431
$t+1$	0.699***	0.169	1428
$t+2$	0.787***	0.181	1422
Corrected Citations/Patents			
t	−0.070	0.394	1212
$t+1$	0.143	0.406	1191
$t+2$	0.194	0.455	1153

Notes: Bootstrapped standard errors with 200 repetitions.
*Significant at 10% level, ** significant at 5% level, and *** significant at 1% level.

logarithm of number of patents and the logarithm of citations received. This specification leads to a lower number of observations, especially for citations. The results are presented in Table 8. The results with respect to patent outcome are hardly affected by the specification change, but for the citation variable, the standard errors are now clearly larger. The estimates also vary considerably; however, the point estimates for the first 2 years are similar to our baseline results. The estimate for 2 years after the start of internationalization is close to zero. The number of observations is nevertheless quite low, which makes inference somewhat problematic.

Another computation of the citation counts is also tested. Patents may receive citations over a long time period, which we only partially observe. The truncation of the citation period may affect patents in different technological fields differently. In some fields, knowledge diffusion may be slower and citations may take longer to arrive than in others. To test whether truncation affects our results, we correct for the truncation using the method suggested by Hall *et al.* (2000) and applied in Hall *et al.* (2007). We allow for different knowledge diffusion processes in eight technological fields[10] and calculate the expected citation lag distribution for each field. Then, we estimate the expected number of citations in 10 years, given the citations observed thus far. The results for the truncation-corrected citation figures are reported in Table 8. Correcting for the truncated citation period leads to higher treatment effect estimates. The results now indicate that beginning international R&D activities increases the number of citations by over 70%. Regardless, after the truncation correction, the average quality of innovations, i.e., number of citations per patent, does not exhibit significant changes. Thus, the main implications remain unchanged although the truncated citation period may produce a slight downward bias in the ATTs.

6.4 Limitations of the study

In this study, we have not attempted to explore whether the gains from R&D internationalization depend on firm characteristics. However, previous studies suggests that the benefits and costs of internationally distributed R&D activities may depend on a firm's capability to integrate new knowledge and other firm characteristics (Singh, 2008; Lahiri, 2010). We also realize that the motivations of firms to engage in international R&D are likely to vary, and these differences may affect how and which firms benefit from international operations (Arvanitis and Hollenstein, 2011). Thus, there can be treatment effect heterogeneity that would be worth studying in further research.

10 The one-digit IPC classes are following: human necessities, performing operations and transporting, chemistry and metallurgy, textiles and paper, fixed constructions, mechanical engineering, physics, and electricity.

Furthermore, innovation performance is only analyzed at the firm level, and possible differences between overseas R&D units are not considered. This question provides interesting and relevant avenues for further research as well.

The key data used in this study are patent data, which only capture new-to-market inventions. Therefore, we are unable to measure the part of R&D internationalization that is conducted to absorb existing knowledge and create imitative innovations that are only new to an individual firm. This type of knowledge sourcing is undoubtedly important to the innovation strategies of many firms; however, it must be addressed with different types of data.

7. Conclusions

Despite the importance of international knowledge sourcing to the innovation strategies of firms, studies on the innovation performance effects of R&D internationalization have been scarce and provided mixed results. They also raise the question of whether the observed relationship between international R&D and innovation performance is due to self-selection into international R&D or to improvements in firms' knowledge sourcing. This question is the main interest of the present study. To provide an answer, this study has analyzed the internationalization of corporate R&D activities among European firms by applying matching and DID methods. Through this analysis, this study has provided novel evidence regarding the self-selection and causal effect of R&D internationalization on the innovation performance of firms.

The results indicate that more-innovative firms self-select to internationalize their R&D activities, which, in our sample, explains 35–100% of the observed quantitative differences in innovation performance between international and domestic firms. After we control for self-selection using matching methods, we observe that firms that begin to internationalize their R&D activities subsequently file approximately 50% more patent applications and receive more citations. At the median, sample firms file only a few patents per year, and thus, the implied increase is approximately two patents per year. R&D internationalization is also found to have a somewhat weaker positive effect on the technological diversity of firms and the breadth of technological impact. This implies that international R&D activities allow firms to diversify their innovation activities to new fields of technology. In contrast to some previous studies, we do not find a statistically significant effect on the average quality of innovations, and in that case, the self-selection process explains the higher average quality of innovations in international firms. The robustness of these results to selection on unobservables is assessed, and the results with respect to quantity of innovations appear strong, whereas the results with respect to technological diversity and breath of technological impact are somewhat more sensitive to possible unobservables. The sensitivity of the results to different matching methods and outcome variable specifications is also tested.

Our findings indicate that empirical research must account for the self-selection of firms to reliably assess the causal innovation performance effects of R&D internationalization. The results also have clear implications for organizing the R&D activities of firms. The innovation performance of firms significantly benefits from international R&D activities in terms of quantity and technological diversity. However, these benefits are not necessarily as large as initially envisaged due to the self-selection process. Moreover, our findings suggest that firms cannot expect improvements in innovation quality during the first years of R&D internationalization; however, firms with long histories of international R&D activities have significantly higher innovation quality, which may imply qualitative improvements later on. Unfortunately, the time frame of the present study does not allow us to analyze potential long run effects.

Moreover, our results relate to the body of literature on drivers of R&D internationalization. The findings indicate that international R&D activities help firms to increase and diversify their innovative output; thus, the results support the knowledge-seeking view of R&D internationalization. The results also show that firms with more innovations and higher innovation quality in the past are more likely to engage in international R&D activity, which is consistent with the capabilities-exploiting view of R&D internationalization. Therefore, both of these views offer important insights into the relationship between the internationalization of R&D and innovation performance of European firms.

This study represents only one step in understanding the causal effects of R&D internationalization. Firm characteristics and motives for engaging in R&D internationalization differ and may affect how the gains from such an activity materialize and are divided among firms. Interesting avenues for further research include the effects on firm productivity and imitative innovation, which cannot be studied using patent data alone.

Acknowledgements

The author would like to thank Valentina Tartari, Tuomas Takalo, Markku Lehmus, two anonymous referees and the participants of the DRUID Academy conference, FDPE Microeconomics and IO workshop, and Jyväskylä XXXI Summer Seminar for Finnish Economists for their helpful comments.

Funding

This work was supported by the Finnish Cultural Foundation, Nissi Foundation, and Wihuri Foundation.

References

Alcácer, J. (2006), 'Location choices across the value chain: how activity and capability influence collocation,' *Management Science*, **52**, 1457–1471.

Almeida, P. (1996), 'Knowledge sourcing by foreign multinationals: patent citation analysis in the U.S. semiconductor industry,' *Strategic Management Journal*, **17**, 155–165.

Argyres, N. and B. Silverman (2004), 'R&D, organization structure, and the development of corporate technological knowledge,' *Strategic Management Journal*, **25**, 929–958.

Arvanitis, S. and H. Hollenstein (2011), 'How do different drivers of R&D investment in foreign locations affect domestic firm performance? An analysis based on Swiss panel micro data,' *Industrial and Corporate Change*, **20**, 605–640.

Athukorala, P.-C. and A. Kohpaiboon (2010), 'Globalization of R&D by US-based multinational enterprises,' *Research Policy*, **39**, 1335–1347.

Audretsch, D. B. and M. P. Feldman (1996), 'R&D spillovers and the geography of innovation and production,' *The American Economic Review*, **86**, 630–640.

Belderbos, R., B. Lokshin and B. Sadowski (2014), 'The returns to foreign R&D,' *Journal of International Business Studies*, **46**, 491–504.

Blundell, R. and M. Costa Dias (2000), 'Evaluation methods for non-experimental data,' *Fiscal Studies*, **21**, 427–468.

Branstetter, L. G. (2001), 'Are knowledge spillovers international or intranational in scope? Microeconometric evidence from the US and Japan,' *Journal of International Economics*, **53**, 53–79.

Breschi, S. and F. Lissoni (2001), 'Knowledge spillovers and local innovation systems: a critical survey,' *Industrial and Corporate Change*, **10**, 975–1005.

Caliendo, M. and S. Kopeinig (2008), 'Some practical guidance for the implementation of propensity score matching,' *Journal of Economic Surveys*, **22**, 31–72.

Chen, C.-J., Y.-F. Huang and B.-W. Lin (2012), 'How firms innovate through R&D internationalization? An S-curve hypothesis,' *Research Policy*, **41**, 1544–1554.

Chung, W. and J. Alcácer (2002), 'Knowledge seeking and location choice of foreign direct investment in the United States,' *Management Science*, **48**, 1534–1554.

De Loecker, J. (2007), 'Do exports generate higher productivity? Evidence from Slovenia,' *Journal of International Economics*, **73**, 69–98.

de Rassenfosse, G., H. Dernis, D. Guellec, L. Picci and B. Van Pottelsberghe de la Potterie (2013), 'The worldwide count of priority patents: a new indicator of inventive activity,' *Research Policy*, **42**, 720–737.

Dehejia, R. (2005), 'Practical propensity score matching: a reply to Smith and Todd,' *Journal of Econometrics*, **125**, 355–364.

DiPrete, T. A. and M. Gangl (2004), 'Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments,' *Sociological Methodology*, **34**, 271–310.

EPO (2013), *PATSTAT - EPO Worldwide Patent Statistical Database, October 2013*. European Patent Office: Vienna.

European Commission (2012), *Internationalisation of Business Investments in R&D and Analysis of Their Economic Impact*. European Commission, Directorate General for Research and Innovation: Luxembourg.

Florida, R. (1997), 'The globalization of R&D: results of a survey of foreign-affiliated R&D laboratories in the USA,' *Research Policy*, **26**, 85–103.

Fors, G. (1997), 'Utilization of R&D results in the home and foreign plants of multinationals,' *The Journal of Industrial Economics*, **45**, 341–358.

Furman, J. L., M. K. Kyle, I. M. Cockburn and R. Henderson (2006), 'Public & private spillovers, location and the productivity of pharmaceutical research,' *NBER Working Paper 12509*, National Bureau of Economic Research, Cambridge, MA.

Ghoshal, S. and C. A. Bartlett (1988), 'Creation, adoption, and diffusion of innovations by subsidiaries of multinational corporations,' *Journal of International Business Studies*, **19**, 365–388.

Granstrand, O., L. Håkanson and S. Sjölander (1993), 'Internationalization of R&D—a survey of some recent research,' *Research Policy*, **22**, 413–430.

Greenaway, D. and R. Kneller (2008), 'Exporting, productivity and agglomeration,' *European Economic Review*, **52**, 919–939.

Hall, B. H. (2002), 'A note on the bias in Herfindahl-type measures based on count data,' in M. Trajtenberg and A. B. Jaffe (eds), *Patents, Citations and Innovations*. MIT Press: Boston, MA.

Hall, B. H., A. Jaffe and M. Trajtenberg (2005), 'Market value and patent citations,' *The RAND Journal of Economics*, **36**, 16–38.

Hall, B. H., A. B. Jaffe and M. Trajtenberg (2000), 'Market value and patent citations: a first look,' *NBER Working Paper 7741*, National Bureau of Economic Research, Cambridge, MA.

Hall, B. H., G. Thoma and S. Torrisi (2007), 'The market value of patents and R&D: evidence from European firms,' *NBER Working Paper No. 13426*, National Bureau of Economic Research, Cambridge, MA.

Hanley, A. and J. Monreal Pérez (2012), 'Are newly exporting firms more innovative? Findings from matched Spanish innovators,' *Economics Letters*, **116**, 217–220.

Harhoff, D., F. Narin, F. M. Scherer and K. Vopel (1999), 'Citation frequency and the value of patented innovation,' *Review of Economics and Statistics*, **81**, 511–515.

Harhoff, D., F. M. Scherer and K. Vopel (2003), 'Citations, family size, opposition and the value of patent rights - evidence from Germany,' *Research Policy*, **32**, 1343–1363.

Heckman, J. J., H. Ichimura and P. E. Todd (1997), 'Matching as an econometric evaluation estimator: evidence from evaluating a job training programme,' *The Review of Economic Studies*, **64**, 605–654.

Hegde, D. and D. Hicks (2008), 'The maturation of global corporate R&D: evidence from the activity of US foreign subsidiaries,' *Research Policy*, **37**, 390–406.

Henderson, R. and I. Cockburn (1996), 'Scale, scope, and spillovers: the determinants of research productivity in drug discovery,' *The RAND Journal of Economics*, **27**, 32–59.

Henderson, R., A. B. Jaffe and M. Trajtenberg (1998), 'Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988,' *Review of Economics and Statistics*, **80**, 119–127.

Hsu, C.-W., Y.-C. Lien and H. Chen (2014), 'R&D internationalization and innovation performance,' *International Business Review*, **24**, 187–195.

Huber, M., M. Lechner and C. Wunsch (2013), 'The performance of estimators based on the propensity score,' *Journal of Econometrics*, **175**, 1–21.

Iwasa, T. and H. Odagiri (2004), 'Overseas R&D, knowledge sourcing, and patenting: an empirical study of Japanese R&D investment in the US,' *Research Policy*, **33**, 807–828.

Jaffe, A. B., M. Trajtenberg and R. Henderson (1993), 'Geographic localization of knowledge spillovers as evidenced by patent citations,' *The Quarterly Journal of Economics*, **108**, 577–598.

Kuemmerle, W. (1999), 'The drivers of foreign direct investment into research and development: an empirical investigation,' *Journal of International Business Studies*, **30**, 1–24.

Lahiri, N. (2010), 'Geographic distribution of R&D activity: how does it affect innovation quality?,' *Academy of Management Journal*, **53**, 1194–1209.

Le Bas, C. and C. Sierra (2002), ''Location versus home country advantages' in R&D activities: some further results on multinationals' location strategies,' *Research Policy*, **31**, 589–609.

Leiponen, A. and C. E. Helfat (2011), 'Location, decentralization, and knowledge sources for innovation,' *Organization Science*, **22**, 641–658.

Lewin, A. Y., S. Massini and C. Peeters (2009), 'Why are companies offshoring innovation? The emerging global race for talent,' *Journal of International Business Studies*, **40**, 901–925.

Miller, D. J. (2006), 'Technological diversity, related diversification, and firm performance,' *Strategic Management Journal*, **27**, 601–619.

Moncada-Paternò-Castello, P., M. Vivarelli and P. Voigt (2011), 'Drivers and impacts in the globalization of corporate R&D: an introduction based on the European experience,' *Industrial and Corporate Change*, **20**, 585–603.

Nieto, M. J. and A. Rodriguez (2011), 'Offshoring of R&D: looking abroad to improve innovation performance,' *Journal of International Business Studies*, **42**, 345–361.

OECD (2008), *The Internationalisation of Business R&D: Evidence, Impacts, and Implications*. OECD: Paris, France.

OECD (2013), *OECD HAN Database*, July 2013. OECD: Paris.

Patel, P. and K. Pavitt (1991), 'Large firms in the production of the world's technology: an important case of "non-globalisation",' *Journal of International Business Studies*, **22**, 1–21.

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

Penner-Hahn, J. and J. M. Shaver (2005), 'Does international research and development increase patent output? An analysis of Japanese pharmaceutical firms,' *Strategic Management Journal*, **26**, 121–140.

Picci, L. (2010), 'The internationalizationof inventive activity: a gravity model using patent data,' *Research Policy*, **39**, 1070–1081.

Rosenbaum, P. R. (2002), *Observational Studies*. Springer: New York.

Rosenbaum, P. R. and D. B. Rubin (1983), 'The central role of the propensity score in observational studies for causal effects,' *Biometrika*, **70**, 41–55.

Rosenbaum, P. R. and D. B. Rubin (1985), 'Constructing a control group using multivariate matched sampling methods that incorporate the propensity score,' *The American Statistician*, **39**, 33–38.

Sanna-Randaccio, F. and R. Veugelers (2007), 'Multinational knowledge spillovers with decentralised R&D: a game-theoretic approach,' *Journal of International Business Studies*, **38**, 47–63.

Schmiele, A. (2013), 'Intellectual property infringements due to R&D abroad? A comparative analysis between firms with international and domestic innovation activities,' *Research Policy*, **42**, 1482–1495.

Singh, J. (2008), 'Distributed R&D, cross-regional knowledge integration and quality of innovative output,' *Research Policy*, **37**, 77–96.

Smith, J. A. and P. E. Todd (2005), 'Does matching overcome LaLonde's critique of nonexperimental estimators?,' *Journal of Econometrics*, **125**, 305–353.

Todo, Y. and S. Shimizutani (2008), 'Overseas R&D activities and home productivity growth: evidence from Japanese firm-level data,' *The Journal of Industrial Economics*, **56**, 752–777.

Trajtenberg, M. (1990), 'A penny for your quotes: patent citations and the value of innovations,' *The RAND Journal of Economics*, **21**, 172–187.

Trajtenberg, M., R. Henderson and A. Jaffe (1997), 'University versus corporate patents: a window on the basicness of invention,' *Economics of Innovation and New Technology*, **5**, 19–50.

von Hippel, E. (2005), 'Democratizing innovation: the evolving phenomenon of user innovation,' *Journal für Betriebswirtschaft*, **55**, 63–78.

von Zedtwitz, M. and O. Gassmann (2002), 'Market versus technology drive in R&D internationalization: four different patterns of managing research and development,' *Research Policy*, **31**, 569–588.

Wagner, J. (2007), 'Exports and productivity: a survey of the evidence from firm-level data,' *The World Economy*, **30**, 60–82.

Wagner, J. (2012), 'International trade and firm performance: a survey of empirical studies since 2006,' *Review of World Economics*, **148**, 235–267.

APPENDIX

Table A1. Balancing test. Covariate means after matching

	1.			2.			3.			4.		
Variable	Treated	Control	*P*-value	Treated	Control	*P*-value	Treated	Control	*P*-value	Treated	Control	*P*-value
log(Patents + 1)	1.742	1.802	0.615	1.732	1.851	0.785	1.774	1.828	0.642	1.747	1.772	0.836
log(Patent stock)	3.129	3.179	0.700	3.149	3.260	0.313	3.175	3.242	0.591	3.137	3.137	0.999
Technological diversity	0.543	0.524	0.700	0.537	0.551	0.764	0.554	0.547	0.879	0.552	0.550	0.958
log(Citations + 1)				1.300	1.351	0.785						
Citations/patents							1.411	1.460	0.899			
Breadth of impact										0.412	0.409	0.941
log(Turnover)	0.082	0.070	0.707	0.081	0.079	0.952	0.084	0.082	0.963	0.082	0.067	0.616
Growth of turnover	0.168	0.179	0.843	0.179	0.163	0.752	0.170	0.175	0.922	0.167	0.176	0.851
Growth missing	6.285	6.305	0.937	6.315	6.363	0.836	6.333	6.268	0.778	6.345	6.328	0.946
R&D intensity	0.030	0.030	0.957	0.030	0.033	0.804	0.030	0.032	0.877	0.031	0.028	0.793
R&D not reported	0.364	0.380	0.810	0.348	0.336	0.850	0.348	0.373	0.698	0.352	0.378	0.686
log(Firm age)	3.815	3.813	0.987	3.857	3.787	0.628	3.849	3.815	0.815	3.820	3.821	0.992

Notes: Covariate balance after the four separate propensity score estimations. *P*-values of tests for equality of means are reported.

Table A2. The results of DID estimation using alternative radii, kernel, and nearest neighbor matching

	Radius 0.05[a]			Radius 0.005[a]			Kernel matching[a]			Nearest neighbor matching[b]		
Log(Patents + 1)	**ATT**	**SE**	**Obs**	**ATT**	**SE**	**Obs**	**ATT**	**SE**	**Obs**	**ATT**	**SE**	**Obs**
t	0.620***	0.071	1492	0.578***	0.092	1469	0.616***	0.077	1492	0.628***	0.081	1492
$t+1$	0.496***	0.093	1492	0.567***	0.125	1469	0.499***	0.092	1492	0.537***	0.088	1492
$t+2$	0.540***	0.095	1492	0.605***	0.136	1469	0.541***	0.093	1492	0.710***	0.093	1492
Log(Citations + 1)												
t	0.486***	0.118	1442	0.588***	0.149	1427	0.494***	0.112	1443	0.422***	0.122	1443
$t+1$	0.397***	0.123	1414	0.527***	0.170	1399	0.397***	0.116	1415	0.456***	0.129	1415
$t+2$	0.316***	0.120	1405	0.421**	0.172	1390	0.322**	0.125	1406	0.385***	0.122	1406
Citations/Patents												
t	−0.076	0.266	1224	0.108	0.374	1204	−0.128	0.255	1224	0.530*	0.308	1224
$t+1$	−0.120	0.251	1067	0.072	0.325	1048	−0.110	0.231	1068	0.244	0.229	1068
$t+2$	−0.013	0.304	1028	0.413	0.381	1008	−0.008	0.259	1029	0.401	0.250	1029
Technological diversity												
t	0.152***	0.041	1492	0.180***	0.060	1469	0.148***	0.039	1492	0.124***	0.045	1492
$t+1$	0.045	0.041	1492	0.086	0.063	1469	0.043	0.038	1492	−0.001	0.045	1492
$t+2$	0.085**	0.040	1492	0.091	0.062	1469	0.083**	0.038	1492	0.031	0.046	1492
Breadth of impact												
t	0.088*	0.052	1490	0.111*	0.062	1476	0.088**	0.042	1490	0.078*	0.045	1490
$t+1$	0.161***	0.048	1490	0.163***	0.058	1476	0.156***	0.039	1490	0.175***	0.044	1490
$t+2$	0.038	0.047	1490	0.067	0.055	1476	0.036	0.040	1490	0.063	0.043	1490

Notes: [a]Bootstrapped standard errors with 200 repetitions.
[b]Subsampling standard errors with 200 draws.
*Significant at 10% level, **significant at 5% level, and ***significant at 1% level.