SEEDS
Sustainability Environmental Economics and Dynamics Studies

Working Paper Series

Universities as sources of information: comparing the role of 'open innovation' and companies' motivations.

by

Massimiliano Volpi

05/2014

The Sustainability Environmental Economics and Dynamics Studies (SEEDS) is an inter-university research centre that aims at developing and promote research and higher education projects in the fields of ecological and environmental economics, with a special eye to the role of policy and innovation in the path towards a sustainable society, in economic and environmental terms. Main fields of action are environmental policy, economics of innovation, energy economics and policy, economic evaluation by stated preference techniques, waste management and policy, climate change and development.

The SEEDS Working Paper Series are indexed in RePEc and Google Scholar.
Papers can be downloaded free of charge from the following websites:
http://www.sustainability-seeds.org/.

Enquiries: info@sustainability-seeds.org

SEEDS Working Paper 5/2014
April 2014
by Massimiliano Volpi.

Universities as sources of information: comparing the role of 'open innovation' and companies' motivations.

Massimiliano Volpi*

Abstract

The paper investigated the role Universities play as sources of information for companies' innovation. This study compared the explanations proposed by the 'open innovation' literature with those suggested by the 'resource view' of the firm, concluding that the way 'open innovation' variables have been constructed should be questioned and the 'resource view' theory should be augmented with innovation motivations, as companies rely on universities to source knowledge not generally available within companies' technological paradigm.

Keywords: green economy; environmental innovation; open innovation; universities

* Natural Environment Research Council, Polaris House, North Star Avenue, SN2 1EU, Swindon, United Kingdom.

1. Introduction

Investment in innovation is fundamental to the economic growth of countries. This central tenet of modern economic policy and of endogenous growth theories (Romer, 1990) has been supported empirically by many authors, including Guellec and de la Potterie (2001). Unfortunately, however, companies' often under-invest in innovation, because innovations are subject to externalities: competitors can copy from innovators and reap the benefits from innovations without having to incur their full costs. In the case of the environmentally motivated innovations, these innovation externalities are coupled to environmental externalities, as companies cannot entirely appropriate the environmental benefits that innovations produce. The presence of these 'twin externalities' leads to significant under-investment in (eco)-innovation (Jaffe et al. 2005). These externalities have justified government intervention, including support to universities for their role in producing and facilitating innovation. However, innovation surveys show that universities are among the least frequently used sources of information for innovating companies, prompting some policymakers to question the effectiveness of such an approach. This paper addresses these questions and informs policies which facilitate innovation through universities' support.

The analysis investigates the determinants of the value companies place on information from universities. It compares the explanations provided by the 'open innovation' approach (Laursen and Salter, 2004) with those proposed by the 'resource view' of the firm (Miotti and Sachwald (2003); Belderbos et al. (2004)). While the present study acknowledges the increasing importance of external information to companies' innovation strategy emphasised by the 'open innovation' approach (Chesbrough, 2003), it is critical of the way the 'open innovation' literature has measured 'openness' when explaining companies' reliance on information from universities (Laursen and Salter, 2004). The analysis highlights empirical and theoretical issues associated with using the openness variables proposed by Laursen and Salter (2004) and proposes alternative explanations. It emphases the importance of search pattern (Nelson and Winter, 1982) and of search motivations (Leiponnen et al, 2010; Arvanitis, 2012) over the sheer number of sources of information used.

Because the analysis focuses on companies for which (environmental) motivations are likely to be important, it is very close to previous analyses of environmental innovation and their determinants (Horbach et al 2008; Horbach et al. 2012; Kesidou and Demirel 2012, De Marchi 2011). These studies suggested introducing innovation motivations to analyse which companies rely on universities. The current paper is also closely related to analyses which investigated the complementarities between environmental and organisational innovations (Antonioli et al. (2013), Cainelli et al. (2012) Arvanitis et al. 2013); the issue of complementarity between environmental and organisational innovations is similar to that of 'open innovation': complementarity suggests adoption of one innovation increases with the number of the other innovations used; likewise, 'open innovation' suggests that the use of universities increases with the number of other sources used.

The analysis innovates upon existing literature both from a substantive and methodological perspective. From the substantive perspective, it explicitly tests the use of the breadth and depth of innovation variables as proxies for openness. The analysis rejects, on theoretical and empirical grounds, the use of these variables. The study proposes use of innovation motivations as an alternative and deeper explanation and

shows that motivations are empirically relevant. From a methodological perspective, the study explicitly takes into account the potential sample selection biases introduced by the structure of the questionnaire. Finally, the study follows a 'general to specific' approach (Hendry, 1995); which allows testing competing explanations against each other.

The paper is structured as follows: Section 2 provides a critical review of previous research, section 3 introduces the theoretical considerations that guide the paper; Section 4 explains the econometric model used in the investigation; Section 5 describes the data and the variables which have been included in the analysis. Section 6 discusses the results of the analysis and Section 7 concludes.

2. A review of the literature

Before addressing the controversies in the empirical literature on the role of universities as facilitators of innovation, it is useful to review the most relevant studies in an historical perspective[2]. This also allows highlighting the variables that are deemed most important.

The initial analyses concentrated on the relation between public research and private R&D. Cohen *et al*. (2002), focussed on manufacturing companies and showed the relation between private R&D and public research depended on companies' industrial sector, size and age. Larger companies and start-ups were found more likely to use information from the public research base[3]. Miotti and Sachwald (2003) investigated the choice of partners for R&D collaboration; they emphasised that the choice between different kinds of partners (public research, clients, competitors, suppliers) depends on complementarity of resources, on the level of R&D and on the distance from the scientific frontier. Belderbos et al. (2004) also showed that the effect of explanatory variables on R&D collaboration depends on the type of partner, as different partners are approached based on the underlying motivations.

Later analyses focused on the value of information from public research. Mohnen and Hoareau (2003) found similar results to those for R&D, with companies' industrial sector, R&D expenditure, and size all being relevant, in addition to having received government support, applied for a patent and being a radical innovator. Remarkably, they controlled explicitly for the potential selection bias introduced by the structure of the questionnaire, but found no evidence of such a bias. Laursen and Salter (2004) added two new variables, called breadth and depth of innovation. These aimed to measure the 'openness' of companies' search strategy. Breadth of innovation was defined as the sum of all information sources, other than universities, which were used to gather information for innovation; depth of innovation was likewise defined as the sum of all sources of information which were rated as important. They found the linear terms for breadth and depth of innovation variables to be positive and significant, but the quadratic terms to be negative, which they interpreted as evidence of decreasing returns[4]. Their results influenced much of the following research (Tether and Tajar,

[2] Comprehensive treatments of the early literature on the relationship between universities and innovation can be found in Mowery and Sampat (2005) and Salter and Martin (2001).

[3] Their sample is, however, a small one and strongly biased toward biological and medical sciences.

[4] Although Larsen and Salter (2004) did not comment this result, many previously statistically significant variables became insignificant with the introduction of the new variables, something that will be discussed later.

2008; Gonzales *et al*., 2009; Bruneel *et al*, 2010; Robin *et al*. 2012) who subsequently included these 'breadth and depth of innovation' variables. The exact specification form of these models differed from one author to the other: while Laursen and Salter (2004) had used both linear and quadratic terms for both variables, Tether and Tajar (2008) included just breadth and its quadratic term. Gonzales et al (2009) included a linear term for breadth only, and so did Brunel et al. (2010) and Robin et al. (2012).

The key difference between the literature based on the 'open innovation' and the analyses which predated this approach – based on the 'resource view' of the firm - is that, while the 'resource view' emphasised heterogeneity in the information sources, all papers that used the 'open innovation' variables implicitly considered the various sources of information homogeneous. In fact, in order to combine all sources of information into a single 'breadth of innovation' variable, it is necessary to assume that these sources are similar and can be used interchangeably. Therefore, it is not surprising that the 'open innovation' literature tended to neglect both the heterogeneity of information sources and the motivations for innovation that were discussed in previous papers (Miotti et al 2003; Belderbos et al. 2004).

Despite the dominance of the 'open innovation' approach in recent years, the introduction of breadth and depth of innovation variables did not completely terminate the debate on heterogeneity versus homogeneity of information sources. Recently, Jensen et al. (2007) proposed a theoretical distinction between two modes of learning for innovation purposes: Science, Technology and Innovation (STI) and Doing, Using and Interacting (DUI). This distinction is based on the notion of heterogeneity in the sources of learning. Fitjar and Pose (2012) tested Jensen's theory on innovation collaboration with different types of partners and found it useful to explain different types of innovation. Similarly, Kohler et al. (2012) analysed the contribution various sources of information bring to the innovation performance of companies. They argued that what kind of information each source provides matters more than the number of information sources it uses, but they neither included innovation motivations explicitly in their model, nor tried to test the 'open innovation' variables.

3 The hypotheses tested in the model and the underlying theory

This study aims to test the use of the breadth and depth of innovation variables made by the 'open innovation' theories (Laursen and Salter, 2004) against theories based on 'directional' searches (Jensen et al, 2007; Kohler et al., 2012) to address the controversy on heterogeneity. The present analysis contends that neglecting motivations and aggregating different sources of information into a single measure, as proposed by Laursen and Salter (2004), is surprising from an evolutionary approach. Innovation motivations and companies' heterogeneity would appear especially important from that perspective. If knowledge is local and tacit, while the capabilities of companies are 'inherited' and evolve only slowly, one would expect that the sources of information used should depend on the type of knowledge being sought and hence on the motivations for innovation. When companies are motivated be 'internal' objectives (i.e. objectives which are either internal to the company or typical of the technological paradigm within which companies operate), companies would be most likely to engage with sources which are close or 'internal' to their own industrial paradigm, i.e. their clients, competitors and suppliers. In that case, the most useful knowledge is likely to be local to the industrial or technological paradigm. However, when the motivations

driving innovations are 'external' (as in the case of environmental or health and safety regulations) companies would be expected to look further afield, because there is probably little useful information within their industrial/technological paradigm; they will then have an incentive to turn to 'external' sources of information, such as public research, universities, consultants and the like. For these reasons, it appears that 'internal' and 'external' motivations should be included in the explanation of the use of information sources. The relation between motivations and specific sources of information is coherent with theories like Jensen (2007) and Kohler et al. (2012) which assume the existence of preferential directions of search for specific problems. Hence, external motivations are expected to be especially important for the use of external sources, which is also coherent with the finding of regulation being a powerful driver for eco-innovation (Mazzanti and Costantini 2012; Snyder *et al*. 2003). Finally, it is well known that all sources of information 'external' to the company and its supply/client networks are used much less than 'internal' sources of information (Laursen and Salter, 2004). The differences are reflected in Table 2 and 3 and are striking: information sources internal to the technological paradigm, such as suppliers, clients, professional associations, and competitors are rated as much more important than information from sources external to the technological paradigm, such as consultants, universities, conferences, standards, and journals . Perhaps, differences in motivations for innovation might explain these differences: it is logical to expect 'internal' motivations (increasing quality, value added, market share, capacity and flexibility of the production process) to occur much more frequently than external ones (health and safety, environmental pressures) and, hence, to lead to a more frequent use of internal sources[5]. For these reasons, the present analysis argues that innovation motivations need to be explicitly included in the model assessing the use of information sources. This leads the study to the first hypothesis that will be tested:

Hypothesis 1: The use of knowledge sources depends on the purpose of innovation and on where the relevant information is most likely to be located. Innovation motivations which are 'external' to the company lead to an increased use of 'external' sources of information, such as universities. Likewise, innovation motivations 'internal' to the company, or its technological paradigm, lead to the use of sources of information 'internal' to the technological paradigm[6].

If different innovation motivations drive the use of internal or external sources, then aggregating all sources of information into a single measure, as the 'open innovation' theory did, would not be defensible because different sources are used for different purposes. Rather, it would be necessary to distinguish between sources which are 'internal' and those which are 'external' to the industry/technological paradigm and re-aggregate them accordingly. This led to the idea of testing the Laursen and Salter (2004) theory by splitting the variables for breadth and depth of innovation between sources which are 'internal' to the industry and those 'external' to it. If all sources are homogeneous, as they proposed, then the number of internal and external sources will

[5] Table 5 confirms the fact that 'external' motivations (such as environmental motivations and health) are considered important by a much lower percentage of companies than internal motivations, such as increasing quality. The formers are considered important by 22% and 26% of the companies, while the latter is deemed important almost by 60% of the companies.

[6] Some of the motivations in the survey are less clear-cut: the need to reduce costs and to increase quality is internal to the industrial sector but external to the individual company, hence their effects are harder to predict.

both have a positive effect on the use of information from universities. If, instead, 'internal' and 'external' sources are heterogeneous, one should expect that the number of internal and external sources used have opposite effects on the use of information from universities. This leads to the second hypothesis.

Hypothesis 2: if innovation motivations determine whether 'internal' or 'external' sources of information are used, because different information bases are best suited to address those innovation motivations, then internal and external sources would be expected to have opposite effects on the use of information from universities.

Splitting the variables used in the 'open innovation' literature between 'internal' and 'external' sources can help testing the homogeneity assumption implicit in the 'open innovation' approach. However, a different aggregation is not enough to address a potential pitfall in using these variables to explain the value of information from universities. The problem with any approach that tries to explain the value of information from universities using the number of other information sources companies rely upon is that it can lead to a circular explanation. A thought experiment will clarify the issue. Imagine a researcher who wanted to simultaneously explain the value of all external sources of information. If she were to follow the open innovation approach, she would probably write a set of equations, one for each source she intends to explain. She would include, in each equation, the number of "other" sources of information, excluding the one that the equation is supposed to explain, as in Laursen and Salter (2004). Finally, she would put all equations into a system. However, doing so leads to a cyclical explanation because, to explain source of information 1 in equation (1), the model relies on whether source 2 is used and, vice-versa, the explanation of source 2 in equation (2) is based on whether source 1 is used. This makes the cyclicality of this reasoning (and hence the endogeneity problem in the econometric specification) clear. The cyclicality of the reasoning would suggest that these variables should be omitted altogether from the explanatory variables, because they are not really a valid explanation[7].

Econometrically, it should also be noted that, if the theory proposed under hypothesis 1 is believed and the motivation variables are relevant to the source of information, but are incorrectly omitted from the individual equations, then even the estimates from each individual equations would be inconsistent. The reason for this is that the omitted variables would be correlated to the remaining regressors (the breadth and depth of innovation variables), by hypothesis 1. In this case, aggregating the different information sources would compound the endogeneity and the omission of variables problems, making both issues even more severe.

In order to test these hypotheses, the models presented here include motivations for innovation and split the number of information sources variable used by Laursen and Salter (2004) between sources 'internal' and 'external' to the industry. The models also include all variables used in previous literature, in keeping with the 'general to specific' approach (Hendry, 1995).

[7] Alternatively, if valid instruments could be found, then these endogenous variables could be instrumented. This, however, requires a large number of valid instruments and there are no such instruments in the data.

4. Econometric model and modelling strategy.

Econometric specification

The econometric specification adopted in the paper is influenced by the structure of the data: as the questionnaire asks only innovating companies about the value of information sources, there are no data on source of information used for companies which did not innovate. However, the researcher can always observe whether companies innovated or not and, hence, whether data are available. Therefore, the possible sample selection bias can be addressed. The estimation procedure mimics this data structure. It is based on two equations, which are tested in a single model. This model is called a sample selection ordered probit. The first equation (the selection equation) estimates whether or not companies innovate (hence, whether data are available) and is modelled with a probit; the second (outcome) equation estimates what value innovating companies place on information from a specific source of information, taking into account that this information is only observed for innovating companies; this is modelled with an ordered probit[8]. The two equations are linked by a correlation factor ρ that can be used for testing whether there is sample selection. If ρ is zero, there is no correlation between the two equations and the model could be modelled in terms of a simple ordered probit, because no bias is being introduced. If ρ is different from 0, however, using a simple ordered probit would render the estimates biased and inconsistent[9]. The bias is caused by the fact that, in this case, the sample which answers the question on the value of information is not random: it is correlated to the mechanism that explains whether companies are selected. One technical aspect of the sample selection ordered model that is important to note is that, in order to achieve identification, at least one of the variables included in the selection equation must be omitted from the outcome equation; this is known as the exclusion restriction condition. For a detailed treatment of the sample selection model the reader is referred to De Luca (2011).

A variety of models are estimated in order to test hypothesis 1 and 2. The first model includes both the motivations for innovation and the variables for 'breadth and depth of innovation' proposed by Laursen and Salter (2004). Two variations of this model are introduced: one includes the variable for collaboration with universities, while the other model excludes this variable[10]. These models are then compared with two models where the variables for 'breadth and depth' are split between 'internal' and 'external' sources, again with or without the collaboration variable, to test hypothesis 2. If the results from this second class of models show the 'open innovation' variables to be problematic, as it is expected on theoretical grounds, then the variables are removed and the third set of models tests the importance of motivations alone (again, with or without the collaboration variable).

5. Data and variables included

The data come from the UK Innovation Survey 2009[11]. The UK Innovation Survey is part of a well-established, Europe-wide Innovation Survey. It is now in its sixth iteration and it is based on the Oslo Manual, which is the standard for innovation surveys. The survey covers both the manufacturing and the service sector, sampling all companies in

[8] The reason for using an ordered probit model for the outcome equation is that the variables of interest are ordered, but they do not lend themselves to a cardinal interpretation. The data section provides a description of the data and the reason why an ordered probit is needed for those data.

[9] Informally, estimates are inconsistent if they remain biased no matter how many data are available. Consistency of the estimates is a highly desirable property.

[10] Robin et al. (2012) showed this variable is potentially endogenous and hence presenting the model with or without the variable is deemed a necessary robustness check, especially as no instruments are available in the data.

[11] Access to the data has been provided by the Office for National Statistics through the Virtual Microdata Laboratory in Newport and, when it became available, via the Secure Data Service.

sections C-K of the Standard Industrial Classification above 10 employees. It takes a stratified sample for UK companies below 250 employees and a census of companies above 250 employees. It is voluntary and implemented by means of a postal questionnaire. It was sent to 28,000 UK enterprises with a response rate just above 50%. The industrial sectors included in this study are listed in Table 1. These were deemed to be the sectors for which external motivations and, in particular, environmental motivations are most relevant. They were identified based on previous research and experts' consultation. For manufacturing, the decision to include specific sectors was based both on the level of environmental expenditure reported in the UK Environmental Expenditure Survey (UKEES) and on expert consultation; for the service sectors, which are not covered by UKESS, the relevant sectors were identified in consultation with personnel from the Natural Environment Research Council and the Environmental Sustainability Knowledge Transfer Network (ESKTN). As, for manufacturing, the outcome of expert consultations and of decisions based on the level of environmental expenditure led to the same conclusions, it is thought that the expert consultation was sufficiently reliable for selection of the service sectors. The sectors included cover a large share of the economy; out of 14,281 companies in the survey, 5,610 are included on the basis of their SIC classification. Including approximately 40% of all companies surveyed should mitigate the effects of possible imprecision in the selection process.

Table 1: Sectors included in the analysis

sector	Freq.	Percent	Cum.
Mineral Fuels and chemicals	256	4.56	4.56
Food	645	11.5	16.06
Utilities	136	2.42	18.48
Construction	1,059	18.88	37.36
Transport	719	12.82	50.18
Support_transport	331	5.9	56.08
Finance	536	9.55	65.63
Computer	461	8.22	73.85
Research	180	3.21	77.06
Management consultancies	575	10.25	87.31
Architectural and engineering consultancies	579	10.32	97.63
Technical testing	133	2.37	100
Total	5,610	100	

Dependent variables

The variable investigated is the importance companies attribute to sources of information for their innovation process. The responses are qualitative: respondents can rate information sources as of high, medium or low importance or as 'not applicable'. These responses were coded into values of 3 for 'high', 2 for 'medium', 1 for 'low' and 0 for 'not applicable', while non-response was treated as missing. As few respondents identify information from universities as of 'high importance', the 'high' and 'medium' categories were merged. Table 2 and 3 show the rating companies give to 'internal' and 'external' sources of information. While most internal sources are rated highly, external sources are not.

Table 2: How companies rate internal sources of information:

Rating of source of info	within business		Suppliers		Clients		Profess assoc		Competitors	
unit	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
Not applicable	152	7.36	227	11.06	125	6.04	465	22.65	207	10.04
Low	139	6.73	509	24.81	147	7.1	718	34.97	556	26.98
Medium	614	29.72	792	38.6	576	27.84	650	31.66	856	41.53
High	1,161	56.2	524	25.54	1,221	59.01	220	10.72	442	21
Total	2,066	100	2,052	100	2,069	100	2,053	100	2061	100

Table 3: How companies rate external sources of information

Rating of source of info	Consultants		University		Public research		Conferences		Standards		Journals	
units	Freq.	%	Freq.	%	Freq.	%	Freq	%	Freq	%	Freq	%
Not applicable	662	32.39	1020	49.78	983	47.93	627	30.53	462	22.5	707	34.54
Low	824	40.31	749	36.55	735	35.84	795	38.7	566	27.57	765	37.37
Medium	419	20.5	204	9.96	252	12.29	492	23.95	687	33.46	454	22.18
High	139	6.8	76	3.71	81	3.95	140	6.82	338	16	121	5.91
Total	2,044	100	2,049	100	2,051	100	2,054	100	2053	100	2,047	100

Explanatory variables

The CIS survey asked questions on the importance of information sources for innovation only to innovating companies. As explained in the econometric section, this leads to the need for an explicit sampling mechanism. Mohnen and Hoareau,(2003) and De Marchi, (2011) relied on a recall question on innovation activities included in the survey to deal with the sample selection. However, this study decided to use the actual responses to the innovation questions, rather than the 'recall' question. Significant inconsistencies were found between the 'recall' question in the survey and the answers to the innovation questions. In particular, 124 companies answered yes to the recall question and to questions on the usefulness of information sources, even though they did not innovate and should not have answered. These companies were omitted from the analysis for consistency.

All variables used as explanations in previous literature have been included in the model, with the exception of the variable for belonging to a group, which was unavailable.

Company size is measured as the natural logarithm of the number of employees and the variable is named LOGEMPLOYM[12]. Age of the company is accounted by a dummy variable, called ESTABLISHED, that takes the value 1 if a company was established in the last 3 years and zero otherwise. Two dummy variables for changes in turnover as a result of mergers and acquisition or sale of part of the business are introduced, to account for the potential effects of mergers and acquisitions on the probability to innovate. These are named respectively: INCR_TURNOVER and DECR_TURNOVER. Phillips and Zhdanov (2012) provided theoretical reasons and empirical evidence on their importance for the innovation decision. However, no previous empirical or theoretical study has found these variables to be significant explanations for the value of information sources. Hence, these variables are used only in the selection equation, while they are excluded from the outcome equations. They provide the exclusion restrictions needed for identification by the sample selection

[12] The exact maximum value of logemployment in table 4 is not shown, due to it being potentially disclosive. This would infringe the rules set by ONS for the use of their data.

model. The dummy variable EXPORTERS is used to distinguish between companies operating in the domestic or international market, as exporting companies have been found to innovate more and to rely more on universities (Bratti and Felice, 2011). The dummy variable NEWGOODORSERV distinguishes companies that introduce either new products or services from those that don't. Similarly, the dummy NEW_PROCESS identifies companies that introduced a new process. A distinction is also made between radical and incremental innovation. The dummy variable for introduction of goods new to the market is called NEW_TO_MARKET, while the dummy for processes new to the industry is dubbed PROC_NEW_IND; radical innovators are deemed more likely to interact with universities. The analysis also includes the type of innovation activities undertaken by each company: internal R&D has been captured by the dummy variable INT_RD; however, because the importance of R&D varies across sectors and is somewhat limited to manufacturing (Huang et al., 2010), all available innovation activities have been included as potential explanations. Innovation activities have been grouped according to the headings in the questionnaire[13]. This led to creating a single dummy variable for the acquisition of machinery, computer and software, called ANYMACHCOMP and set to one if any such activity was undertaken. Similarly, the dummy MKTINTROINN covers all market introductions of innovation, i.e. changes to product design, market research, marketing innovation and launch of advertising. Other innovation activities include the acquisition of external research, named EXT_RD; TRAINING, a dummy for training geared at introducing innovations and DESIGN, which captures engagement in activities preliminary to the introduction of R&D. The analysis also includes dummy variables for constraints on innovation. These take the value of one if the constrain is significant and zero otherwise. The constraints included in the analysis are whether the market is dominated by established business DOM_ESTABL, and the importance of UK and EU regulation: CONS_REG_UK and CONS_EU_REG, respectively. These are chosen because they are potentially relevant to the innovation decision and largely exogenous to the company. Not all constraints are included, because companies' assessment of some constraints on innovation has been found to depend on their previous innovation experience and some of the constraints were found to have signs opposite to those expected (Pellegrino and Savona, 2013). Companies which have already innovated might be more aware of constraints to innovation. Also, previous innovation activities affect the importance of some constraints; for example, firms which have already innovated are less likely to be constrained by lack of information, as they have already gathered the information they needed. Once carefully assessed, many of the constraints on companies are far less 'external' (exogenous) than they look at first sight. The model proposes to include only those constraints which are reasonably exogenous to the company; whether their market is dominated by established companies and the importance of regulatory constraints on innovation are such constraints. The presence of constraints on innovation from government and EU regulation is especially interesting because it is a policy variable. Table 4 summarises all of these variables. Industry dummies are also included into the analysis. The sectors are those listed in table 2 and construction is the reference (i.e. omitted) sector.

[13] The text of the questionnaire is available from BIS http://www.bis.gov.uk/assets/BISCore/corporate/docs/C/cis6-2006-2008-questionnaire.pdf .

A number of variables are only available for innovating companies, due to the structure of the questionnaire. They can, therefore, be included only in the outcome equation. These are the variables on the motivation to innovate, breadth and depth of innovation and collaboration with others. Motivations have been coded as dummies, equal to one where their importance was high or very high and zero otherwise. This allows looking only at the effect of key motivations for innovation. The following motivations were included and are summarised in table 6: increasing range of goods, *Drange*; entering new markets. *Dnew_mkt*; increasing market share *Dmkt_size*; improving quality, *Dqual*; improving flexibility for production, *Dflex*; increase capacity, *Dcap*; increasing value added, *Dva*; reducing costs *Dcost*; improving health and safety, *Dhealth*; reducing environmental impact, *Denv*; replacing outdated products or processes, *Doutdat*; meeting regulatory requirements, *Dreg*.

Breadth and depth of innovation sources are defined as in Laursen and Salter (2004). They represent – respectively – the number of external information sources that a company uses at least to some degree and the number of external sources that the company considers important. In both cases universities are excluded. Information from within the company is also excluded from breadth and depth. In order to test the validity of breadth and depth of innovation variables proposed by Laursen and Salter (2004) both variables are also split between sources 'external' and 'internal' to the industry, as explained in the theoretical section. The variable *biiuni* indicates breadth of sources internal to the industry (the sum of clients, competitors, suppliers and professional associations); while *beiuni* indicates breadth of sources external to the industry (the sum of public research institutes, universities, consultants, standards and journals). Likewise, *diiuni* and *deiuni* indicate depth of use of internal and external sources of information from universities. The empirical results of these variables will allow testing hypothesis 2[14].

A variable for whether or not companies collaborated with universities is also available for innovating companies only. This is introduced as a dummy variable called *colluni*, following D'Este (2010) and Robin et al. (2012). Robin et al. (2012), however, highlighted this variable could be endogenous. As no valid instruments are available, models with or without these variables are presented as a robustness check. A summary of all variables subject to sample selection is presented in Table 5.

[14] The way external and internal sources are defined could be seen as potentially arbitrary. The inspiring criteria have been whether the source is part of the industry being considered or whether the source is external to the industry, as discussed in the theoretical section. Exclusion of professional associations from the internal sources produced roughly similar results. However, the key point is that if all sources are homogeneous any arbitrary reclassification should yield the same results. As such the potential arbitrariness of the re-aggregation is somewhat irrelevant.

Table 4: Information on the variables:

Variable	Obs	Mean	Std. Dev.	Min	Max
logemploym	5610	4.098	1.539	2.303	> 11
exporters	4678	0.340	0.474	0.000	1
established	5610	0.071	0.256	0.000	1
incr_turnover	5610	0.075	0.264	0.000	1
decr_turnover	5610	0.065	0.247	0.000	1
nochange	5610	0.660	0.474	0.000	1
newgoodorserv	5594	0.278	0.448	0.000	1
new_to_mkt	5430	0.123	0.328	0.000	1
new_process	4792	0.183	0.387	0.000	1
proc_new_ind	4758	0.055	0.228	0.000	1
Abinc	4764	0.113	0.316	0.000	1
int_RD	4776	0.398	0.490	0.000	1
ext_RD	4756	0.146	0.353	0.000	1
anymachcomp	4746	0.540	0.498	0.000	1
ext_know	4762	0.147	0.354	0.000	1
training	4769	0.330	0.470	0.000	1
Design	4748	0.230	0.421	0.000	1
mktintroinn	4761	0.384	0.486	0.000	1
skills_science	5603	8.319	19.641	0.000	100
skills_other	5605	8.818	18.187	0.000	100
Cdominated	5104	0.093	0.291	0.000	1
Creg_UK	5096	0.085	0.279	0.000	1
Creg_EU	5092	0.070	0.256	0.000	1

Table 5 Summary of variables for motivations, and openness of innovation

Variable	Obs	Mean	Std. Dev.	Min	Max
breathuni	2079	6.743	2.634	0	9
colluniv	2484	0.154	0.361	0	1
depthuni	2079	1.552	1.579	0	9
biiuni	2078	3.470	0.995	0	4
beiuni	2059	3.306	1.832	0	5
diiuni	2078	1.158	1.039	0	4
deiuni	2059	0.398	0.807	0	5
Dnew_mkt	2060	0.344	0.475	0	1
Dcapac	2050	0.265	0.441	0	1
Dflexy	2052	0.315	0.465	0	1
Dhealth	2051	0.272	0.445	0	1
Dmkt_size	2060	0.383	0.486	0	1
Dqual	2066	0.598	0.490	0	1
Dva	2058	0.436	0.496	0	1
Doutdat	2050	0.250	0.433	0	1
Drange	2060	0.414	0.493	0	1
Dcost	2045	0.333	0.471	0	1
Denv	2055	0.227	0.419	0	1
Dreg	2054	0.406	0.491	0	1

6. Discussion of results

The sample selection ordered probit model is made of two equations: one for the value of information from the source being assessed, the other for the selection process; Table 6 present the results for the value of information from universities, which is the

model of main interest, while Annex 1 presents the results for the sample selection equation. Table 6 compares the six models discussed in the econometric section. All models include the variables for motivation[15]. Models (1) and (2) present estimates with the traditional definition of 'breadth and depth of innovation' variables. Model (1) includes collaboration with universities, while model (2) excludes it. Model (2) is closest to the original estimates presented by Laursen and Salter (2004)[16]. Models (3) and (4) instead split the breadth of innovation variables between internal and external sources. Model (3) includes the collaboration variable which is omitted from model (4). Models (5) and (6) omit all variables for 'breadth and depth of innovation' and present results, respectively, with and without the collaboration variable.

Comparing models (1) and (2) with models (3) and (4) allows to test hypothesis 2 on information sources' heterogeneity. When these variables are defined as in Laursen and Salter (2004) and are not split between 'internal' and 'external' variables, models (1) and (2) show that breadth has the expected (positive) sign and is significant, while depth is positive but insignificant. However, when the variables for breadth and depth are split between internal and external sources, as in models (3) and (4), breadth for 'internal' sources becomes negative and significant, while breadth of 'external' sources remains positive. Depth remains insignificant. This result is at odds with what would be expected from Laursen and Salter's (2004) approach. The use of the 'breadth and depth of innovation' variables proposed by Laursen and Salter would imply that *both* 'internal' and 'external' sources are positive and significant, as they implicitly assume that all variables are homogeneus (Kohler et al. 2012). Although Laursen and Salter predict that, once a certain number of sources is used, additional sources will generate decreasing returns, there is nothing in their theory that distinguishes between sources which are internal and external to the industry. The negative effect of the number of internal sources suggests that information sources are not homogeneous and casts doubts on the procedure to aggregate them into a single variable. The opposite signs of internal and external sources of information suggest that companies tend to look for information in specific direction, as proposed by Jensen et al. (2007) and Kohler et al. (2012). Some companies rely on sources from the same industrial sector, while others look for information outside of their sector. These results provide support for the assumptions underlying hypothesis 2.

Results from model (3) and (4) raise the question of whether the rejection of the sources' homogeneity hypothesis should lead to the use of separate variables for breadth and depth of 'internal' and 'external' sources. This study suggests not. The key reason is theoretical. Section 2 showed that variables for breadth and depth, generated as the sum of the other sources used by companies, are very likely to be endogenous, leading to a circular argument. A similar endogeneity issue would still apply to the proposed 'internal' and 'external' sources of information, because any sum of endogenous

[15] Removing these variables or using the constraints as exclusion restrictions, rather than the mergers and acquisition variables, does not change the conclusions on the role of breadth and depth of innovation. Results are available from the Author.

[16] It differs in that it does not include quadratic terms. However, when quadratic terms were used in the model, this produced non-sense results, with the quadratic term positive and significant and the linear term negative and significant, the opposite of Laursen and Salter original results. Results are available on request from the Author. It should also be noted that a very large share of the 'open innovation' literature which investigated the role of universities used only linear terms: while Laursen and Salter (2004) used quadratic terms, most others did not, as shown in the literature review.

variables will still be endogenous, no matter how these are aggregated. These variables were created only to test the use of the breadth and depth variables popularised by Laursen and Salter (2004), not as a valid alternative explanations.

If anything, the heterogeneity in sources of information that models (3) and (4) expose might suggests motivations are a driver of what sources are used, in accordance with hypothesis 1. If external motivations have a positive effect on the use of all external sources (including universities), then including the number of these external sources in the equation for use of universities would simply mask the effect of motivations, as companies which use a large number of external sources are also those which are driven by external motivations. Likewise, companies which are driven by internal motivation will rely mainly on internal sources. Again, including the number of internal sources would mask the effect of internal motivations.

The most obvious way to explore the possibility of common drivers behind use of internal and external sources is to look at the correlation amongst the individual sources of information used. If 'external' and 'internal' motivations are the common determinants of the use of 'external' and 'internal' sources, one would expect that use of all 'external' sources and use of all 'internal' sources are strongly correlated within each other. At the same time, use of any internal and external source should be weakly correlated. This can be investigated by creating dummy variables for whether each source of information is used or not (breadth) and by looking at their correlation patterns. If a high correlation is found within each group of 'internal' and 'external' sources, while the correlation between 'internal' and 'external' sources of information is low, this would be additional evidence that companies tend to use one or the other 'group' of information sources and that use of both sub-groups could be driven by the same underlying forces. The correlation pattern for breadth is shown in table 7. Table 7 shows that the correlation between information from clients and information from competitors is 0.64. Similarly, there is a very high correlation (0.76) between use of information from universities and use of information from public research institutes. At the same time, the correlation between universities and clients is just 0.26. Correlation between universities and suppliers is also fairly low, at 0.28. This suggests that the presence of common factors (motivations) driving use of all sources is a very real possibility.

The theoretical reasoning and the empirical evidence presented above both suggest that there could be serious problems with the breadth and depth variables. Because the number of internal and external sources of information is very likely to be endogenous and their introduction can mask the effect of motivations, models (5) and (6) dispense with them altogether[17]. The two models differ only in the inclusion or exclusion of the collaboration variable. Results from models (5) and (6) show that, once the endogenous variables are removed, the significance of motivations for innovation emerges. Reducing environmental emissions, improving health and safety, improving quality and reducing costs – all become significant determinants of the importance companies place on information from universities[18]. This is what would be expected if 'external' and

[17] An alternative approach would be to instrument these variables, but unfortunately there are no valid instruments for these variables, so this is not possible.

[18] Cost reduction and quality improvements do not necessarily fit the dichotomy between 'external' and 'internal' motivations; there can be both external pressures to reduce costs and improve quality, due to

'internal' sources masked the effect of motivations. The significance of these 'external' motivations for 'external' sources of information contrasts with the lack of significance of 'internal' motivations, such as increasing value added, market share, improving flexibility or capacity. Importantly, models (5) and (6) also show that, once the endogenous variables are removed, some other variables that were insignificant in previous models (size, export status), become significant again. These are variables that theory predicts to be significant and that, nonetheless, are found to be insignificant in 'open innovation' models. Their insignificance in models (1)-(4) is not surprising, as the inclusion of endogenous variables in a regression equation makes all estimates inconsistent; including breadth and depth of innovation causes problems on these variables through the same mechanism (endogeneity) that caused problems to the motivation variables. Model (5) and (6) therefore confirm previous findings from the 'resource view' literature. Their results are broadly similar: the percentage of employees with a degree in science is positively associated with the perceived usefulness of information universities. The study also confirms differences amongst industrial sectors but, due to space limitations, these are neither presented in the tables, nor discussed[19]. Start-ups do not appear to value information from research more than established companies, contrary to common belief but in line with most findings. Constraints on innovation are not found to be a significant explanation of the use of information from universities. Radical innovators are not more likely to benefit from information generated from the public research base than other companies. The result is slightly unexpected, but it depends on the use of the conventional 5% significance level; if a 10% significance level were considered, the positive sign would become statistically significant again[20]. One difference between model (5) and (6) relates to innovation activities: model (5) shows internal R&D is important, while model (6) emphasises external R&D and acquisition of external knowledge. This is expected, as model (5) captures the contribution of external sources through the collaboration with universities, while model (6) does not.

All of the results presented are from the ordered probit sample selection model. Hence, it is useful to discuss whether there is any evidence of sample selection. The econometric section explained that, if $\rho = 0$, then there is no sample selection and the model can be simplified to a simple ordered probit. There is no evidence of sample selection in any of these models: the relevant parameter is *atrho* and it is never significantly different from zero. This result is similar to Mohnen and Hoareau (2003), who also did not find evidence of sample selection. It would therefore be possible to estimate simple ordered probit models over the same samples.

Overall, removing the 'openness for innovation' variable allows the analysis to re-discover many results that were established in the early literature and have been largely neglected by the 'open innovation' literature[21]. More importantly, it allows shedding

---

competition, and internal ones, due to increased profitability; however, it is plausible that the external element dominates.

[19] Results are available on request from the author.

[20] It is therefore likely that the weaker than usual statistical significance is the result of introducing a number of distinctions into the analysis, which differentiates between product and process innovation and between radically new process and product innovations.

[21] Omission of relevant variables and the endogeneity of the breadth and depth variables are the most likely explanations for the lack of significance of these 'traditional' variables in the literature following Larsen and Salter (2004).

light on new powerful explanations of the value companies place on information from universities, focusing on the heterogeneity of the motivations to innovate for information sourcing. Innovation motivations also help to make sense of the large gap between the ratings of 'internal' and 'external' sources of information: 'external' motivations for innovation obviously occur less frequently than 'internal' motivations; it is therefore not surprising that external sources are cited as useful by a lower percentage of companies. The results also suggest that it is not the number of information sources used by a 'representative' company that matters most for the choice of information sources. Rather, it is the heterogeneity of companies that matters: as companies have different needs and motivations and are different in their characteristics (size, export status, etc), they turn to the information sources which are most suited to them, with significant differences between companies that source their information from within the industry and those which rely mostly on external sources[22].

[22] The validity of this approach was also tested by running a similar regression on clients, which is not included due to space limitations. The results showed that, for clients, internal sources have a positive effect, while external sources have a negative one. Also, innovation motivations for clients are mainly internal ones. This is coherent with hypotheses 1 and 2a.

Table 6: Results for information from universities, outcome equation

info_univrs	(1)	(2)	(3)	(4)	(5)	(6)
logemploym	0.005	0.011	0.005	0.012	0.066**	0.071**
	(0.025)	(0.024)	(0.025)	(0.025)	(0.023)	(0.023)
exporters	0.099	0.133	0.090	0.121	0.161*	0.175**
	(0.075)	(0.073)	(0.076)	(0.075)	(0.066)	(0.065)
established	-0.168	-0.050	-0.137	-0.023	-0.180	-0.092
	(0.129)	(0.125)	(0.131)	(0.128)	(0.112)	(0.110)
int_RD	0.085	0.072	0.049	0.037	0.192*	0.176
	(0.110)	(0.107)	(0.113)	(0.111)	(0.098)	(0.105)
ext_RD	0.098	0.245**	0.067	0.203*	0.070	0.201**
	(0.085)	(0.082)	(0.086)	(0.083)	(0.076)	(0.074)
anymachcomp	-0.078	-0.062	-0.063	-0.043	0.091	0.090
	(0.089)	(0.087)	(0.090)	(0.088)	(0.078)	(0.078)
ext_know	0.022	0.068	0.024	0.066	0.113	0.149*
	(0.083)	(0.081)	(0.084)	(0.082)	(0.074)	(0.073)
training	-0.078	-0.026	-0.082	-0.033	-0.020	0.026
	(0.080)	(0.078)	(0.081)	(0.079)	(0.071)	(0.071)
design	-0.129	-0.098	-0.111	-0.081	-0.038	-0.017
	(0.081)	(0.079)	(0.082)	(0.080)	(0.073)	(0.074)
mktintroinn	-0.223	-0.197	-0.224	-0.193	0.038	0.066
	(0.125)	(0.121)	(0.128)	(0.125)	(0.116)	(0.132)
Cdominated	-0.021	0.018	-0.026	0.014	0.000	0.034
	(0.115)	(0.112)	(0.116)	(0.114)	(0.104)	(0.104)
Creg_UK	0.091	0.158	0.105	0.166	0.198	0.248
	(0.162)	(0.157)	(0.164)	(0.160)	(0.148)	(0.145)
Creg_EU	0.046	0.033	0.015	0.006	-0.025	-0.026
	(0.174)	(0.168)	(0.176)	(0.170)	(0.159)	(0.156)

Table 6: Results for information from universities, outcome equation (continued)

info_univrs	(1)	(2)	(3)	(4)	(5)	(6)
newgoodorserv	-0.106	-0.073	-0.115	-0.086	-0.055	-0.041
	(0.087)	(0.085)	(0.088)	(0.087)	(0.075)	(0.074)
new_process	-0.062	-0.096	-0.044	-0.075	-0.106	-0.132
	(0.080)	(0.078)	(0.081)	(0.080)	(0.071)	(0.070)
new_to_mkt	0.084	0.124	0.097	0.139	0.106	0.141
	(0.085)	(0.083)	(0.086)	(0.084)	(0.076)	(0.074)
proc_new_ind	0.036	0.125	0.026	0.108	0.098	0.176
	(0.113)	(0.109)	(0.114)	(0.111)	(0.103)	(0.100)
skills_science	0.004**	0.006***	0.003*	0.005**	0.005***	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
skills_other	0.003	0.003*	0.004	0.004*	0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
finsupp	0.108	0.321***	0.106	0.308***	0.216**	0.401***
	(0.091)	(0.086)	(0.091)	(0.087)	(0.082)	(0.079)
breathuni	0.553***	0.538***				
	(0.027)	(0.025)				
depthuni	0.008	0.016				
	(0.024)	(0.023)				
biiuni			-0.152*	-0.236**		
			(0.076)	(0.074)		
beiuni			0.780***	0.793***		
			(0.042)	(0.040)		
diiuni			-0.045	-0.051		
			(0.039)	(0.038)		
deiuni			0.085	0.108*		
			(0.046)	(0.045)		

Table 6: Results for information from universities, outcome equation (continued)

info_univrs	(1)	(2)	(3)	(4)	(5)	(6)
colluniv	1.215***		1.141***		1.082***	
	(0.091)		(0.092)		(0.081)	
Drange	0.114	0.076	0.113	0.079	0.136*	0.110
	(0.074)	(0.072)	(0.075)	(0.073)	(0.066)	(0.064)
Dnew_mkt	0.008	0.012	0.032	0.041	0.112	0.118
	(0.074)	(0.072)	(0.075)	(0.073)	(0.066)	(0.065)
Dmkt_size	-0.048	-0.063	-0.045	-0.060	-0.059	-0.066
	(0.074)	(0.072)	(0.075)	(0.073)	(0.066)	(0.065)
Dqual	0.096	0.077	0.115	0.101	0.154*	0.136*
	(0.079)	(0.077)	(0.080)	(0.078)	(0.069)	(0.068)
Dflexy	-0.086	-0.053	-0.130	-0.104	-0.014	0.012
	(0.081)	(0.079)	(0.083)	(0.081)	(0.073)	(0.072)
Dcapac	0.173*	0.135	0.205*	0.172*	0.091	0.066
	(0.083)	(0.081)	(0.084)	(0.082)	(0.074)	(0.073)
Dva	0.063	0.062	0.092	0.094	-0.007	-0.001
	(0.074)	(0.072)	(0.075)	(0.073)	(0.065)	(0.064)
Dcost	0.002	0.009	-0.014	-0.003	0.172*	0.175**
	(0.076)	(0.074)	(0.077)	(0.076)	(0.067)	(0.066)
Dhealth	0.058	0.106	0.047	0.094	0.129	0.168*
	(0.095)	(0.093)	(0.097)	(0.095)	(0.086)	(0.084)
Denv	0.395***	0.389***	0.345***	0.334***	0.311***	0.308***
	(0.094)	(0.092)	(0.096)	(0.094)	(0.085)	(0.083)
Doutdat	0.038	0.001	0.027	-0.013	-0.019	-0.040
	(0.084)	(0.082)	(0.085)	(0.083)	(0.075)	(0.074)
Dreg	-0.072	-0.057	-0.076	-0.063	0.045	0.060
	(0.081)	(0.079)	(0.082)	(0.081)	(0.072)	(0.071)

Table 7: correlation between use of source of information (breadth)

	dclients	Dcompet	dconf	dconsult	djournal	dprof_ass	Dpublic	dstandard	Dsuppl	duniv	Dwithin
dclients	1.00										
dcompet	0.64	1.00									
dconf	0.31	0.41	1.00								
dconsult	0.28	0.33	0.46	1.00							
djournal	0.28	0.34	0.59	0.49	1.00						
dprof_ass	0.35	0.41	0.58	0.42	0.60	1.00					
dpublic	0.23	0.30	0.55	0.53	0.60	0.49	1.00				
dstandard	0.38	0.43	0.50	0.45	0.61	0.64	0.47	1.00			
dsuppl	0.48	0.45	0.33	0.37	0.35	0.36	0.27	0.41	1.00		
duniv	0.21	0.29	0.53	0.59	0.58	0.42	0.76	0.44	0.28	1.00	
dwithin	0.55	0.49	0.28	0.30	0.27	0.31	0.22	0.33	0.54	0.23	1.00

7. Conclusions

The analysis showed that some of the variables currently used to explain the importance of information from universities are problematic, from a theoretical and empirical perspective. Variables such as breadth and depth of innovation are affected by endogeneity problems and end up obscuring important determinants of the value of information from universities. The analysis also showed that company's characteristics and motivations are more important than the number of information sources used. Relevant characteristics include both companies' features (size, innovation activities, absorptive capacity and export status) and innovation motivations. External motivations to innovate, such as reducing pollution and improving health and safety emerged as one of the key determinants of which companies rely on universities.

These findings have important managerial and policy implications: the number of information sources previously used in 'open innovation' suggested that managers should widen the number of sources, to make the most of each individual source. However, the present study found there is no reason to do so. Using too many sources is likely to unjustifiably increase the costs companies incur, as they need to integrate information from a wide variety of sources. Moreover, focussing on the number of sources suggests these are homogeneous and interchangeable, while they are not. Homogeneity obscures the fact that 'internal' (industry) and 'external' (public) sources of information each cater for specific needs and that universities provide information which could not readily be substituted by resorting to other sources.

Overall, the study emphasised the importance of diversity *amongst* companies, rather than diversity in information sources *within a 'representative' company*, which is a strongly evolutionary message.

While the study showed that universities do not necessarily play a direct role for *all* companies, they remain important to the whole economic system. Universities provide information to large, export-oriented, innovative and highly-skilled companies and help companies to cope with pressures for change generated by society. Acknowledging the heterogeneity amongst information sources shows universities have a unique role to play in the innovation system.

The analysis also 're-discovered' the importance of companies' characteristics (size, export status, etc) to the use of information from universities, as proposed by the 'resource view' of the firm. This result was obscured by the introduction of 'open innovation' variable. These results also carry important policy implications: if the usage of information from universities depends on firms' characteristics, then the impact of academia no longer depends only on the quality of research, but also on the skills, activities and motivations of the companies which may use it. Because companies' characteristics and motivations have an important role, the uptake of information from universities can be increased via the regulatory landscape (especially environmental policy), as well as by increasing their absorptive capacities (Cohen and Levinthal, 2004).

In closing, a further development is suggested. The study is based on a cross section. Therefore, it would be interesting to extend the analysis to a panel setting, which would enable to draw conclusions on causality rather than correlations. This will be left to future work.

Acknowledgements

The analysis has been developed during my secondment to the Department for Business, Innovation and Skills (BIS) and has been developed to inform the knowledge exchange strategy of NERC. Financial support from BIS and NERC is gratefully acknowledged. The data from the UK Innovation Survey have been made available to me initially by the Virtual Micro-data Laboratory at the Office for National Statistics and later by the Secure Data Service, project number 53841. The use of these data does not imply the endorsement of the data owner or the UK Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. The paper has benefited from comments from participants at the ENEF 2012 conference in Bologna, Keith Smith, Rosa Fernandez, Massimiliano Bratti and Phil Heads and from editorial help by Thomas Marshall. As usual, all remaining mistakes are my sole responsibility.

Appendix 1: Results for information from universities, selection equation

sel eq univ.	(1)	(2)	(3)	(4)	(5)	(6)
logemploym	0.115***	0.115***	0.115***	0.115***	0.116***	0.116***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
exporters	0.069	0.069	0.069	0.069	0.070	0.070
	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
established	-0.032	-0.032	-0.030	-0.032	-0.032	-0.032
	(0.087)	(0.087)	(0.087)	(0.087)	(0.087)	(0.087)
incr_turnover	0.305***	0.305***	0.301***	0.302***	0.305***	0.305***
	(0.088)	(0.088)	(0.088)	(0.088)	(0.088)	(0.088)
decr_turnover	0.087	0.088	0.087	0.088	0.084	0.087
	(0.088)	(0.088)	(0.088)	(0.088)	(0.089)	(0.089)
int_RD	0.534***	0.534***	0.535***	0.535***	0.534***	0.534***
	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)
ext_RD	0.218*	0.218*	0.219*	0.219*	0.215*	0.215*
	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)	(0.085)
anymachcomp	0.215***	0.215***	0.215***	0.215***	0.214***	0.214***
	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)
ext_know	0.012	0.012	0.013	0.012	0.015	0.014
	(0.081)	(0.081)	(0.081)	(0.081)	(0.081)	(0.081)
training	0.223***	0.222***	0.224***	0.222***	0.221***	0.221***
	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)
design	0.402***	0.403***	0.401***	0.401***	0.402***	0.403***
	(0.070)	(0.070)	(0.070)	(0.070)	(0.070)	(0.070)
mktintroinn	0.888***	0.888***	0.889***	0.889***	0.888***	0.888***
	(0.059)	(0.059)	(0.058)	(0.058)	(0.059)	(0.059)

Appendix 1, continued: Results for information from universities, selection equation

sel eq univ.	(1)	(2)	(3)	(4)	(5)	(6)
skills_science	0.002	0.002	0.002	0.002	0.002	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
skills_other	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cdominated	-0.350***	-0.350***	-0.351***	-0.351***	-0.350***	-0.350***
	(0.084)	(0.084)	(0.084)	(0.084)	(0.084)	(0.084)
Creg_UK	0.120	0.119	0.124	0.123	0.119	0.119
	(0.147)	(0.147)	(0.147)	(0.147)	(0.147)	(0.147)
Creg_EU	0.190	0.190	0.183	0.184	0.193	0.193
	(0.159)	(0.159)	(0.159)	(0.159)	(0.159)	(0.159)
cons	-1.605***	-1.604***	-1.605***	-1.603***	-1.608***	-1.608***
	(0.092)	(0.092)	(0.092)	(0.092)	(0.092)	(0.092)
cut1	4.193***	4.182***	2.296***	2.131***	1.069**	1.122**
	(0.456)	(0.427)	(0.473)	(0.449)	(0.332)	(0.392)
cut2	6.183***	5.973***	4.356***	4.013***	2.584***	2.497***
	(0.496)	(0.456)	(0.517)	(0.481)	(0.344)	(0.398)
athrho	-0.206	-0.165	-0.257	-0.211	-0.071	-0.020
	(0.192)	(0.183)	(0.201)	(0.194)	(0.176)	(0.219)
N	4044.000	4044.000	4044.000	4044.000	4044.000	4044.000
ll	-2889.548	-2985.515	-2845.911	-2929.836	-3313.089	-3405.194
chi2	754.456	674.980	773.751	715.095	475.373	331.401
p	0.000	0.000	0.000	0.000	0.000	0.000
bic	6426.886	6610.514	6356.222	6515.767	7257.358	7433.262
aic	5935.097	6125.030	5851.822	6017.673	6778.179	6960.387
rho	-0.203	-0.163	-0.252	-0.208	-0.071	-0.020
p_c	0.338	0.429	0.264	0.350	0.688	0.929

References

Antonioli D., Mancinelli S., Mazzanti M., 2013 Is environmental innovation embedded within high-performance organisational changes? The role of human resource management and complementarity in green business strategy. Research Policy, in press.

Arvanitis S., 2012 How do different motives for R&D cooperation affect firm performance? An analysis based on Swiss data. Journal of Evolutionary Economics 22, 981-1007

Arvanitis S., Lokshin B., Mohnen P., Worter M., 2013. Impact of external knowledge strategy on innovation: a comparative study based on Dutch and Swiss Panel Data. UNU-Merit Working Paper 2013-03.

Belderbos R., Carree M., Lokshin B., 2004. Cooperative R&D and firm performance. Resarch Policy 33, 1477-1492.

Bratti M. and Felice G., 2011. Are exporters more likely to introduce product innovations? Centro Studi Luca D'Agliano Working Papers.

Bruneel, J. D'Este, P. Salter, A., 2010 Investigating the factors that diminish the barriers to university-industry collaboration. Research Policy 39, 858-868.

Cainelli G., Mazzanti M., Montresor M., 2012 Environmental innovation, local networks and internationalization. Industry and Innovation 19(8), 697-734.

Chesbrough H.W., 2003 The era of open innovation. MIT Sloan Management Review 44(3), 34-41.

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

Cohen, W.M. Nelson, R.R. Walsh J.P. 2002. Links and impact: the influence of public research and industrial R&D. Management Science 48(1), 1-23.

De Luca G. and Perotti V., 2011. Estimation of ordered response models with sample selection. The Stata Journal 11(2), 213-239.

De Marchi V., 2011 Environmental innovation and R&D cooperation: empirical evidence from Spanish manufacturing firms. Research Policy, in press.

Fitjar R.D. and Pose A.R., 2013. Firm collaboration and modes of innovation in Norway. Research Policy, 42(1), 128-138.

Gonzales G.S., Alvarez N.G., Nieto M., 2012. Sticky information and heterogeneous needs as determining factors of R&D cooperation with customers. Research Policy 38, 1590-1603.

Guellec D. and van Pottelsberghe de La Potterie B., 2001. R&D and productivity growth: panel data analysis of 16 OECD countries. OECD Economic Studies.

Hendry, D. F., 1995. Dynamic Econometrics, Oxford University Press, Oxford

Horbach J., 2008. Determinants of eco-innovation- new evidence from German panel data sources. Research Policy 37, 163-173.

Horbach J., Rammer C., Rennings K., 2012. Determinants of eco-innovation by type of environmental impact – The role of regulatory push/pull, technology push and market push. Ecological Economics 78, 112-122.

Huang C., Arundel A., Hollanders H., 2010 How firms innovate: R&D, non-R&D and technology adoption. UNU_MERIT working paper 027.

Jaffe A.B., Newell R., Stavins R.N., 2005. The tale of two failures: technology and environmental policy. Ecological Economics, 54: 164-174.

Jensen M.B., Johnson B., Lorenz E., Lundvall B.A., 2007. Forms of knowledge and modes of innovation. Research Policy 36, 680-693.

Kesidou E. and Demirel P., 2012. On the drivers of eco-innovation: empirical evidence for the UK. Research Policy 41, 862-870

Kohler C., Sofka W., Grimpe C, 2012. Selective search, sectoral pattern, and the impact of product innovation performance. Research Policy 41, 1344-1356.

Laursen K. and Salter A., 2004. Searching high and low: what types of firms use universities as a source of innovation? Research Policy 33, 1201-1215.

Leiponnen A. Helfat C.E., 2010. Innovation objectives, knowledge sources and the benefits of breadth. Strategic Management Journal 31, 224-236.

Mazzanti M. and Costantini V., 2012. On the green and innovative side of trade competitiveness? The impact of environmental policies and innovation on EU exports. Research Policy, 41, 132-153.

Miotti L. and Sachwald F. 2003. Cooperative R&D: why and with whom? An integrated framework of analysis. Research Policy, 1481-1499.

Mohnen P. and Hoareau C., 2003.What type of enterprise forges close links with universities and government labs? Evidence from CIS2. Managerial and Decision Economics 24(2-3), 133-145

Mowery D.C. and Sampat B.N., 2005.Universities in national innovation systems. In The Oxford handbook of innovation. Oxford University Press.

Nelson R.R. and Winter S.G., 1982. An Evolutionary Theory of Economic Change. Harvard University Press

Pellegrino G. and Savona M., 2013. Is money all? Financing versus knowledge and demand constraints to innovation. UNU-Merit working paper 2013-029.

Phillips G. and A. Zhdanov, 2012.R&D and the incentive from mergers and acquisition activity.National Bureau of Economic Research WP 18346.

Robin S., Schubert T., 2013. Cooperation with public research institutions and success in innovation: evidence from France and Germany. Research Policy 42, 149-166.

Romer P., 1990. Endogenous Techical Change, The Journal of Political Economy 98(5) S71-102

Salter A.J., Martin B.R., 2001. The economic benefits of publicly funded basic research: a critical review. Research Policy 30, 509-532.

Snyder L.R, Miller N.H. and Stavins R.N., 2003. The effects of environmental regulation on technology diffusion: the case of chlorine manufacturing. American Economic Review 93(2), 431-435.

Tether B.S. and Tajar A., 2008. Beyond industry-university links: sourcing knowledge for innovation from consultants, private research organisations and the public science-base. Research Policy 37, 1079-1095.