Study on Higher Education Institutions and Local Development

A. Bonaccorsi
D. Biancardi
M. Sánchez-Barrioluengo
F. Biagi

2019
Abstract ................................................................................................................................. 3
Acknowledgements .............................................................................................................. 4
1. Introduction ..................................................................................................................... 5
2. Survey of the literature ................................................................................................. 7
   2.1 Pathways through which HEIs impact on the economy ......................................... 7
      2.1.1 Higher education and the creation of human capital ..................................... 8
      2.1.2 The impact of universities through research activities ............................. 11
      2.1.3 Academic entrepreneurship and startup creation ...................................... 14
      2.1.4 Attractiveness for foreign investment ....................................................... 15
      2.1.5 High value added procurement .................................................................. 15
      2.1.6 Cultural and social externalities ............................................................... 15
      2.1.7 University expenditure ............................................................................ 16
   2.2 Contextual factors moderating the impact of university education and research . 17
      2.2.1 Complementarity between formal education and cognitive and non-cognitive skills ........................................................................................................ 17
      2.2.2 Matching between the competences of graduates and the demand of firms 18
      2.2.3 Quality of education and university-work transition .................................. 20
      2.2.4 Graduate migration ................................................................................... 20
      2.2.5 Quality of academic research .................................................................... 22
      2.2.6 Regional pattern of specialization and absorptive capacity ...................... 24
      2.2.7 Complementarity and substitution effects between the different pathways .. 26
      2.2.8 Complementarity between human capital accumulation and innovation .... 28
3. Empirical research strategy ............................................................................................ 30
   3.1 Research questions ................................................................................................... 30
   3.2 Description of data: ORBIS, ETER ...................................................................... 31
      3.2.1 Orbis dataset ........................................................................................... 31
      3.2.2 ETER ....................................................................................................... 36
   3.3 Description of main variables and methodological issues ....................................... 36
      3.3.1 Independent variables ............................................................................. 36
      3.3.2 Dependent variables ............................................................................... 39
      3.3.3 Control variables ..................................................................................... 40
   3.4 Treatment of spatial effects ..................................................................................... 42
4. Main results ..................................................................................................................... 45
   4.1 Models of growth of firm performance and aggregate independent variables ...... 45
   4.2 Models of growth of firm performance. (a) models with control variables ........ 49
      4.2.1 General results ......................................................................................... 49
Abstract

In this report we first carry out an extensive survey of the literature that is relevant for the analysis of the regional impact of Higher Education Institutions (HEIs). Based on the literature we specify some broad research questions and we develop a battery of testable hypotheses. We then present the details of the data used in the empirical analysis. Summary statistics for the available datasets and a preliminary discussion of the econometric approach are also presented.

In the second part of the Report we describe the structure of models and show the findings of several alternative specifications. We discuss the main results and illustrate further opportunities for research and policy making.
Acknowledgements

This report has been done in the framework of the Knowledge-Hub of Higher Education, an Administrative Agreement between DG EDUCATION and CULTURE (EAC) and DG JOINT RESEARCH CENTRE (JRC). Its content was developed under the Service Contract number 934514-2017 A08-IT with FBK-IRVAPP (Trento, Italy).
1. Introduction

It is widely recognized that the creation of human capital, mainly at tertiary level, is a condition for economic growth and local and regional development. It is also recognized that knowledge spillovers originated in the research sector generate new ideas and opportunities, fuelling the entrepreneurial process and supporting the innovation activities of firms.

Yet most of these conclusions have been reached through aggregate macroeconomic analyses at country and regional level, as well as through surveys at firm level.

In this Report we investigate the impact of human capital formation and knowledge spillovers by assuming as unit of analysis the individual Higher Education Institution (HEI). We examine the extent to which HEIs generate, directly or indirectly, positive impacts on all firms located in their surroundings. Our approach is innovative as we integrate various datasets, solving several conceptual and technical issues when merging heterogeneous sources.

We start by measuring how many graduates (from bachelor up to doctoral), and academic staff attend or are employed at each individual institution, by field of education, in the year 2015. For this we use a census of all HEIs in Europe, called ETER. We then integrate this dataset with scientific publications from the Scopus database, and measure how many publications are produced at each institution and how many citations they receive, by scientific discipline, in the time interval 2011-2015.

By combining these two sources we finally obtain a full picture of the sources of human capital at tertiary level and of research outputs. We geo-referentiate this data at NUTS 2 and NUTS 3 level. This gives us a rich representation of the flows that are generated by the higher education sector at local and regional level.

We then ask whether these flows have an impact on those firms that are located close to HEIs. For this we geo-referentiate all firms included in the ORBIS database across almost all European countries. We then calculate the geographic distance, using GIS coordinates, between firms and HEIs and define a radial structure, from 10km to 100km distance.

Finally, we use a large number of control variables from Eurostat, OECD and other sources.

Having integrated all these datasets, we run a large number of regression models.

The main results can be summarized as follows.

The covariation of firms’ performance and activities carried out by higher education institutions on firms is confirmed beyond any reasonable doubt. After controlling for many structural factors (industry, size of the firm, legal form) and location factors (country, regions) it appears that HEIs produce outputs, in terms of human capital (graduates) and knowledge (publications), that in the large majority of cases positively influence firms located in the neighbourhood. This effect is even more remarkable since it comes from flows (yearly production of graduates and publications), that is, from relatively short term phenomena, not from established, slowly created stocks.

At the same time, the impact is strongly dependent on specific disciplines, or bases of knowledge, with Engineering and Business showing the largest impact, Basic Sciences and Medicine mixed results, and Social Sciences and Humanities a positive impact only in joint models of human capital and research. This is another important result, which confirms the notion that knowledge is heterogeneous and follows different pathways to become productive in the economic system. At the same time, this finding sheds light on the necessity to consider a variety of impact models. Assuming an impact model for which the only dependent variables to be considered are those that refer to firms (whatever the dimension of performance) can be severely misleading.
The report is structured as follows. Section 2 offers a survey of the literature on the impact of higher education on regional and local development.

Section 3 discusses the main research questions, develop specific testable hypotheses, introduce the data sources and describes the variables extracted from the dataset. It also presents the main features of the ETER data, a list of the variables that have been used in the study and a preliminary analysis of these variables. Furthermore this Section identifies the sources of information that have been integrated into the main dataset.

Section 4 describes the structure of the regression models used in the empirical analysis. Section 5 presents the main results. Section 6 summarizes the results and concludes.
2. Survey of the literature

The idea that the knowledge produced by universities has a measurable impact on the wider context of economy and society has a long tradition and has been empirically confirmed many times (Mansfield 1991; 1995; Feldman, 1994; Varga, 1998; Arbo and Benneworth, 2007; Geiger and Sá, 2008; Kenney and Mowery, 2014). An influential literature has addressed the challenge of producing econometric estimates of the impact (Anselin et al. 1997).

Governments and policy makers at national, regional and local level are keen to reap the benefits of higher education for the economy (Goldstein and Renault, 2004; Drucker and Goldstein, 2007; Drucker, 2016). Especially in less developed regions this has taken the form of a rush for creating and supporting higher education institutions, as well as of policies aimed at establishing intense relations between HEIs and the economic and social environment.

There has been a growing interest in measuring the impacts of higher education on regional economies. However, evaluating the socio-economic impacts of universities is complex and the research literature lies at the intersection of different fields such as higher education, regional studies, urban economics, economic geography and policy studies. The literature has followed a large variety of directions.

In terms of empirical methods, the literature on HEIs and local development is quite heterogeneous (Drucker and Goldstein, 2007). Arbo and Benneworth (2007) examine the literature in economic geography and regional economics. Drucker and Goldstein (2007) offer a review of the studies from a methodological perspective. Considerable attention is paid to the methodological advantages and shortcomings of four major research designs evidenced in the literature:

— single-university impact studies
— surveys
— knowledge production functions
— cross-sectional and quasi-experimental designs.

Among these approaches, it can be said that the more popular approaches are the knowledge production functions, associated to standard cross-sectional econometric models. Time series are seldom used. There is an increasing adoption of spatial econometric techniques. A quasi-experimental design (in particular, difference-in-difference, regression discontinuity design, or synthetic control) is followed in the most recent literature, although with a small number of studies and with data from single countries or regions. Single-university impact studies are not generally considered valid (see below on models of impact based on expenditure), while surveys are largely used in order to address the need for data on flows (for example, migration or university-work transition), institutional features (for example, university-industry contractual relations), but not for the estimation of impact.

We structure the survey as follows. First, we identify the specific pathways through which HEIs impact on the economy. Second, we examine the factors that influence the impact (directly or indirectly, positively or negatively), that are related to, respectively, universities, firms, or their external context (at regional and national level).

2.1 Pathways through which HEIs impact on the economy

Universities can affect local development through several distinct channels. The recent literature has done a remarkable effort in identifying specific ways in which HEIs can impact the economy.
2.1.1 Higher education and the creation of human capital

To start with, HEIs increase the level of education of the region in which they are located and in this way contribute greatly to the creation of human capital, in general, and for the available workforce. In other words, HEIs contribute to the increase of the average educational attainment of a population. Although this variable is not synonymous of skills (see below), it is crucial for the process of economic growth. A large and consistent literature has shown a strong positive relation between the level of human capital and the absolute performance and rate of growth of countries, regions, and cities (Rauch, 1993; Henderson et al. 1995; Glaeser et al. 1995, Eaton and Eckstein 1997, Black and Henderson 1999, Glaeser and Saiz 2004; Rodriguez-Pose and Vilalta-Bufi 2005). Sianesi and Van Reenen (2003) provide a summary of the empirical evidence on the effects of education on economic growth.

This literature has gained prominence with the endogenous growth theory (Lucas, 1988; Romer, 1990), according to which the accumulation of human capital produces positive externalities and contributes to the growth of productivity, as well as to the development of new products and processes (Mankiw et al. 1992; Barro, 1991).

What are the specific mechanisms at work?

First, workers with higher educational qualifications are able to perform work activities with higher complexity. This is an established fact in labor economics and in official statistics. The International Standard Occupational Classification (ISOC), managed by the International Labor Organization (ILO) classifies work activities on the basis of their complexity and allows the comparison of national statistics. For example, the US Bureau of Labor Statistics maintains an Employment Requirement Table that shows the Training requirement for hundreds of work positions. All data tell that graduates occupy the work positions with the highest levels of autonomy, self-direction and ability to cope with complexity.

Second, by hiring graduates firms get access to knowledge that would not be available via other means (e.g. by investing in equipment developed by other companies). In some industries this is simply a precondition for operations. Hiring engineers or science graduates means getting access to entire branches of knowledge that contribute to the creation of value. In other words, there is a strong linkage between human capital and innovation, a phenomenon already discussed by Nelson and Phelps (1966).

Third, by hiring graduate staff companies generate internal spill-overs of knowledge. Graduate staff increases the productivity of non-graduate staff by organizing the work and pushing production towards higher value added activities.

Fourth, by hiring graduate staff companies get access to the knowledge stored in the university in which they took a degree. This is a highly informal kind of impact, but one most effective in many cases. After graduation, students work in companies but keep some informal relations with their former professors. Very often professors themselves are keen to maintain informal relations with their former students. When opportunities arise and companies have the need to access external knowledge, they may leverage on these relations instead of entering into a formal process of scouting and selection of universities.

A related issue is whether education contributes to firm performance even if students do not complete their degree. In some cases students undertake courses and pass exams but do not finish their career (or finish it late, after years of employment). We suggest that this process of incomplete human capital creation may be valuable, insofar as students are exposed to forms of higher level knowledge (that is, codified, abstract and general), although in partial ways.

The strong positive effect of human capital on productivity and growth is not without qualifications. As we will examine more precisely below, the positive relation is moderated by some contextual factors.
To what extent does this general effect depend on the proximity of universities? In other words, do local economic systems, at lower level of geographic disaggregation (i.e. regions and cities) benefit from having universities at a close distance?

Several studies confirm that the positive effect has a spatial dimension, confirming positive and significant effects of university proximity on workers’ productivity and local development.

These studies examine the regional impact of human capital with an aggregate approach, by taking the proportion of active population with tertiary education as explanatory variable. This is a standard approach in the human capital literature, given the lack of microdata on HEIs at regional level. It is important to remark the difference between this literature and the approach followed in this study. In this literature human capital is proxied by the average educational attainment of the population and the relationship between human capital and growth is typically aggregate. This means that:

— the educational attainment of the population is the result of many disparate flows at any given point in time (e.g. educated or less educated workforce entering the pool of workers, workforce retirement, immigration and outmigration of workers);

— the relation between the stock of education of population and the flow from higher education is left unobserved.

In our approach we examine the flows of human capital, in the form of graduates from HEIs located in a geographic space. We also examine the impact of the stock of highly qualified human capital represented by academic staff. To these dimensions we add the flows of production of scientific research. We do all this using microdata, which allow an unprecedented level of granularity in spatial analysis. The main limitation of our microdata approach is that we cannot specify the proportion of students or graduates who migrate in order to work in a different territory than the one in which they have studied.

With these clarifications is nevertheless useful to review briefly some results of the literature using the stock of human capital as the main independent variable.

Fischer et al. (2009) examines productivity data across European regions and find a strong effect of human capital. Acemoglu and Dell (2010) show that differences in human capital account for half of between-municipality differences in output. Gennaioli et al. (2013) also find that education is the most important determinant of regional income and productivity.

Buendia-Azorin and del Mar Sanchez-de la Vega (2015) examine the relation between human capital and labor productivity for European regions in the 2000-2009 period. Labor productivity is defined as Gross Value Added per worker, while human capital is proxied by the share of active population with the first, second and third stages of higher education (International Standard Classification of Education ISCED, levels 5–8). Using a spatial econometric approach (SDM, spatial Durbin model) they examine the impact on productivity of spillovers from spatially contiguous regions. They find a strong positive effect on labor productivity: an increase of 10% in human capital in a region has an average direct impact of 3.0% in close regions.

Other studies, on the other hand, explicitly examine the presence of universities in the local and regional context (Beason and Montgomery, 1993; Andersson et al., 2004; Liu, 2015; Abel and Deitz, 2012; Benander et al. 2016; Pfister et al. 2017).

Andersson et al. (2004) investigate the economic effects of the HEI decentralization policy – i.e. the creation of university sites or branches at dispersed geographical locations - on productivity and output per worker. They find important and significant effects of this policy upon the average productivity of workers, suggesting that the economic effects of decentralization on regional development are economically important. Liu (2015) presents evidence of spillovers from universities and examines the short- and long-run effects of university activities on geographic clustering of economic activity,
labor market composition and local productivity. He treats the designation of land-grant universities (1) as a natural experiment after controlling for the confounding factors with a combination of synthetic control methods and event-study analyses. Three key results are obtained. First, the designation substantially increases local population density. Second, the share of manufacturing workers in the population, an indicator of labor market composition, is not affected by the designation. Third, the designation greatly enhances local manufacturing productivity, as captured by local manufacturing output per worker, especially in the long run.

Abel and Deitz (2012) examine the spatial distribution of universities in the USA and the link with the occupational composition of the industry. They find a strong positive relationship between a metropolitan area’s higher education activities and the share of workers in high human capital occupations. This relation is stronger when higher education activities are produced by institutions with high research intensity as measured by the Carnegie classification.

Bonander et al. (2016) study the impact on the regional economy of granting research university status to three former university colleges in three different regions in Sweden. Granting the status of university does not mean a formal attribution of credentials but the admission to research funding and the creation of attractiveness for students. They analyze the development in the treated regions compared to a set of control regions that are created using the synthetic control method. They find small or no effects on the regional economy, casting doubts on the effectiveness of research universities in fostering regional growth and development. They claim that a more credible identification strategy in assessing the effects of universities on the regional economy should be used, as opposed to what has been used in previous studies. Given the specificity of the legal framework it is difficult to generalize from these results.

Besides the systematic evidence of the impact of HEIs on the creation of human capital at regional and local level, it has been found that the presence of HEIs positively affects wages of workers with a lower level of education following a strong spill-over effect. Moretti (2004) estimates spill-overs from college education by comparing wages for otherwise similar individuals who work in cities with different shares of college graduates in the labour force and finds that a percentage point increase in the supply of college graduates raises high school drop-outs’ wages by 1.9%, high school graduates’ wages by 1.6%, and college graduates wages by 0.4%.

This happens because the wage of educated and uneducated workers is affected by the share of college graduates in two different ways. For both categories of workers there is a positive spillover effect similar to the one discussed in the literature on human capital: “human capital externalities arise because workers learn from each other, and they learn more from more skilled individuals” (Moretti, 2004, 179). For educated workers, however, the increase in college graduates means a shift in the supply curve that lowers the equilibrium wages. The combined effect is still positive but lower. For uneducated workers there is no supply effect. Their wage level is increased because the increase in the college share enhances the productivity of all firms, that is, the productivity of both educated and uneducated workers. If there is imperfect substitution between educated and uneducated workers the net effect is positive for the latter categories: “standard neoclassical model suggests that if educated and uneducated workers are imperfect substitutes, an increase in the share of educated workers will raise productivity of uneducated workers” (Moretti, 2004, 178).

The empirical finding that the presence of HEIs at local level fundamentally changes the structure of local labor markets is confirmed by other studies (Beeson and Montgomery, 1993; Abel and Deitz, 2012). As Abel and Deitz (2012) show for the US case, the structural effect of the presence of universities with research activity on local labor markets can be observed in terms of changes in the composition of the occupational

(1) The Land-Grant University emerged in the US in the 1860s to provide low-cost higher education and to meet local technical needs, especially those relating to agriculture and the ‘mechanical arts’.
structure towards work roles with higher requirements for competences. The entire occupational structure is shifted upwards in terms of higher wages.

The impact of higher education on human capital and, through this channel, on productivity and growth, takes place mainly via the educational activities of universities. It is education (that is, teaching and associated activities) that increases the level of knowledge of a large number of people, so that they have a higher productivity when entering the job market and produce positive externalities via learning effects. Universities, however, not only teach students and produce graduates, but also do research.

The degree to which education and research are separable is an important theoretical and empirical question. To a certain extent, one can argue that teaching from professors who are active in research is completely different from teaching from instructors who are not. In this sense the creation of human capital would be the result of a joint production of universities. At the same time, however, research activities produce spillovers that are distinct from those that are channelled through education. For this reason it is important to keep the impact of education and the impact of research separate in the analysis. We now turn to the literature that examines more closely the impact of research, as (at least) partially separate from education.

2.1.2 The impact of universities through research activities

The linkage between university research and economic performance is well established in the literature. Universities produce publicly available new knowledge, which advances the frontier of knowledge and open opportunities for technological advancements.

University-based research is therefore a source of large positive externalities to the overall economy. Knowledge from university-based research may reach the economy following a variety of pathways that are different from the channels of education of students.

A recent survey on UK universities (Hughes and Martin, 2012) suggests the following list of pathways (see also Siegel et al. 2007; Rasmussen and Wright, 2015; Larsson et al. 2017):

- contract research
- collaborative or federated research in joint organizations
- licensing of IPRs
- student entrepreneurship
- consulting
- laboratory test and proof-of-concept
- mobility of researchers into companies and viceversa
- short term secondment of company staff
- support in the selection of talented graduates
- doctoral students and post-doc
- seminars
- industry visits

While for many of them there is a dedicated literature, it is useful to start with a broader appreciation of the issues associated to the definition and measurement of the impact of research on the economy. This has taken the form of estimating rates of return, and more recently of estimating coefficients in reduced form econometric models.

From this literature we derive an important theoretical problem that is crucial for our study: in principle the causality relation between research and regional development may
go in both directions. On the one hand, research generates opportunities for innovation and growth and creates qualified human capital; on the other hand, rich countries and regions may support a higher level of public investment in R&D (for a short review of the literature on R&D and productivity see Appendix 1) and also have a larger and more articulated industrial base, leading to a higher level of private R&D. Having large and liquid markets, rich countries and regions create also incentives for companies to invest into R&D in order to develop marketable innovations. Therefore it is not surprising that the studies that adopted a causality approach found both directions of unilateral causality, or bidirectionality of causality (Lee et al., 2011; Inglesi-Lotz and Paris, 2013; Jin and Jin, 2013; Inglesi-Lotz et al., 2014; 2015; Yaşgül and Guriş, 2015; Kumar et al., 2015). These studies examine the relationship in single countries (USA, China, Turkey, BRIC countries, South Africa) or in samples of countries, usually those for which complete time series based on OECD data are available. A large sample (34 OECD countries) is examined in Ntuli et al. (2015).

The most comprehensive and updated study is Solarin and Yen (2016). They examine, using adequate econometric techniques, the causal relation between research output (as measured by the volume of Scopus publications) and the rate of growth of real GDP per capita in 169 countries for the period 1996-2013. Controlling for a variety of factors, such as physical capital, financial development and institutional quality, they find strong support for a positive relationship. Interestingly, when they run separate regressions for countries at various levels of scientific achievement (using a threshold at 100 and 1000 publications per year per country) they confirm the positive relationship for all countries, irrespective of the articulation of the research system.

What about knowledge spillovers and the role of HEI? Knowledge spillovers can exist both within and across countries/regions, and they can involve directly human capital (i.e. migration of skilled workers or researchers in other countries or regions). Several authors have tried to estimate directly the magnitude of the spillover effects. The idea is that the economic activity of a given unit (firm, industry, region, country) may be influenced by the investment into R&D of other units, which support the cost. Thus innovative firms may generate spillovers in favour of competitors located in the same geographical area through non-voluntary disclosure of technical information via social ties, mobility of workers and technicians, joint use of suppliers. Or a country may benefit from spillovers from knowledge generated abroad by hosting Foreign Direct Investments of multinational companies (see Coe and Helpman, 1995, and Verspagen, 1997).

Spillover effects take a variety of concrete forms and involve many agents at several levels of analysis (Goto and Suzuki, 1989; Los and Verspagen, 2000; Eberhardt et al., 2013). The "quest" for spillover effects has generated a vast literature, in which the specification of dependent and independent variables is highly diverse, as are the time windows, geographic scale, and institutional features. In addition this issue is central to several disciplines, from economics of innovation to economic geography, from regional studies to urban economics, to spatial econometrics. Therefore it is not an easy task to summarize a large and fragmented literature.

Karlsson et al. (2013) offer a survey of a limited number of studies covering 75 separate observations. The studies examined are Bottazzi and Peri (2000; 2004); Crescenzi (2007); Crescenzi and Rodriguez-Pose (2008); Maggioni et al. (2007); Moreno et al. (2005); Pinto (2010); Varga et al. (2010); Pinto and Rodriguez (2010); Greunz (2003); Krammer (2009).

Their summary of the available evidence is that “investment in knowledge related activities (e.g. in form of R&D expenditure) tends to augment the local patent production. On the other hand, the analysis shows that the spillovers from R&D investments in non-local regions induce a positive, but marginally small effect on local patent production. Spatial knowledge spillovers tends to be concentrated to regions characterized by same technological attributes and infrastructure development” (Karlsson et al. 2013, 167). Another interesting conclusion is that “total local R&D expenditure is more efficient for local patent production when allocated via private
Finally, and closer to the focus of this report, a large and consistent literature has stressed the somewhat counterintuitive fact that the impact of (University) research on the economic system has a strong spatial dimension, that is, is localized. A number of well-crafted studies, originated in the early '90s, showed that companies use preferentially research results produced by universities located close to their facilities (Jaffe, 1989; Acs et al., 1992; Jaffe et al., 1993; Anselin et al., 1997; Varga, 2000; Adams, 2002). These findings have originated a large literature on spatial spillovers and localized knowledge flows. Feldman (1994) and Varga (1998) develop a full scale analysis of the spatial dimension of innovation, in general, and of university knowledge spill-over more specifically. Rosenthal and Strange (2008) offered a review of studies on proximity effects. Bottazzi and Peri (2003) offer an econometric estimate of the decay of spillover effects with distance, using patent data. According to this literature, largely based on spatial econometric techniques, knowledge spillovers reach external audiences with an intensity that decreases with geographic distance (Karlsson and Anderson, 2006; Abramovsky et al., 2007; Brostrom, 2010; De Fuentes and Dutrenit, 2016). Boschma (2005) is a classical treatment of various kinds of proximity. Among the most recent studies, Ponds et al. (2009) studied the spillovers from academic research to regional innovation, finding an important role for university-industry collaborations (Fischer and Varga, 2003; Fritsch and Franke, 2004). Aggregate evidence at regional level confirms the importance of this pathway. Vértessy et al. (2013) find strong correlation between university research performance and independent territorial variables such as territorial competitiveness, labour market efficiency, and innovation capability of European regions. Cowan and Zinovyeva (2012) analyze empirically whether the expansion of a university system affects local industry innovation. They examine how the opening of new universities in Italy during 1985-2000, affected regional innovation. They found that creation of new universities increased regional innovation activity already within five years. On average, an opening of a new university led to a seven percent change in the number of patents filed by regional firms. The evidence suggests that the effect is mainly generated by high quality scientific research brought to the region by the newly established university.

A recent study exploited the opportunity offered by the creation of Universities of Applied Sciences in Switzerland to adopt a difference-in-differences estimation (Pfister et al., 2017). They found that the creation of a new HEI, although a non-university one, has a strong impact on the regional economy: 8.5 to 14 percent increase in regional patent activity, and 2 to 3.6 percent increase in citation per patent. This is one of the few studies adopting a counterfactual approach in order to determine a causal impact and constitutes a benchmark for our approach as well. These findings are confirmed in the studies by Bonaccorsi et al. (2013; 2014) on the localized impact on innovative entrepreneurship.

Finally, it has been observed that a distinct channel for the impact of HEIs is the informal one, that is, the face-to-face interaction aimed at addressing specific problems. This impact takes the form of informal meetings, consulting activities, exchange of students. There are many impact pathways from universities to the economic system, and those pathways that originate from research results are not necessarily the most important (Salter and Martin, 2001; D'Este and Patel, 2007). Other pathways, in fact, refer to the application and adaptation of existing knowledge (as opposed to new knowledge) via consulting, personnel exchange, visits and informal interactions (Landry et al., 2007; Bishop et al., 2011; Ghinamo, 2012; Bianchini et al., 2015). According to a large survey on UK academics, informal relations are more important than formal ones, since they involve many more researchers. Informal relationships, consultancy activities and training of undergraduate and Ph.D. students are crucial for SMEs (Wright et al. 2008). To the extent that knowledge spill-overs are based on face-to-face interactions, they require frequent and regular personal exchanges (Storper and Venables, 2004). In turn, these take place at a distance which is constrained by the maximum distance that people
may travel in a working day. It is too expensive to meet regularly if people convene from a distance which cannot be travelled during the usual shuttle distance of daily work activity.

What these studies tell us is that knowledge spillovers are localized, that is, produce benefits in a limited spatial range. Therefore the main concern for policy makers should not be whether to spend public resources in R&D (which is always good), but to prepare the infrastructures that permit the local appropriation of these positive spillovers.

The spillover effects from research carried out at HEIs are not limited to the collaboration with local industry. HEIs bring to the local economy the benefits of their extended networking with firms and other organizations on a large geographic scale. Ponds et al. (2009) analyse the effect of knowledge spillovers from academic research on regional innovation. The effect of university–industry collaboration networks on knowledge spillovers are modelled using an extended knowledge production function framework applied to regions in the Netherlands. They find that the impact of academic research on regional innovation is not only mediated by geographical proximity but also by networks stemming from university–industry collaboration.

These findings tend to reconcile the debate that has taken place within the economics of innovation on the relative importance of different notions of proximity. While spatial proximity has been extensively studied, several authors questioned the emphasis given to it, suggesting that the impact of research may follow other dimensions of proximity, such as proximity defined in terms of scientific communities (who do not have a geographic basis) or defined in terms of social networks (which are not entirely bounded by geographic boundaries) (Breschi and Lissoni, 2001; Bathelt et al., 2004; Maggioni et al., 2007). From this perspective we might conclude that universities bring to the regional economy not only the advantages that come from the spatial proximity, in terms of student education and research, but also the access to international networks of knowledge. Regions with active universities are more internationalized and enjoy more access to international networks.

**2.1.3 Academic entrepreneurship and startup creation**

Among the various pathways through which HEI can affect the local economic performance, a special place is that of entrepreneurship. There are two main pathways: academic entrepreneurship, or the direct creation of new firms by academic staff, and startup creation, or the indirect contribution of universities to the generation of entrepreneurial ideas and opportunities. More recently, a new category has been added, i.e. student entrepreneurship.

HEIs might be a direct source of creation of economic activities, following an entrepreneurial process (Carlsson et al. 2009; Colombo et al., 2010; Colombo and Grilli, 2010). Companies directly created by universities and/or based on specific research assets generated in university laboratories are labelled spinoff companies and are the object of a dedicated literature. More broadly, universities greatly enhance the entrepreneurial process via the creation of start-up companies by university graduates and, increasingly, by students during their study period. The presence of universities has a significant impact on the creation of new firms, or the natality rate (Audretsch et al., 2005; 2012; Acosta et al. 2014; Fritsch and Aamoucke, 2013). According to the knowledge filter hypothesis, new knowledge finds its way to the market through the creation of new ventures when established firms fail to recognize its potential and exhibit inertia in investing in it (Acs and Plummer, 2005; Acs et al., 2009).

According to Carree et al. (2014) new entrepreneurial ventures may represent a viable and effective mechanism to transform academic knowledge into regional economic growth. They test this notion for the Italian provinces between 2001 and 2006. They evaluate three outputs of academic activities: teaching, research and Intellectual Property Rights (IPR) activities management. New ventures may be able to transform the mentioned outputs into improved economic performance. The findings show that the
effects of academic outputs on provincial economic growth (all sectors) are appreciable when they are associated with sustained entrepreneurial activities in the province.

Bonaccorsi et al. (2013) examine how the scientific specialization of universities impacts new firm creation across industries at the local level. They estimate negative binomial regression models separately for each industry category to relate new firm creation to the scientific specialization in basic sciences, applied sciences and engineering, and social sciences and humanities of neighboring universities. They find that universities specialized in applied sciences and engineering have a broad positive effect on new firm creation in a given province, this effect being especially strong in service industries. No effect of humanities and social sciences was found.

In a companion paper Bonaccorsi et al. (2014) investigate how far in space university knowledge goes to breed the creation of knowledge-intensive firms (KIFs), depending on the nature (either codified or tacit) and quality of this knowledge. They consider the impact of knowledge codified in academic patents and scientific publications and tacit knowledge embodied in university graduates on KIF creation in Italian provinces in 2010, while distinguishing between local university knowledge created by universities located in the same province and external university knowledge created by universities located outside the province. The econometric estimates indicate that the positive effects of scientific publications and university graduates are confined within the boundaries of the province in which universities are located. The above effects are confined to high-quality universities; low-quality universities have little effect on KIF creation.

2.1.4 Attractiveness for foreign investment

The presence of HEIs in a region might influence location decisions of companies, in particular the decisions to locate R&D activities by multinational companies (MNCs). Abramovsky et al. (2007) combine establishment-level data on R&D activity with information on levels and changes in research quality from the Research Assessment Exercise in UK. The strongest evidence for co-location is for pharmaceuticals R&D, which is disproportionately located near to relevant university research, particularly 5 or 5* rated chemistry departments. Abramovsky and Simpson (2011) investigate evidence for spatially mediated knowledge transfer from university research. They examine whether firms locate their R&D labs near universities, and whether those that do are more likely to co-operate with, or source knowledge from universities. Confirming the results from Abramovsky et al. (2007), they find that pharmaceutical firms locate R&D near to top quality chemistry departments.

2.1.5 High value added procurement

A recent stream of literature examines universities as sources of purchasing, in particular of high added value items, such as scientific instrumentation, software, equipment and the related specialized knowledge-intensive services. It turns out that the demand for products and services from universities activates a distinct channel of impact on the local economy.

Goldschlag et al. (2018) examine data on purchases from 13 research-intensive universities in the USA, referring to more than 1,2 million transactions, based on the UMETRICS consortium collaboration. They show that “Establishments physically closer to each university are disproportionately more likely to be vendors to the university's researchers than are other U.S. establishments” (Goldschag et al. 2018, 9). The vendor relation tends to be stable over time. If vendors establish new facilities, this takes place more likely close to the university for which they do more business.

2.1.6 Cultural and social externalities

Universities create a social and cultural climate in which valuable non-university activities find a favorable environment. The presence of a population of young and educated people brings with itself favorable conditions for cultural activities, as well as for
entertainment and leisure. This effect is at the core of the literature that has been
developed around the popular notion of creative class, introduced by Richard Florida
(2002a; 2002b; 2005).

Although Florida's creative class approach has been criticised sharply for a variety of
substantive and methodological reasons (see, for example, Glaeser 2004; Lang and
Danielsen 2005; Peck 2005; Boyle 2006; Hansen et al. 2005; Markusen 2006; Scott
2006), it has attracted the attention on the spatial contiguity between occupational roles
for which formal, tertiary-level education is a requisite, and activities that do not require
credentials but share some features of autonomy, self-direction and learning.

While the empirical validity of the notion of creative city is highly controversial, there is
no doubt that universities contribute to the creation of creative environments.

2.1.7 University expenditure

Finally, there is a classical Keynesian multiplier effect, represented by the additional
expenditure of the student population in the territory, in terms of accommodation, food,
transport, sport and leisure activities. To this student-based effect it should be added the
expenditure of the national or regional government that is channeled to the university,
for example for the payroll of academic and nonacademic staff. In other words,
universities trigger additional demand effects upon the local economy. This is the object
of a dedicated literature aimed at quantifying the multiplier, often with the goal of
advertising it to the media and the audiences.

It is worthwhile to review briefly a few of the most cited studies (Elliott et al., 1988; Kott,
1987-88; Beeson and Montgomery, 1993; Blackwell et al., 2002).

The direct positive impact of universities in the short term comes from:

— direct expenditures of students and faculty;
— expenditure of students and faculty who move into the area from other regions;
— expenditures by the institution;
— funds for research, salaries, and other expenditures (equipment, infrastructure,
supplies).

In studies that estimate the local impact of universities these expenditures are then
multiplied for a number, which reflects the Keynesian effect of additional aggregate
demand.

Long term direct and indirect effects include most of the effects we discuss in this survey
of the literature:

— enhancement of workers' skills;
— relationship between research and local industry;
— positive effects on business location (attraction of foreign investment);
— business creation (creation of new firms and technologies).

There are a number of methodological issues when trying to estimate multiplier effects.
First, the decisions to move and locate into the area of students and of faculty and staff
are not easy to interpret. Second, the allocation of expenditures to institutions is often
problematic, due to fixed infrastructure costs or shared facilities and activities. Third, the
quantification of multipliers is subject to subjective distortions. Perhaps more
importantly, these estimates ignore the general equilibrium effect of the impact of
universities. Attracting students from region A to region B gives a boost to the
expenditures and income in region B, but the net effect is to be calculated by observing
the real increase in welfare measures (for example due to the creation of better human
capital if students move to better universities, or the creation of a more effective
matching between individual skills and the type of education in the host university).
Nevertheless, almost all the published studies show a very high local impact of the presence of universities, with a multiplier much greater than one on average.

2.2 Contextual factors moderating the impact of university education and research

In this section we review some of the most important contextual factors that influence the positive impact of universities on the local economy. These factors may have a direct or indirect, and positive or negative, influence on the relation. Finally, they mainly refer to education and research.

In the case of education, for example, a classical debate refers to the conditions under which the strong positive relationship with economic growth holds. While the main thrust of the literature has been on establishing a strong positive impact of human capital on economic performance and growth, it must be noted that several authors have raised objections. Famously, Benhabib and Spiegel (1994) found a weak correlation between human capital and economic performance. Pritchett (2001) confirmed this finding and addressed several possible explanations. A large literature followed this line of analysis (Hanushek and Kimko, 2000; Bils and Klenow, 2000; Hanushek and Woessmann, 2009).

With respect to our discussion there are several important qualifications, which we can interpret as contextual factors, on which the involved agents have partial or no control, be universities or firms.

They are listed as follows.

(a) Contextual factors moderating the impact of creation of human capital
- complementarity between formal education and cognitive and non-cognitive skills
- matching between the competences of graduates and the demand of firms (and associated potential mismatch, under- or over-education effects)
- quality of education and university-work transition
- migration of graduates.

(b) Contextual factors moderating the impact of research
- quality of academic research
- regional pattern of specialization and absorptive capacity
- complementarity and substitution effects between the different research pathways
- complementarity between human capital accumulation and innovation.

2.2.1 Complementarity between formal education and cognitive and non-cognitive skills

First, there is a distinction between formal education and the acquisition of skills and competencies. The literature on human capital has taken the numbers of years of education as the main independent variable. This is due to the lack of other large scale and reliable data. According to Hartog: "Schooling is taken to be the only link from ability to productivity. This ignores all other learning, such as learning from general life experience, learning in the home environment and learning in the work environment from on-the-job training and work experience. These are serious omissions, but including them would simply be too much for a single survey". (Hartog, 2001, 517).

For the sake of the truth, the seriousness of this limitation has been recognized since long time: "Frankly, I find it hard to conceive of a poorer measure of the marketable skills a person acquires in school than the number of years he has been able to endure a
A major flaw of this approach is, therefore, the lack of consideration for the difference between formal education and the bundle of competencies needed in the work environment (Busato et al. 2000). It is this bundle of competences that leads to the increase in productivity and hence to economic growth.

An important missing variable is the extent to which people are subject to on-the-job training. On-the-job training is a major source of learning and acquisition of skills and competencies. It may take place with formal training initiative, or in informal ways.

This issue has gained prominence after data on cognitive skills have been made available by international surveys such as IALS (International Adult Literacy Survey), ALL (Adult Literacy and Lifeskills Survey) and PIAAC (Programme for the International Assessment of Adult Competences). Several authors have started to use standardized data to examine the impact of skills on earnings and the rate of return of education. Hanushek et al. (2015) use PIAAC data on cognitive skills in 23 countries to examine variability in earnings, showing large heterogeneity.

These studies have opened a debate on the relative role of formal education and cognitive skills, given that the latter may depend also on non-educational factors. A recent statement of the problem, based on PIAAC data for 20 European countries, strongly confirm the importance of formal schooling (Cappellari et al. 2016). Nevertheless, it is important to keep this distinction into the analysis of the regional impact of HEIs.

### 2.2.2 Matching between the competences of graduates and the demand of firms

The impact of human capital on firm productivity and growth depends on whether the competencies are deployed in an occupational role and within an organization. In turn, this requires that firms show a demand for workers’ competencies that is consistent with those supplied by graduates.

The process of matching between the skills and competencies of people and the requirements of the employee is subject to several possible failures, or mismatch.

There are several types of mismatch (Cedefop, 2010; Sloane, 2014): vertical mismatch (level of education or skills is lower or higher than the requirements of a job position), horizontal mismatch (level of education or skills adequate, but not the field of education), over- or underqualification (a person has a higher/lower level of qualification than required for a job position), over- or underskilling (a situation where a person is unable to fully use their skills and abilities in the current job position, or lacks necessary skills and abilities for performing the current job).

There is a large literature that examines the matching, or lack thereof (mismatch) between the supply of competencies, as measured by the titles and degrees of graduates, and the demand of competencies, as measured by the job vacancies (Cedefop 2010, 2018; Garcia-Aracil and Van der Velden, 2008; Caroleo and Pastore, 2015).

In particular, an alarming possibility is that the accumulation of human capital and the investment of firms into R&D mutually reinforce each other, but not in the virtuous way predicted by endogenous growth theory. This is called “low education-low innovation” trap and has been studied with particular reference to Southern European countries (e.g. Italy: Schivardi and Torrini, 2011). Endogenous growth theory posits that both human capital accumulation and use of new knowledge are subject to positive externality effects and increasing returns. Colonna (2016) examines the possibility that these effects work jointly in bringing the economy downward to a low investment equilibrium, given an initial specialization in low technology sectors. In practice, the population does not invest enough (or invest wrongly) because the demand of skills by firms is low, while firms find
it easier to hire low qualified workers rather than re-skilling graduates who miss the practical skills needed for low technology manufacturing settings.

This issue is related to another well-studied phenomenon in higher education: the dropout of students (i.e. students that decide to abandon their studies before completing them). Dropout rates are monitored by the OECD (OECD *Education at a glance*, various years) and have been the object of several studies at European level (NIFU-Cheps, 2014; Schopf, 2014; see Zotti 2015 for a comprehensive survey).

For these reasons an interesting empirical question is whether the impact of higher education on the economy is mediated by the institutional design of the system, which can amplify/reduce mismatches. In particular, it is well known that not all European countries adopt a distinction between university training and vocational training. In many European countries (namely, Southern European and, with some qualifications, France) the vocational training associated to service industries and to medium-level qualifications is not managed by separate institutions, but is managed by universities, in collaboration with other actors. In other words, they do not have a strong non-university sector. In these countries universities absorb the largest share of students.

Other countries, namely Germany and countries under the German linguistic and cultural influence (Austria, Switzerland, part of Eastern European countries, some Scandinavian), have developed a vigorous, rich and professional segment of higher education, non-university institutions, in charge of developing curricula with strong practical and professional orientation. This is called dual or binary model of higher education (Clark, 1983; Kywik, 2004). In dual systems a large share of students attend non-university HEIs, such as Fachhochschulen (in Germany, Switzerland, Austria) or Hagenschulen (in Netherlands) or colleges, in several Scandinavian and Eastern European countries.

The two systems address differently the long term impact of mass higher education (Trow, 1974; Clark, 1978). It is a matter of debate whether non-dual European countries are moving towards a convergence with the dual model, or whether they maintain a unitary system (Teichler, 1988; Huismann, 1995; Meek et al. 1996; Huismann et al. 2007). Most recent studies show the inertia of institutional systems and the persistence of significant diversity in the higher education landscape (Kehm and Stensaker, 2009; Fumasoli and Huisman, 2013).

It is therefore interesting to empirically examine the difference between universities and non-university institutions with respect to their respective impact on the economy, by using microdata. Examining the differential impact of higher education institutions on the labor market in the UK, Faggian and McCann (2009) find that the so-called post-1992 universities, or Coalition of Modern Universities (CMU, that is, former Polytechnics that received the status of universities after the Thatcher reform, are not research-intensive and are mostly oriented towards the local economy) have the largest impact: “Furthermore, when we examine how the mobility of graduates is related to the university attended and the characteristics of both the students and the local economic environment environment, we find that CMU universities are the universities with the strongest local labour market effects. These are the universities which have a strong local orientation and mission. As such, from the perspective of the local region, the CMU universities perform a different role to the other types of universities and colleges, most of which tend to produce graduates who move away to other areas for employment” (Faggian and McCann, 2009, 221). CMU universities in the UK play the same role as non-university institutions in dual systems.

The matching between education and demand for skills depends largely on the sectoral specialization of the industry. For these reasons the impact of human capital on the regional economy depends on the mitigation of the mismatch, which in turn depends on supply and demand factors in different ways. As it has been suggested: “*demand factors generally play a major role in reducing educational mismatch in technologically more advanced countries, whereas supply factors are more important in countries that are*
lagging behind in the international division of labour” (Ghignoni and Verahschagina, 2014, 670).

It is also likely that the specialization of universities in specific fields of education and research has heterogeneous impact. For example, Bonaccorsi et al. (2013) find that the impact of universities on the creation of new firms is significant only for the fields of Medicine and Engineering, while the impact of Natural sciences and, moreover, of Social sciences and Humanities is negligible.

2.2.3 Quality of education and university-work transition

As already noted, most of the literature on the impact of human capital on growth uses a proxy based on the number of years of education, or formal schooling. Besides ignoring the issue of other forms of learning, this approach completely ignores the differences in quality of education. The same year of education may have a largely different impact on the learning of students, hence on their competencies and skills and their potential impact on productivity, across countries. The importance of quality of education has been demonstrated with respect to secondary education, based on comparable data on education outcomes across countries, available from standardized test on student performance (Salas Velasco, 2014).

We are not aware of similar studies on higher education. The issue of student learning outcomes at university level has been initially examined in the OECD context, but we are far from a consensus on conceptual and methodological grounds. The pioneering studies carried out under the AHELO framework (Assessment of Higher Education Learning Outcomes) come to the conclusion that comparable measures of student learning as an outcome of education can be achieved for soft or transversal skills, such as problem solving. They are much more difficult to achieve for domain or disciplinary knowledge. Furthermore, the measurement of learning outcomes is meaningful only if one carefully controls for the level of skill at the enrolment date, following a notion of “value added”. However, while internationally data on cognitive skills, such as the PISA survey, take students of the same age for reasons of comparability, the age at which students are enrolled varies across countries. The degree to which same-age data on cognitive skills can be taken as a reliable measure of the level of entry in higher education is therefore dubious in comparative studies.

A related factor is the way in which graduates enter the job market. This is called university-work transition. It has been shown that graduates find their job more easily and with a shorter waiting time if the institutional framework of higher education facilitates the matching between their competences and the demand by firms before they complete their studies. Therefore initiatives such as stages, plant visits, company projects, internships and the like may have large impact on the labour market outcomes of graduates.

Large scale surveys in several European countries, such as the CATEWE project and the CHEERS (Careers after Higher Education. A European research survey) have highlighted the importance of these factors (Allen and de Weert, 2007; Garcia-Aracil and Van der Velden, 2008; Salas Velasco, 2017)

2.2.4 Graduate migration

Finally, the effect of student and graduate mobility must be taken into account.

Several studies find that graduates are more mobile than individuals with lower qualifications (Kodrzycki, 2001; Gottlieb and George, 2006; Whisler et al. 2008). Kodrzycki (2001) estimates that college graduates are at least twice as likely to move from the US state where they attended college or high school.

Graduate mobility weakens the link between the local presence of universities and the local economy. Students who study and graduate in a region may migrate to another
region. The creation of human capital takes place in a different place than the exploitation of human capital.

The extent to which this phenomenon fundamentally alters the impact of universities is a matter of empirical estimation. Most likely, it depends on institutional features of national higher education systems, as well as other structural factors leading to people mobility, such as, among others, local development, local labour market conditions, family structure, cost of transportation, and markets for house rentals (Antolin and Bover, 1997).

As a general remark, students that attend universities in the same area of their domicile will tend to look for a job in the same area, while students who migrated in order to attend universities may stay for a job in the same area, come back home, or migrate to another region. Using census data on Italy, Ciriaci (2014) has shown that out of a cohort of graduates, 89% studied in the same regional macroarea of origin, while 11% migrated to another area. Interestingly, of those that studied in the same area, 93.5% stayed in the same area also after graduation, while only 6.5% migrated. By contrast, among those who migrated for study, 53.4% move to another macroarea after graduation. There is hence correlation between migration before and after graduation.

The factors that lead students to migrate for study are the object of a large dedicated literature (Desjardins et al., 1999; Dotti et al. 2013; Sá et al., 2004; Caruso and de Wit, 2014; Cattaneo et al. 2017). Among them many studies emphasize the quality of university and their reputation as a major element of attractiveness, in addition to personal (student and family) and institutional factors: “one of the key drivers of migration by university students is the presence of a good university in a given province” (Dotti et al. 2013).

A recent JRC Report (Sánchez Barrioluengo and Flisi, 2017) offers the most update analysis of student mobility. Combining ETER data with Unesco-OECD-Eurostat data they show that institutional characteristics of individual universities are an important determinant of attractiveness. In particular, “among institutional characteristics, better quality universities and those with a higher reputation are associated with a higher share of mobile students, while research orientation and excellence are more relevant for degree mobile PhD students” (Sánchez Barrioluengo and Flisi, 2017, 2). Among regional factors, high density regions have higher mobility rates.

In a series of papers, Faggian and McCann have examined the mobility of UK students from domicile to university and from university to work (Faggian and McCann, 2006; 2009; Faggian et al., 2007). They find positive correlation between mobility from domicile to university and subsequent mobility. They also find an average strong attraction effect from other regions to the London region, which is characterized by high quality, research-intensive universities. The implications for the impact on the local economy are not as clear-cut: “While regional innovation performance is reduced in regions where universities are not research active, there is very little evidence to suggest that regional innovation performance is directly related to the local presence of international quality research universities. (…) Universities attract students into a region, and many of these subsequent graduates will remain in the local region for employment if the local economy is strong. This process itself also contributes to the future innovation performance of the local region, and this cumulative migration-innovation effect is critical in the case of the London regions. This labour hysteresis effect is itself a form of indirect information spillover, and appears to be far more important in contributing to regional innovation performance than any direct information spillovers between universities and local Industry” (Faggian and McCann, 2006, 496-497).

At the same time, a study on Finland reaches somewhat different conclusions. Haapanen and Tervo (2012) find that most of graduates do not move from their region of studies after 10 years after graduation. Mobility from the place of graduation is higher from universities located in peripheral regions and is lower for students that attend university in their place of origin. Venhorst et al. (2011) add another perspective to the problem, by
showing that Dutch regions have increased the share of retention of graduates in the last decade. The size of the local labor market is considered the main determinant for the retention effect.

In a recent survey of the literature, Faggian et al. (2017) show that there are positive as well as negative effects on the origin region from migration of skilled graduates and conclude that the evidence is not large enough to argue that the relation between the local presence of universities and the local economy is fundamentally disrupted by migration of the high skilled.

While this general conclusion may be valid in general, it is likely that the relation between local economic conditions and universities takes a specific direction in the case of laggard regions.

Laggard regions are net exporters of student population and are not able to attract university students from other regions. As it has been clarified by a recent literature, migration does not lead to convergence among regions in the level of GDP, and, in fact, it may exacerbate spatial inequalities. This happens because migration of highly skilled people follows just one direction, from poor regions to rich ones, and (almost) never the other way round. This is called selective migration (Berry and Glaeser, 2005; Kanbur and Rapoport, 2005) and is at the origin of persistent regional inequalities in human capital (Fratesi and Riggi, 2007). As Dotti et co-authors have argued, “there may be a reinforcing mechanism between university attractiveness and economic development, since universities benefit from the dynamism of the local labour market in terms of superior student attractiveness, and, in turn, richer provinces benefit from the presence of attractive universities in order to gain bright students from lagging areas of the country” (Dotti et al. 2013, 459).

2.2.5 Quality of academic research

The bulk of the evidence points to the notion that the impact of universities on the regional economy is stronger if the academic quality of research, as measured through publication indicators, is higher. We borrow from Bonaccorsi (2016) for a brief survey of the literature.

First, there is large evidence that the new science-based technologies and the industries they have generated in the last three decades have been the direct or indirect outcome of world class academic research. Without any completeness due to the huge literature, one can easily record that biotechnology (Zucker et al. 1998; Colyvas et al. 2002), information technology (Agrawal, 2001), semiconductors (Lim, 2000), advanced materials (Baba et al. 2009) or nanotechnology (Bonaccorsi and Thoma, 2011) were all originated in (or hugely benefited from) laboratories in top universities or research institutes, mainly in the United States, in a few cases in Europe (e.g. the invention of Scanning Tunneling Microscope at IBM and ETH in Zurich and graphene at Manchester).

Second, firms prefer to collaborate with top-tier universities rather than second-tier universities (Mansfield 1991, 1995; Mansfield and Lee, 1996; Laursen et al. 2011; Mora-Valentin et al. 2004; Muscio and Nardone, 2012; Hong and Su, 2013). Reputation and prestige of the university is a powerful attractor for industry cooperation (Schartinger et al. 2002; Link and Scott, 2005; O’Shea et al. 2005; Powers and McDouglas, 2005; Fontana et al., 2006). University reputation, in fact, solves a fundamental adverse selection problem: external actors are likely to engage into relation with prestigious universities because they believe the latter make available their best technology, contrary to what private companies would do, keeping the best technology for internal development (Effelbein, 2006). Consequently, high quality of academic research gives rise to positive signaling effects (Podolny and Stuart, 1995; Sine et al., 2003) and mobilizes additional industry funding (Bruno and Orsenigo, 2003; Muscio and Nardone, 2012).

Moreover, in the case of cutting-edge knowledge, firms search for collaboration with top quality universities irrespective of the distance (Zucker et al., 1998; Darby and Zucker,
A plausible interpretation of this evidence is that firms with high absorptive capacity look for high quality departments, irrespective of the distance, while firms with low absorptive capacity need proximity, somewhat irrespective of, or less sensitive to, research quality (D’Este and Iammarino, 2010; Laursen et al. 2011).

Third, large companies base their R&D location decisions on the quality of the university research located in the host country and site. They systematically scan the scientific landscape in search of potential partners, rank research laboratories around the world, identify research stars and use this information in their locational models. Only the best graduate programs attract industrial R&D facilities (Malecki, 1987). R&D location and subsequent attraction of talents at local level is triggered by the presence of star scientists.

A consistent literature, originated with the studies of Zucker and Darby, has shown how large firms in science-based industries in the USA locate their R&D laboratories close to star scientists (Zucker et al., 1998; 2002; Zucker and Darby, 1998; 2006; 2007). Their findings have been confirmed with respect to other countries, including European ones (Mariani, 2002; Adams, 2005; Karlsson and Andersson, 2006; Athey et al., 2007; Maier et al., 2007; Tripl, 2009; Laursen et al. 2011).

In a survey of 250 MNCs from US, Europe and other countries, Thursby and Thursby (2009) observed that large companies locate their R&D centres close to university campuses in which star scientists were active. While a “faculty with special scientific or engineering expertise” is one of the top criteria for location in developed countries, it is still highly relevant for location in emerging countries, more than cost-related factors. Attractiveness of talent is associated to strong local research (Beeson and Montgomery, 1993; Audretsch et al., 2005; Woodward et al., 2006). This relations is somewhat mediated by industry-level variables: pharmaceutical, chemistry and materials science R&D tends to be located close to star departments in UK, machinery and communication equipment R&D is co-located with lower-quality rated research departments (Abramovsky et al., 2007), while mechanical and civil engineering firms make use of university cooperation while not being science-based sectors (Bekkers and Bodas-Freitas, 2008).

Fourth, the weight of the evidence is towards a positive association between research productivity and spillovers that may contribute to regional growth, such as academic patenting and academic entrepreneurship (Thursby and Thursby, 2002; Breschi et al., 2007; Carayol, 2007; Lissoni et al., 2008; Toole and Czarnitzky, 2010; Crespi et al., 2011; van Looy et al., 2011; Lawson, 2013; Perkmann et al., 2013). Faculty quality is positively related to the number of university invention disclosures and their commercialization (Friedman and Silberman, 2003). With respect to entrepreneurship, there is large evidence that start up firms created from high quality research are more numerous, and survive and grow with higher probability (Di Gregorio and Shane, 2003; Audretsch and Lehmann, 2005; O’Shea et al. 2005; Link and Scott, 2005; Wright et al., 2008; Colombo et al., 2010; Avnimelech and Feldman, 2011; Audretsch et al., 2012; Fritsch and Aamoucke, 2013; Conceição et al., 2014). Researchers from top universities have easier access to critical resources (van Looy et al., 2011). Fini et al. (2016) show that the rate of birth of university spinoffs is influenced by the reputation of the university in research, measured with national rankings. There is also some evidence that start up firms originated from high quality research go more frequently into IPOs and receive a higher initial stock evaluation (Powers and McDougall, 2005). As Hill (2006) puts it, “universities with the greatest economic impacts are generally those with the highest quality research programs” (p. 4).

The spatial dimension of research excellence is a matter of considerable policy attention. One important issue, recently raised in the literature on Higher Education systems in Europe, is that universities have large internal variability with respect to their research quality. In one word, we observe very low correlation between indexes of research quality across departments, so that it is not true that excellent departments are found
systematically together, under the same university. In a typical European university one find excellent departments in a few disciplines (sometimes just one) under the same umbrella with poor or average departments in other disciplines (Abramo and D’Angelo, 2014; Bonaccorsi et al., 2016). Contrary to what happens in US and Anglo-Saxon countries, in European countries good quality researchers are not concentrated in top quality departments, which themselves are under the umbrella of top universities, but are scattered almost everywhere (Bonaccorsi et al. 2017).

This means that the interaction between individual universities and firms located in the surrounding may take different intensity depending on the field, following field-specific factors that are often the result of idiosyncratic and historical factors. Recent bibliometric evidence shows that European universities (more generally, European science) falls behind US universities in the upper tail of research excellence (Albarran et al. 2010; Herranz and Rui Castillo, 2013; Rodriguez Navarro, 2016; Rodriguez Navarro and Narin, 2018). Europe produces a larger volume of research with a good average quality, as measured in terms of citation per paper, but is not leading research at the top.

The spatial implications of this situation have not been explored so far. On the one hand, a clear implication is that Europe lacks a good number of large excellent universities, or universities that excel across a large spectrum of fields. These are found only in UK (Oxbridge and London area), in Switzerland (EPFL and ETH), and to a lesser extent in the Netherlands. At the same time it is true that “islands” of excellence can be found in many universities, scattered throughout Europe. To the extent that the impact on the regional economy is mainly based on attractiveness (i.e. attraction of top scientists and students, attraction of R&D laboratories due to the presence of star scientists) this feature might be a source of weakness. To the extent, on the contrary, that the main impact takes place through creation of human capital and direct interaction with local firms, the distributed model might generate a diffused effect. More research is needed to address this issue from an empirical point of view.

2.2.6 Regional pattern of specialization and absorptive capacity

Yet the impact of university research on the economy does not depend only on the university side, but also on the firm side. Firms are not distributed evenly in the space. The spatial distribution of firms is sticky, given the path dependency of location decisions, agglomeration economies, and the costs of relocation.

Thus at any point in time the firm structure at regional level is largely given. This opens the way for the possibility of another mismatch, between the pattern of sectoral specialization of firms and the flow of knowledge that originates from university research. A recent literature examines more closely the issue of sectoral differences in the extent and impact of knowledge spill-over (Schartinger et al. 2002; Landry, Amara and Ouimet, 2007; Bekkers and Freitas, 2008; Anselin et al. 2010; Bodos Freitas et al. 2013). An established tradition in the economics of innovation argues that firms largely differ in the way in which they produce, absorb and transform knowledge (Pavitt, 1984). A fundamental difference is in place between those industries in which innovation is generated via systematic investment into R&D (science-based) and those industries in which it comes from various non-R&D sources, such as interaction with industrial users (specialised suppliers), or from suppliers and producers of equipment and raw materials (supplier-dominated). More recently, the distinctive role of industrial design and aesthetic design as a source of innovation has been recognized.

These structural differences are clearly visible in the sectoral distribution of R&D intensity (expenditure in R&D/turnover). Another classical concept in this tradition is absorptive capacity. According to Cohen and Levinthal (1990) the ability of companies to absorb external knowledge (for example, produced by universities) is a positive function of the internal, in-house investment into formalized R&D. Consequently, only companies that have a separate budget and dedicated personnel with a higher education background benefit from the interaction with universities (Veugelers, 1997).
This problem is particularly delicate in catching up regions, in which the sectoral pattern of specialization is usually tilted towards traditional sectors, with a very low investment into R&D.

The recent literature on the role of universities for regional development, particularly for peripheral regions, calls for a more critical approach, overcoming excessive expectations and misguided policies (Drucker, 2016; Brown, 2016; Bonaccorsi, 2016). It is shown that universities per se are not enough to trigger opportunities for economic growth in the short term and that their impact is conditional on the presence of complementary factors. The main role of universities is the creation of human capital in the very long term, to be measured in the time scale of several decades, if not centuries. In the short to medium term, their impact on growth is highly selective. In particular, it originates mainly from “pocket of excellence”, that is, areas of concentration of talent in highly dynamic scientific fields, that can be linked to industrial activities. This critical approach is consistent with studies on catching up of lagging countries and regions at world level. Park and Lee (2015) show that latecomer rely more on recent and scientific knowledge than incumbent firms, but only in short-cycle technological fields, in which there is rapid obsolescence of technology. Thus the implication is that not all scientific research matters for generating opportunities, but only that that can enter rapidly growing sectors not dominated by incumbents. Ramanayake and Lee (2015) show that only export growth matters for triggering the economic development of lagging countries. Thus the challenge is how to relate excellent research to export-oriented new industrial activities.

The importance of the matching between human capital and the productive structure of regions has been recently highlighted by Teixeira and Queirós, who examine the case of Spain for half a century (1960-2011) and of other countries. They find that “the effect of human capital via specialization in high-tech and knowledge-intensive activities is negative. The latter result indicates that the lack of industrial structures able to properly integrate highly educated individuals into the productive system leads countries to experience disappointing economic returns” (Teixeira and Queirós, 2016, p. 1636).

They confirm the important findings of Vandenbussche et al. (2006) who show that the contribution of human capital to growth is larger if countries are close to the technological frontier (in OECD countries). Since innovative activities are more skill-intensive than traditional ones, countries with a specialization pattern in which innovative sectors have a larger role benefit more from investment in human capital.

The "mismatch" between the quality of the university and the pattern of local industrial development can explain the negative relationship that some find (Maietta, 2015; Barletta et al. 2017; Maietta et al. 2017) between academic excellence and localized impacts, possibly due to the fact that the most productive relations with industry are not managed by top level, or high quality researchers, who are instead mainly motivated by the need to publish in good journals. More precisely, Maietta (2015) finds that the research quality of the closest academic institution has a negative impact on product innovation of firms. Research quality is measured using bibliometric indicators on publications and national research assessment indicators. In another paper she finds that the number of citations has a negative effect on the intensity of university-industry collaboration (Maietta et al. 2017). Barletta et al. (2017), on the other hand, examine data at research group level and find a negative relation between the number of publications per capita and technology transfer activities.

The latter effects are driven by industry-specific effects: mature, or medium-technology industries (such as for example the agro-food industry) benefit more from research of lower level, that is, from the interaction in close proximity with universities that are not necessarily able to compete at the top of research quality. In particular, it is argued that in medium and low-tech industries the most important contribution played by universities is not the production of top quality research, but rather the education of students with

(2) One might remind of the famous sentence attributed to US Senator Daniel Patrick Moynihan: “If you want to build a great city, build a great university, and wait 200 years”.
practical orientation and the engagement into trouble-shooting and problem solving activities of firms, which do not require high levels of research creativity (Laursen and Salter, 2004; Laursen et al., 2011; Bodas-Freitas et al. 2013; Maietta, 2015). This is particularly true for SMEs, which lack absorptive capacity for research but often exhibit high levels of creativity in their business orientation.

According to Mansfield and Lee (1996), universities of lower quality may still be very active in promoting firms' innovation, if this does not require advancements of research. They asked managers in seven high tech industries to mention five academics that contributed the most to their innovation. As expected, top departments were cited more often, but a good deal of citations were also received by lower-rank departments.

Here the quality-proximity trade-off is clearly visible. Firms in mature industries, which have in general lower absorptive capacity due to minimal R&D investment, take advantage from interacting with universities only if these are located close to their plants. They do not search distant universities that are top quality in specific research fields, but rather are satisfied with interaction with local researchers, irrespective of their status in international research (Brostrom, 2010; Hong and Su, 2013). It seems that large and/or science-based and technology-based firms may engage into long distance interaction with universities. This is what has been found empirically: Abramovsky et al (2007) show that in the pharmaceutical and chemical industries (i.e. in science-based industries) firms locate their R&D laboratories close to universities with a high rank in research, even if they are distant from the headquarters, while in middle-technology industries, such as motor vehicles, they interact with universities located in their vicinity, irrespective of the quality.

This evidence also points to a major issue in estimating the impact of HEIs on local economies. If the local economic system includes large companies and/or science-based and technology-based firms, then it is likely that local graduates will find an employment in the same place where they have studied. Similarly, researchers based at universities will find it easy to interact with local companies. If the local economic system, on the contrary, is only based on SMEs and/or firms in low-to-medium technologies that only have access to knowledge spill-overs generated at short distance (Fritsch and Franke, 2004; Hong and Su, 2013), then there is large room for mismatch. A certain number of graduates will migrate to other cities or regions in search of better job opportunities, and local researchers will search for industry collaborations not with local firms, but with distant firms. It is likely that this effect will be greater the higher the quality of education and research of universities. In the absence of microdata on mobility flows, these effects could be addressed by introducing several classifications that try to capture the absorptive capacity of the region.

2.2.7 Complementarity and substitution effects between the different pathways

The impact of HEIs on the regional economy is not without costs, however. The impact is generated when HEIs put effort in interacting with external actors, often allocating dedicated resources (such as technology transfer offices, specialised intermediaries, professional roles), affecting other important university activities (positively or negatively). This creates tensions between the various dimensions of university missions. Traditionally, education/teaching is considered the first mission (by historical origins), research the second one, and interaction with the external economy and society the third mission.

While in the past the role of complementarity between the three missions has been stressed, recently some contributions have also pointed to potential substitution effects, in particular in the areas of:

a) research vs. external engagement
b) research vs. teaching.
These tensions originate by the fact that these missions require activities that compete on the same time-budget of researchers. In addition, the tension between research and external engagement depends also from the potential mismatch in procedures and the ethos of public research.

With respect to the first tension, it is argued that external activities of universities directly benefit research quality (Calderini et al., 2007; 2009). Among the most recent studies Bonaccorsis et al. (2014), Calcagnini et al. (2016), Fini et al. (2016) and Szucs (2018) find evidence of a complementarity relationship. Similarly, meta-analysis of research results indicates that there is a weak but positive relationship (Hattie & March, 1996; Verburg, et al., 2007). Wang et al. (2016) conclude that academic commercialization and engagement yield a combined positive effect on teaching.

On the one hand, maintaining close relations with firms may offer universities the opportunity to meet challenging operational problems, which in some cases may be of scientific interest. This is particularly true in frontier research, or for research cross-cutting the boundaries between fundamental research and applied solutions, which is typical of emerging industries. Second, collaboration with industry offers academic researchers additional resources, often with a scale that is not easily offered by academic funding (Perkmann et al. 2013).

However the literature also highlights that the decision to strengthen particular activities might have a negative influence on other activities. This is what the literature calls the "substitution" effect between university missions. Some examples in this regard suggest the "substitutability" between research quality and third mission, or impact on society and economy. Producing high quality research is a time-consuming activity, as it requires maintaining operational relations with firms and social actors. As already stated, these two activities use the same (relatively) fixed amount of academic human resources. The research question here refers to the existence and magnitude of both substitution and complementarity effects on firm performance.

Fini et al. (2011) analyze the extent to which University-Level Support Mechanisms (ULSMs, such as Technology Transfer Offices and dedicated administrative staff) and Local-Context Support Mechanisms (LCSMs, such as incubators, technology parks, or specialized intermediaries) complement or substitute each other in fostering the creation of academic spin-offs. Similarly, time dedicated to research also competes with the number of hours allocated to teaching.

In a more recent paper, Fini et al. (2018) examine the impact of universities faced with the challenge of meeting simultaneously the high standards of global research competition and the local demands. This creates tensions that are solved only through the creation of a dedicated professional layer of agents that promote intense relations while delivering benefits on both sides of the missions of universities.

Moreover, in most European countries governments have adopted policies that link the funding of research to the outcome of research production, as measured mainly by publications. This creates a strong pressure on researchers for publishing their results, even if this is detrimental to the time and energy allocated to teaching, knowledge transfer and consulting activities.

Also, it is possible that the kind of research needed by firms is not top quality, but rather ordinary and consolidated knowledge. This argument depends crucially on the absorptive capacity of firms (see above for a fuller discussion). Large firms have internal R&D departments, composed by researchers with postgraduate degrees, who have an intimate understanding of the research generated by universities. These firms are in a position to absorb the knowledge generated by academic researchers, for example by reading and consulting the scientific literature. On the contrary, firms that rely more on non-R&D innovation do not take benefit from interacting directly with top researchers but need an effort to translate and adapt the knowledge. Finally, several forms of interaction between university and industry directly prevent the publication of results. By granting research contracts to universities, in fact, private companies may ask that some results
are not published (industrial secret), or are published only after a certain delay, or embargo period.

The second type of tension, between teaching and research, is also the object of a dedicated literature.

Universities carry out teaching and research activities together. In economic terms, there are economies of scope between teaching and research (Izadi et al., 2002; Johnes, 2004; Longlong et al., 2009; Burke and Ron, 2010; Worthington and Higgs, 2011; Bowden and Gonzalez, 2012). However, the existence of economies of scope does not exclude the possibility that the relation is not linear. Two possible effects are of interest here: whether an excess emphasis on research may lead to inadequate efforts in teaching, or whether an excess commitment to teaching may be harmful to research.

It turns out that the answers to these two questions are remarkably different.

Feldman (1987) carried out the first extensive survey of studies that compare the teaching evaluations of professors that are active in research with those that publish little or nothing. He found that only one study out of thirty concluded that research had a negative impact on the quality of teaching. Bok (2013) has recently extended the survey to more recent studies (Mc Caughey, 1992; Olsen and Simmons, 1996; Sullivan 1996) and concludes that "the soundest conclusion to draw from the existing studies is that engaging in research has no significant demonstrable effect, either positive or negative, on the quality of undergraduate teaching" (Bok, 2013, 333).

The reverse causality direction is also important but here we find different results that point to a negative relation. Several studies show a negative relation between teaching load (usually measured through the student/staff ratio) and research output and/or quality, suggesting that beyond a certain threshold teaching and research become substitutes, rather than complement (Marsh, 1984; Hattie and Marsh, 1996; Becker and Kennedy, 2006; Fayerweather, 2002; 2005; Green, 2008; Fender, Taylor and Burke, 2011; Bak and Kim, 2015).

From an empirical point of view, it is important to control for some measure of teaching load, in order to take into account the possibility of a negative effect of excess teaching on research.

In sum, the literature criticizes the vision of ‘one-size-fits all’ university model that characterizes HEIs as centres of excellence in teaching, research and outreach activities (Sánchez-Barrioluengo, 2014), mainly because this model assumes that missions are carried out in an interconnected way and universities combine them to fulfil expectations in their contribution to society without taking into account the differences between higher education systems and the heterogeneity among regions and countries.

### 2.2.8 Complementarity between human capital accumulation and innovation

The above discussion has shed light on several specific mechanisms through which the activities of HEIs impact on firms. In this section we point to an additional element, which has recently gained attention: in absorbing external knowledge produced by HEIs firms have the problem of addressing jointly the improvement of their human capital and the management of the expanding technological knowledge. If there is complementarity between human capital accumulation and innovative activities, then the overall impact may depend significantly on the management capabilities of firms.

Firms hire graduate workers (hence absorb knowledge in embodied form) and interact with researchers in a large variety of ways, from informal face to face interaction to consulting, from contract research to co-publication and co-invention (hence absorb new knowledge).
How do firms make productive use of this knowledge? This is an important question, which has triggered lot of research effort. We summarize some of the main findings, as follows:

i) there is large heterogeneity among firms in the way they carry out innovative activities, as well as in the way in which innovation, human capital accumulation and investment in physical capital are (or not) patterned and sequenced

ii) the external orientation of firms (towards other companies as partners, as well as towards universities and research centres) is an antecedent of innovative orientation

iii) there is complementarity between investment into the accumulation of human capital and the introduction of innovation

iv) there is complementarity between product and process innovation

v) innovation and internationalization are dynamically linked, so that, once one of them is triggered (sometimes for contingent reasons) and goes beyond a given threshold, the other is also triggered and the two activities mutually reinforce each other

vi) innovation and accumulation of physical capital can be both complements and substitutes in generating permanent increase in productivity

vii) multinational corporations (MNCS) have an advantage in the management of human resources (hence the productive use of human capital) insofar as they transfer across countries the best practices, in order to increase productivity

Keeping into account all these factors in models of spill-over effects at territorial level is extremely difficult. We discuss below the potential and limits of the available datasets in order to address these issues.
3. Empirical research strategy

3.1 Research questions

Based on the survey of the literature, we delineate in this section the overall research strategy and the main research questions, highlighting the empirical approach that we propose.

We address four main research questions. We also discuss the main limitations of the data available and some possible solution (but also their limits). The main research questions are as follows:

   a) Impact of HEIs education activities on firm performance
   Does having a HEIs located close to firms have an impact on firm performance due to the access to qualified human capital?

   b) Impact of HEIs research activities on firm performance
   Does having a HEIs located close to firms have an impact on firm performance due to the access to new knowledge?

   c) Impact of quality of education and research
   To what extent the impact of HEIs on firms depend on the quality of their education and research production?

   d) Role of geographic proximity
   Does the impact of HEIs on firms decay with geographic distance? If yes, which is the maximum distance for the effect of spill-overs? If yes, do we see differences in the importance of proximity between education and research, and between activities of different quality?

In the context of the present study by firm performance we mean observable changes in several economic and financial variables, observed at microlevel, that is, at the level of individual firms. The variables include turnover, employment, total assets, total intangible assets. The choice of these variables, as discussed at length below, depends largely on the availability of data and the need to have a balanced coverage across countries. Taken together they offer a rich and articulated representation of the performance of firms in both manufacturing industry and service sectors.

Another key preliminary remark refers to the notion of impact.

In the endogenous growth literature the explanatory variables related to human capital and R&D are formulated in terms of stock. Empirical studies relate the stock of human capital, as measured by the number of years of education accumulated by the population (hence the stock of knowledge embodied in the population at any given date), with an appropriate rate of discount or decay, representing the effect of obsolescence of knowledge acquired during formal education. Similarly, in the large literature on the impact of R&D on growth, the R&D intensity measure (or the ratio between expenditure in R&D and GDP) is used in order to build up a measure of R&D stock, or stock of knowledge. This is consistent, of course, with the adoption of a knowledge production function approach, in which the traditional variables labor and capital are qualified by using the stock of human capital and the stock of knowledge.

We are not in a position to measure the stock of human capital or R&D at the microlevel of firms. We do not have reliable data on employment with a breakdown by level of education. Hence we cannot disentangle between the current employment situation (stock) and the flow of new hires, by distinguishing those newly recruited units that enter the firm in any given year by their level of education. We only observe the pool of potential candidates offered each year by HEIs located close to the firm. Therefore our notion of “impact” is somewhat looser than the notion of impact in the tradition of
counterfactual causal analysis. We control for a large number of factors, but not to the point to establish causality in a formal sense.

Similarly, we do not have data on the stock of knowledge derived from some form of capitalization of past investment into R&D. Microdata on R&D investment are only available based on surveys, not on official censuses or administrative databases. Therefore, we only observe the pool of available knowledge, but we do not have certainty about the fact that any given company in the dataset has, actually and demonstrably, made use of the knowledge generated by a university in the nearby.

These limitations make it safe to argue that we will be able to find rich descriptive evidence and some broad generalization on association of variables. At the same time, since we are dealing with the universe of higher education in Europe (i.e. ETER data cover all HEIs by an officially validated census) and with a large sample of firms, we believe that these generalizations hold large validity.

3.2 Description of data: ORBIS, ETER

3.2.1 Orbis dataset

We plan to evaluate the effects of HEIs on local development mostly using microdata on firms' performance.

To do this we choose to work with ORBIS-Bureau van Dijk, the single largest database with microdata on firms across European countries.

In order to select the performance variables for our exercise we first examine the set of variables available in Orbis (Table 1).

We have 18,069,103 observations in the data, which include firms with some information for at least one of the variables described in Table 1 in the period 2011-2015. However, not all these variables are covered equally well.

Table 2 shows the size of the sample for a subset of variables that might be used as performance variables, along with the mean and standard deviation of the variable itself.

For most candidate dependent variables the coverage improves significantly during the sample period.

We therefore select the performance variables according to two criteria: (a) survey of the literature; (b) coverage in the dataset. Based on these criteria we identify the following performance variables:

   a) Operating turnover
   b) Number of employees
   c) Total assets
   d) Fixed intangible assets

The coverage for some variables is very poor (Current market capitalisation and Research and Development expenditures) and limited to very large companies. Therefore, it would be difficult to use them in the regression analysis.
<table>
<thead>
<tr>
<th>Variable name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>bvdidnumber</td>
<td>Bureau van Dijk ID number – firm’s identifier</td>
</tr>
<tr>
<td>nacerev2corecode4digits</td>
<td>NACE code for industrial sector (4 digits)</td>
</tr>
<tr>
<td>countryisocode</td>
<td>Country ISO code</td>
</tr>
<tr>
<td>consscode</td>
<td>Consolidation code for financial data</td>
</tr>
<tr>
<td>lastavailyear</td>
<td>Last year of available information for a firm</td>
</tr>
<tr>
<td>City</td>
<td>Firm location – city</td>
</tr>
<tr>
<td>Latitude</td>
<td>Firm location – latitude</td>
</tr>
<tr>
<td>longitude</td>
<td>Firm location – longitude</td>
</tr>
<tr>
<td>nuts2</td>
<td>Firm location – Nuts 2 region</td>
</tr>
<tr>
<td>nuts3</td>
<td>Firm location – Nuts 3 region</td>
</tr>
<tr>
<td>standardisedlegalform</td>
<td>Standardised legal form (e.g. private limited companies, non profit...)</td>
</tr>
<tr>
<td>numberofavailableyears</td>
<td>Number of available years</td>
</tr>
<tr>
<td>date_incorporation</td>
<td>Year of firm creation</td>
</tr>
<tr>
<td>categoryofthecompany</td>
<td>Category of the company – Size (4 classes from small to very large)</td>
</tr>
<tr>
<td>operatingrevenueturnoverlastvalue</td>
<td>Operating revenue (Turnover) – last available value (th EUR)</td>
</tr>
<tr>
<td>numberofemployeeslastvalue</td>
<td>Number of employees – last available value</td>
</tr>
<tr>
<td>currentmarketcapitalisationtheur</td>
<td>Current market capitalisation – th EUR</td>
</tr>
<tr>
<td>noofcompaniesincorporategroup</td>
<td>Number of companies in corporate group</td>
</tr>
<tr>
<td>noofrecordedshareholders</td>
<td>Number of recorded shareholders</td>
</tr>
<tr>
<td>noofrecordedsubsidiaries</td>
<td>Number of recorded subsidiaries</td>
</tr>
<tr>
<td>noofrecordedbranchlocations</td>
<td>Number of recorded branch locations</td>
</tr>
<tr>
<td>numberoftrademarks</td>
<td>Number of trademarks</td>
</tr>
<tr>
<td>Operatingrevturnovertheur201*</td>
<td>Operating revenue (Turnover) th EUR 2011-2015</td>
</tr>
<tr>
<td>Numberofemployees201*</td>
<td>Number of employees 2011-2015</td>
</tr>
<tr>
<td>plbeforetaxtheur201*</td>
<td>Profit and loss before tax – th EUR 2011-2015</td>
</tr>
<tr>
<td>Piforperiodnetincometheur201*</td>
<td>Profit and loss for period (Net income) th EUR 2011-2015</td>
</tr>
<tr>
<td>Totalassetstheur201*</td>
<td>Total assets – th EUR 2011-2015</td>
</tr>
<tr>
<td>Intangiblefixedassetstheur201*</td>
<td>Intangible fixed assets – th EUR 2011-2015</td>
</tr>
<tr>
<td>RDtheur201*</td>
<td>Research and development expenses - th EUR 2011-2015</td>
</tr>
</tbody>
</table>

Source: Orbis dataset, 2018.
Table 2. Summary statistics for relevant variables

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Sample size</th>
<th>Mean</th>
<th>Standard dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>currentmarketcapitalisationtheur</td>
<td>13,988</td>
<td>833531.2</td>
<td>5305622</td>
</tr>
<tr>
<td>nofcompaniesincorporategroup</td>
<td>18,054,972</td>
<td>8.588</td>
<td>122.3</td>
</tr>
<tr>
<td>Noofrecordedshareholders</td>
<td>18,069,103</td>
<td>1.18</td>
<td>4.14</td>
</tr>
<tr>
<td>Noofrecordedsubsidiaries</td>
<td>18,069,103</td>
<td>0.18</td>
<td>4.50</td>
</tr>
<tr>
<td>noofrecordedbranchlocations</td>
<td>18,069,103</td>
<td>0.15</td>
<td>5.20</td>
</tr>
<tr>
<td>Numberoftrademarks</td>
<td>18,069,103</td>
<td>0.04</td>
<td>1.94</td>
</tr>
<tr>
<td>Operatingreventurnovertheur2011</td>
<td>5,143,615</td>
<td>5986.61</td>
<td>408881.6</td>
</tr>
<tr>
<td>Operatingreventurnovertheur2012</td>
<td>5,776,296</td>
<td>5677.92</td>
<td>388061.5</td>
</tr>
<tr>
<td>Operatingreventurnovertheur2013</td>
<td>7,910,807</td>
<td>4359.02</td>
<td>317388.1</td>
</tr>
<tr>
<td>Operatingreventurnovertheur2014</td>
<td>8,402,280</td>
<td>4385.60</td>
<td>310283.8</td>
</tr>
<tr>
<td>Operatingreventurnovertheur2015</td>
<td>8,749,030</td>
<td>4519.98</td>
<td>254420.3</td>
</tr>
<tr>
<td>Numberofemployees2011</td>
<td>4,960,693</td>
<td>23.75</td>
<td>962.1</td>
</tr>
<tr>
<td>Numberofemployees2012</td>
<td>5,316,395</td>
<td>23.17</td>
<td>908.4</td>
</tr>
<tr>
<td>Numberofemployees2013</td>
<td>7,386,561</td>
<td>18.05</td>
<td>800.5</td>
</tr>
<tr>
<td>Numberofemployees2014</td>
<td>8,461,385</td>
<td>16.52</td>
<td>750.4</td>
</tr>
<tr>
<td>Numberofemployees2015</td>
<td>8,897,479</td>
<td>16.69</td>
<td>734.4</td>
</tr>
<tr>
<td>plbeforetaxtheur2011</td>
<td>4,120,162</td>
<td>603.4</td>
<td>61478.2</td>
</tr>
<tr>
<td>plbeforetaxtheur2012</td>
<td>4,550,902</td>
<td>464.0</td>
<td>54940.5</td>
</tr>
<tr>
<td>plbeforetaxtheur2013</td>
<td>5,141,902</td>
<td>518.1</td>
<td>51528.5</td>
</tr>
<tr>
<td>plbeforetaxtheur2014</td>
<td>5,619,067</td>
<td>519.7</td>
<td>52015.4</td>
</tr>
<tr>
<td>plbeforetaxtheur2015</td>
<td>5,772,141</td>
<td>485.6</td>
<td>52154.2</td>
</tr>
<tr>
<td>plfperiodnetincome2011</td>
<td>4,117,993</td>
<td>475.2</td>
<td>48389.1</td>
</tr>
<tr>
<td>plfperiodnetincome2012</td>
<td>4,546,054</td>
<td>360.9</td>
<td>45858.4</td>
</tr>
<tr>
<td>plfperiodnetincome2013</td>
<td>5,135,408</td>
<td>447.5</td>
<td>59678.5</td>
</tr>
<tr>
<td>plfperiodnetincome2014</td>
<td>5,618,103</td>
<td>432.3</td>
<td>47904.4</td>
</tr>
<tr>
<td>plfperiodnetincome2015</td>
<td>5,776,457</td>
<td>406.5</td>
<td>49841.1</td>
</tr>
<tr>
<td>totalassetstheur2011</td>
<td>6,884,954</td>
<td>19133.3</td>
<td>2868351</td>
</tr>
<tr>
<td>totalassetstheur2012</td>
<td>7,657,086</td>
<td>18313.9</td>
<td>2768057</td>
</tr>
<tr>
<td>totalassetstheur2013</td>
<td>8,644,091</td>
<td>17421.3</td>
<td>2441091</td>
</tr>
<tr>
<td>totalassetstheur2014</td>
<td>9,646,028</td>
<td>17465.8</td>
<td>2517263</td>
</tr>
<tr>
<td>totalassetstheur2015</td>
<td>10,378,260</td>
<td>17206.5</td>
<td>2391344</td>
</tr>
<tr>
<td>Intangiblefixedassetstheur2011</td>
<td>5,898,932</td>
<td>559.2</td>
<td>79824.1</td>
</tr>
<tr>
<td>Intangiblefixedassetstheur2012</td>
<td>6,496,861</td>
<td>537.4</td>
<td>71050.4</td>
</tr>
<tr>
<td>Intangiblefixedassetstheur2013</td>
<td>7,273,725</td>
<td>499.7</td>
<td>67746.7</td>
</tr>
<tr>
<td>Intangiblefixedassetstheur2014</td>
<td>8,105,163</td>
<td>506.4</td>
<td>70733.9</td>
</tr>
<tr>
<td>Intangiblefixedassetstheur2015</td>
<td>8,611,094</td>
<td>573.4</td>
<td>79439.2</td>
</tr>
<tr>
<td>RDtheur2011</td>
<td>9,953</td>
<td>13565.5</td>
<td>147963.5</td>
</tr>
<tr>
<td>RDtheur2012</td>
<td>11,041</td>
<td>12860.6</td>
<td>140342.6</td>
</tr>
<tr>
<td>RDtheur2013</td>
<td>12,171</td>
<td>11140.1</td>
<td>121384.1</td>
</tr>
<tr>
<td>RDtheur2014</td>
<td>11,238</td>
<td>11809.8</td>
<td>135918.8</td>
</tr>
<tr>
<td>RDtheur2015</td>
<td>9,997</td>
<td>14868.2</td>
<td>162518.2</td>
</tr>
</tbody>
</table>

Source: Orbis dataset, 2018.

The four performance variables capture different dimensions of performance:

- Operating turnover is a measure of effectiveness in operating in the market; the growth in turnover is an indicator of relative advantage with respect to other firms and of expansion;
- Employment captures the degree to which the expansion in turnover is reflected in, or anticipated by, the growth in the labour factor;
- Total assets deliver a measure of investment, which may grow over time due to investment decision into long term assets (fixed assets) or as a result of the
expansion of operational activities (working capital, such as cash, credit, and inventory);

- Intangible fixed assets is a proxy for innovative activities, since it is largely formed by intellectual property (patent, copyright, trademark, design), capitalization of long term R&D expenditures, and goodwill.

Table 3 offers an analysis of the coverage of the dataset for the central year of the period (i.e. 2013; data are available on each year) for the four variables.

Table 3. Coverage by country of selected dependent variables

<table>
<thead>
<tr>
<th>Country</th>
<th>Total active today</th>
<th>Total with some information</th>
<th>Turnover 2013</th>
<th>Employees 2013</th>
<th>Total Assets 2013</th>
<th>Intangible Assets 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>1,120,283</td>
<td>203,590</td>
<td>36,315</td>
<td>79,783</td>
<td>117,166</td>
<td>103,839</td>
</tr>
<tr>
<td>Belgium</td>
<td>2,518,679</td>
<td>414,075</td>
<td>42,255</td>
<td>128,504</td>
<td>356,100</td>
<td>336,265</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>1,559,807</td>
<td>494,538</td>
<td>331,331</td>
<td>364,838</td>
<td>254,518</td>
<td>254,497</td>
</tr>
<tr>
<td>Croatia</td>
<td>268,167</td>
<td>103,276</td>
<td>81,130</td>
<td>59,396</td>
<td>81,131</td>
<td>81,084</td>
</tr>
<tr>
<td>Cyprus</td>
<td>211,736</td>
<td>3,109</td>
<td>1,851</td>
<td>1,038</td>
<td>2,407</td>
<td>2,077</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>2,548,099</td>
<td>1,430,291</td>
<td>1,069,619</td>
<td>225,327</td>
<td>159,084</td>
<td>158,952</td>
</tr>
<tr>
<td>Denmark</td>
<td>962,704</td>
<td>215,914</td>
<td>37,604</td>
<td>23,015</td>
<td>159,084</td>
<td>157,391</td>
</tr>
<tr>
<td>Estonia</td>
<td>266,160</td>
<td>122,444</td>
<td>319,230</td>
<td>298,318</td>
<td>139,557</td>
<td>122,796</td>
</tr>
<tr>
<td>Finland</td>
<td>1,371,903</td>
<td>463,893</td>
<td>225,327</td>
<td>660,597</td>
<td>456,844</td>
<td>400,099</td>
</tr>
<tr>
<td>France</td>
<td>12,215,179</td>
<td>903,450</td>
<td>720,752</td>
<td>720,110</td>
<td>720,110</td>
<td>720,110</td>
</tr>
<tr>
<td>Germany</td>
<td>3,129,042</td>
<td>1,077,510</td>
<td>348,825</td>
<td>660,597</td>
<td>456,844</td>
<td>400,099</td>
</tr>
<tr>
<td>Greece</td>
<td>844,114</td>
<td>31,592</td>
<td>24,834</td>
<td>189,340</td>
<td>332,463</td>
<td>238,318</td>
</tr>
<tr>
<td>Hungary</td>
<td>1,382,423</td>
<td>372,746</td>
<td>295,631</td>
<td>189,340</td>
<td>332,463</td>
<td>238,318</td>
</tr>
<tr>
<td>Iceland</td>
<td>246,062</td>
<td>137,212</td>
<td>20,040</td>
<td>27,325</td>
<td>110,973</td>
<td>75,663</td>
</tr>
<tr>
<td>Italy</td>
<td>4,486,990</td>
<td>3,007,688</td>
<td>2,270,397</td>
<td>718,300</td>
<td>716,405</td>
<td>716,405</td>
</tr>
<tr>
<td>Latvia</td>
<td>236,239</td>
<td>110,127</td>
<td>88,969</td>
<td>88,695</td>
<td>88,677</td>
<td>88,677</td>
</tr>
<tr>
<td>Lithuania</td>
<td>150,703</td>
<td>101,725</td>
<td>53,537</td>
<td>82,924</td>
<td>10,670</td>
<td>10,281</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>161,856</td>
<td>19,408</td>
<td>7,272</td>
<td>12,228</td>
<td>11,425</td>
<td>11,425</td>
</tr>
<tr>
<td>Malta</td>
<td>56,135</td>
<td>16,928</td>
<td>14,164</td>
<td>400,137</td>
<td>498,489</td>
<td>498,489</td>
</tr>
<tr>
<td>Netherlands</td>
<td>2,987,438</td>
<td>1,456,983</td>
<td>14,164</td>
<td>595,488</td>
<td>498,489</td>
<td>498,489</td>
</tr>
<tr>
<td>Poland</td>
<td>1,891,237</td>
<td>1,308,394</td>
<td>99,374</td>
<td>687,049</td>
<td>107,182</td>
<td>80,806</td>
</tr>
<tr>
<td>Portugal</td>
<td>688,966</td>
<td>344,654</td>
<td>257,201</td>
<td>231,737</td>
<td>279,664</td>
<td>231,123</td>
</tr>
<tr>
<td>Romania</td>
<td>1,506,607</td>
<td>606,258</td>
<td>492,187</td>
<td>492,196</td>
<td>492,158</td>
<td>492,158</td>
</tr>
<tr>
<td>Slovakia</td>
<td>637,935</td>
<td>330,496</td>
<td>237,392</td>
<td>185,030</td>
<td>142,702</td>
<td>142,609</td>
</tr>
<tr>
<td>Slovenia</td>
<td>279,377</td>
<td>118,018</td>
<td>92,551</td>
<td>43,757</td>
<td>98,382</td>
<td>71,956</td>
</tr>
<tr>
<td>Spain</td>
<td>3,562,261</td>
<td>815,983</td>
<td>621,567</td>
<td>479,840</td>
<td>686,245</td>
<td>641,361</td>
</tr>
<tr>
<td>Sweden</td>
<td>1,933,854</td>
<td>629,899</td>
<td>334,352</td>
<td>375,610</td>
<td>52,724</td>
<td>44,279</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>4,252,518</td>
<td>1,136,951</td>
<td>172,593</td>
<td>80,372</td>
<td>1,598,526</td>
<td>1,121,221</td>
</tr>
</tbody>
</table>

Source: Orbis dataset, 2018.

The table shows the differences in coverage across countries in year 2013. We marked in bold the cases in which the number of non-missing observations is below 10% of the total number of firms in the country, reflecting the fact that it would be problematic to make inference on the population with this relatively small (and usually non-random) sample. For some countries (e.g. Cyprus and Greece) only a small share of companies is registered in the Orbis database (see Column (2) to Column (1) ratio), while for some others (e.g. Italy) coverage is above 50%. There are also large differences across variables, for example using Turnover as a variable measuring firms performance would work well for Italy or Czech Republic but do not for Netherlands or the United Kingdom.

For each of the selected performance variable we will calculate the following operationalizations:
i) Absolute growth
   i. Difference in variable level 2011-2015
   ii. Difference in variable level per each year 2011-2015

ii) Rate of growth
   i. Difference in variable level/ level at initial year 2011-2015
   ii. Difference in variable level/ level at initial year per each year 2011-2015

In order to take into account the size-dependence of rates of growth (i.e. the well-known fact that rates of growth are larger for small units) we include in the control variables the value of the performance variable at initial year. Absolute growth and rates of growth can have negative or positive sign.

Figure 1 shows the maps for the country average values of the selected variables capturing firm performance. Darker colours indicate higher values for the average of turnover, employment, total assets and fixed assets in each country.

**Figure 1. Country average of firm performance in 2013**

<table>
<thead>
<tr>
<th>a) Turnover</th>
<th>b) Employment - 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Map of Turnover" /></td>
<td><img src="image2.png" alt="Map of Employment" /></td>
</tr>
<tr>
<td>c) Total Assets</td>
<td>d) Intangible Fixed Assets</td>
</tr>
<tr>
<td><img src="image3.png" alt="Map of Total Assets" /></td>
<td><img src="image4.png" alt="Map of Intangible Fixed Assets" /></td>
</tr>
</tbody>
</table>

Figure 1a presents average Turnover in 2013 in EU countries. Average turnover is higher in richer countries (they have, on average, larger and more productive firms) and lower in countries with poorer coverage (countries with a low coverage include only relatively large firms in the data).
Figure 1b shows a very similar figure for average employment. In Figure 1c we map average values for total assets of companies. Finally, Figure 1d represents the average level of intangible fixed assets in EU countries.

3.2.2 ETER

The explanatory variables of interest in our regressions are measures of HEIs volume and quality of both research and teaching. In order to find trustworthy measures we rely on two different sources of data. The first one is the ETER database, collecting information on European HEIs in terms of students, academic staff and financing disaggregated by field of education. The data include some information about 2,895 European Institutions. ETER includes several dozen variables. A subset of them, described in Table 4, can be used for our purposes.

Based on the theoretical discussion in Section 2, we use as independent variables the number of graduates at ISCED level 6-8 and the number of academic staff. The location variables are used to calculate the geographic distance between the HEIs and firms.

ETER data are integrated with data on publications and citations disaggregated by university and field of study for the period 2011-2015, extracted from Scopus.

The number of publications refer to the production in the 2011-2015 period, while the number of citations refer to the citations received in the 2011-2015 period to publications authored by the institution, published in the same period. In the regression model we make use of the citation variable, which is a measure of quality of research largely accepted in the literature.

3.3 Description of main variables and methodological issues

Before entering into the analysis of independent and dependent variables, it is useful to clarify the aggregation strategy. As already stated, one of the main advantages of the microdata approach is the level of granularity of observations. This however creates an obvious problem of spatial aggregation of data.

The variables are aggregated at spatial level by merging together HEIs that are located at a given distance from firms. We build up a distance matrix between all HEIs in the ETER dataset and all firms in the Orbis dataset. Then we start from the location of the individual firm and we ask which (if any) HEIs are located within fixed distance intervals (see below for the details). If we find HEIs located within these circles, we aggregate their variables (by using sum or averages depending on the nature of the variable). The resulting aggregation is a sort of “virtual” institution, whose components share the feature that they are at a similar distance from the focal firm. We might label them “isodistance” HEIs, when the prefix “iso” is interpreted in the terms of geographic areas of fixed radius from the centre. With this clarification in mind, let us examine the variables more in detail.

3.3.1 Independent variables

(a) Education activities of HEIs

To capture educational activities of HEIs we use the number of graduates, which well reflects the output of higher education.

Graduates are classified in two ways (see the ETER Handbook: all definitions and classifications are available in Appendix 2):

— by level of education (ISCED level). distinguishing between short courses (2 years) at ISCED5, bachelor (three years) at ISCED6, master (3+2 years) at ISCED7 and doctoral education at ISCED8;
— by Field of Education (FoE), based on the UNESCO definition of classes.
Table 4. Variables selected from ETER

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eterid</td>
<td>ETER ID – HEI identifier</td>
</tr>
<tr>
<td>englishinstitutionname</td>
<td>English Institution name</td>
</tr>
<tr>
<td>countrycode</td>
<td>Country code</td>
</tr>
<tr>
<td>foundationyear</td>
<td>Foundation year</td>
</tr>
<tr>
<td>regionofestablishmentnuts2</td>
<td>Location – NUTS 2</td>
</tr>
<tr>
<td>regionofestablishmentnuts3</td>
<td>Location – NUTS 3</td>
</tr>
<tr>
<td>nameofthecity</td>
<td>Location – Name of the city</td>
</tr>
<tr>
<td>geographiccoordinateslatitude</td>
<td>Location – latitude</td>
</tr>
<tr>
<td>geographiccoordinateslongitude</td>
<td>Location – longitude</td>
</tr>
<tr>
<td>postcode</td>
<td>Location – postcode</td>
</tr>
<tr>
<td>multisiteinstitution</td>
<td>Multi-site institution</td>
</tr>
<tr>
<td>nuts3codesofothercampuses</td>
<td>NUTS 3 codes of other campuses</td>
</tr>
<tr>
<td>academicstaff*</td>
<td>Academic staff (disaggregated by university and field of study for years 2011-2014)</td>
</tr>
<tr>
<td>graduatesatisced5*</td>
<td>Graduates at Isced 5 (disaggregated by university and field of study for years 2011-2014)</td>
</tr>
<tr>
<td>graduatesatisced6*</td>
<td>Graduates at Isced 6 (disaggregated by university and field of study for years 2011-2014)</td>
</tr>
<tr>
<td>graduatesatisced7*</td>
<td>Graduates at Isced 7 (disaggregated by university and field of study for years 2011-2014)</td>
</tr>
<tr>
<td>graduatesatisced7ld*</td>
<td>Graduates at Isced 7 long degree (disaggregated by university and field of study for years 2011-2014)</td>
</tr>
<tr>
<td>graduatesatisced8*</td>
<td>Graduates at Isced 8 (disaggregated by university and field of study for years 2011-2014)</td>
</tr>
<tr>
<td>researchactiveinstitution</td>
<td>Research active institution dummy</td>
</tr>
<tr>
<td>rdexpenditureeuro*</td>
<td>Research and development expenditure (euros)</td>
</tr>
<tr>
<td>rdexpenditureppp*</td>
<td>Research and development expenditure (ppp)</td>
</tr>
</tbody>
</table>

Given this articulation, we use as independent variables the sum of the total number of undergraduate and master graduates (level 5-7), plus the number of graduates at doctoral level (ISCED 8). When relevant, this sum is also expressed by field of study, following the Field of Education (FoE) classification:

- Social Sciences and Humanities
- Business and Economics
- Medicine
- Technology
- Natural Sciences

It is useful to remark that, for the time being, it is not possible to control directly for the quality of education. Surveys on employability of HEIs graduates are still fragmented and non-comparable across countries. Quality Assurance data on the quality of curricula are even more fragmented, although there are efforts at European level for standardizing definitions and possibly integrating data.

(b) Research activities of HEIs

By research we mean, broadly speaking, the production of new knowledge, and there are several qualifications that point to a very articulated definition of variables.

First, knowledge production differs widely by disciplines. This is an established point in the literature: knowledge is a highly specific economic good, which is not infinitely
malleable. Producing new knowledge in molecular biology differs dramatically from producing knowledge in political science. We use several forms of classification of citations based on Subject Categories of publications (corresponding broadly to scientific fields, and largely utilized in the bibliometric literature). The classification is summarized in few categories that might be reconciled with the classification of educational activities (Field of Education, FoE), and which makes it possible to distinguish between STEM (Science, Technology, Engineering, Mathematics) and SSH (Social Sciences and Humanities).

Second, new knowledge is embedded into several channels.

Embodied knowledge can be approximated by academic staff. People that have an institutional relation with HEIs (full time or part time, at any level of the academic career) have a mandate to produce and circulate new knowledge. They produce research, publish books and papers, go to conferences, consult companies and public administrations, talk in newspapers and radio programmes, teach to students- in one word, knowledge circulates with them all the time.

Disembodied knowledge, by contrast, is delivered in formal textual form, and becomes a piece of codified knowledge that circulates without the personal involvement of their authors. Once published, knowledge becomes the property of everybody that reads the publication, everywhere in the world.

Third, knowledge production is heterogeneous in itself. The most useful and productive kind of knowledge is the one associated to discoveries, that is, to scientific evidence that opens new directions for exploration. Another important type of knowledge is given by methodologies and techniques, in particular experimental ones. In contrast, many research activities just confirm and support existing theories and models, providing cumulative evidence within established directions. Summing up, the impact of research is highly variable.

It has become common practice, after an extensive theoretical and empirical investigation started in the '60s, to approximate the impact of research with the citations received by any single piece of knowledge- in practice, by any article published in scientific journals. Aggregating citations is a commonly used measure of research impact at various scales- individual, research team, department, university and country. Citations are to the measurement of knowledge what GDP is for the measurement of economic activity: they allow the aggregation of measures that come from largely heterogeneous research activities. Standardized citations allow the comparison of the research quality across fields, insofar as they give a measure of the relative impact of a given unit (article, author, university, country) with respect to all the others.

In this report we aggregate citations at the level of universities, that is, we sum all citations received by articles published by authors affiliated to the university in subsequent years. This aggregation gives a fairly accurate picture of the overall impact of research produced at any given university.

Consequently, with respect to research activity, the independent variables are as follows:

— Number of citations
— Number of academic staff

Citations can also be classified by field, using Subject Categories, an international standard for the classification of journals:

- Social Sciences and Humanities
- Basic Sciences
- Medicine
- Technology
A similar classification by field for academic staff is not available.

We have a slight distinction between the classification for graduates (education) and that for citations (research). For citations, Social Sciences and Humanities include Business and economics, while Basic Sciences include Mathematics, Physics and Chemistry. Conversely, graduates in Natural sciences include graduates in Mathematics, Physics or Chemistry (i.e. Basic Sciences), and graduates in Business and Economics are classified separately.

It should be noted that the dataset allows further exploration of models that combine several variables, as for example using per capita output variables (number of publications or citations per unit of academic staff) or combining education and research activities (number of student per unit of academic staff). All these analyses are left for future research.

3.3.2 Dependent variables

All models are defined by using microdata, with firm’s performance as a unit of analysis. This is perhaps the most important advantage of the dataset with respect to the large literature on regional impact of knowledge spillover. By using microdata on firms we are able to disentangle several dimensions of the knowledge spillovers.

As already stated we identify the following performance variables:

a) Operating turnover
b) Number of employees
c) Total assets
d) Fixed intangible assets

The growth in the dependent variables has been examined with several alternative operationalizations:

- Absolute growth (difference 2011-2015 in monetary value in euro; difference by year in the 2011-2015 period) in natural and log form;
- Relative growth (total rate of variation 2011-2015; rate of variation by year in the 2011-2015 period)

In the following we present a selection of results.

The growth over time of the variable is interpreted as a “performance variable” and captures several elements of firms' performance: economic, intellectual capital and employment. Firms that grow in their total assets exhibit an increasing level of activity, resulting in either fixed investment (tangible and intangible), or working capital, or both. The meaning of total assets is established in the literature. Since total assets must be covered by financial measures (equity or debt), which are costly (explicitly or implicitly), this means that the growth in total assets follows an increase in operational activities. The growth in total assets has proven to be a reliable measure across several European countries in the study by Bonaccorsi et al. (2018).

Intangible assets are used as a proxy for intellectual capital and for the creation of intellectual property rights. They include patents, trademarks, copyright, design, as well as goodwill. There are national variations in the way in which these assets are defined and measured (mainly due to fiscal reasons), but there is sufficient agreement on the interpretation of data. Growth in intangible assets is associated to innovative activities.

Finally, we use employment data, for which sufficient coverage is available. Growth in employment is a performance variable, although in the interpretation of data it is
necessary to consider the potentially confounding effects of changes in legislation and labour regulation.

These dimensions of performance are only partially correlated since they refer to different aspects of activities of firms and move in asynchronous ways. The relation between them is very complex and is the object of a dedicated literature, which has not yet come to a unifying explanation. One might say that firms first observe whether their turnover increases over time, usually waiting for the stabilization of short term fluctuations, then increase their total assets (in particular, working capital in terms of credits, cash and stocks), only at the end increase their employment and make investments into intangible assets. Another important difference is that turnover depends both on demand and supply (plus product pricing decisions) while assets mostly depend upon firms' decision, incorporating expectations and the price of capital. In the years covered by the analysis the price of capital has been always quite low. However expectations for many countries have not been very good. In fact current turnover can be taken as a signal for future profitability and hence used to make investment decisions.

On the other hand, the increase in turnover may be generated from previous investments into intangible assets, for example by trademarks as elements of a long term marketing strategy. At the same time the increase in total assets is mediated by several variables, such as the average time of payment (which may depend on national legislation, business practices, and bank attitudes), the rate of rotation of stocks (which is influenced by the length of the production process and the adoption of just in time and lean management approaches), and the availability of unused productive capacity. In the same line of reasoning, the increase in employment is mediated by the features of the labour market (wage determination processes, hiring costs, firing costs, flexibility) and the national legislation, as well as the supply of labour. The same complex dynamics might be in place when the turnover decreases, temporarily or for many periods in line. Also, the behaviour of firms may also change, depending on the volatility of the economic environment in which they operate.

Summing up, the identification of structural effects of the presence of HEIs would require a more elaborated modelling strategy for the performance of firms and the observation of firm data for a long period. This goes beyond the scope of the present project.

An important methodological issue to be discussed here refers to the heterogeneity of firms with respect to the impact of HEIs activities. As largely discussed in the literature and reported above, these effects depend critically on industry-specific variables and on firm size.

There are several potential models to be considered here: a) an aggregate model; and b) a sectoral model. In case a) we assume that what matters is the total number of students or graduates (that is, we implicitly admit some level of inter-sectoral mobility of students and graduates: i.e. student with a degree in Physics may work not only in science-based industries but, say, in the Leisure industry). In case b) we might take into account the statistical classification of economic activities of the firms and run separate models at different NACE main section level and/or use industry dummy variables. In this Report we do not pursue modelling strategy b), but only use sectoral variables as control variables (dummies).

### 3.3.3 Control variables

Introducing control variables has two main effects. First, it wipes out the variability of the dependent variable originated by the variability of the control. Second, it allows to identifying possible effects from the control to the dependent variables that deserve close scrutiny.

All control variables are taken as dummies, so that it is possible to include a long list of variables without detracting from the coefficients.

The control variables are defined as follows.
Size of the firm (reference = Small company)
- Small company
- Medium-sized company
- Large company
- Very large company

Source: ORBIS dataset

Sector of the firm (reference = A-Agriculture)
A - Agriculture
B - Mining
C - Manufacturing
D - Electricity
E - Water supply
F - Construction
G - Wholesale and retail
H - Transportation
I - Accommodation
J - Information
K - Financial
L - Real estate
M - Professional services
N - Administration
O - Public administration
P - Education
Q - Human health
R - Arts, entertainment
S - Other services
T - Other activities

Source: ORBIS dataset based on NACE-rev2 main industry classification

When controlling for different sectors, we have used two approaches: (a) include dummies for individual sectors; (b) include dummies that captures the aggregation of sectors into four macro-sectors (Agriculture, Non-manufacturing industry; Manufacturing industry; Services). The aggregation does not change the sign of coefficients of the dependent variables. In the tables below we refer to the specification with individual sectors.

Legal form of the firm (reference = Foreign companies)
- Non-profit organisations
- Other legal forms
- Partnerships
- Private limited
- Public authority
- Public limited
- Sole traders

Source: ORBIS dataset

Country (reference= Austria)
- Belgium
- Bulgaria
- Czech Republic
- Germany
- Estonia
- Spain
- Finland
- France
- United Kingdom
- Greece
- Hungary
- Ireland
- Italy
- Lithuania
- Latvia
- Luxembourg
- Malta
- Cyprus
- Netherlands
- Poland
- Portugal
- Sweden
- Slovenia
- Slovak Republic

Region: NUTS 2 classification (reference = AT11, Burgenland)
- List of regions omitted

Source: EUROSTAT

Institutional quality
- Dummy 1= countries with high institutional quality
- Dummy 0= countries with low institutional quality

Source: European Quality of Government Index

R&D intensity
- Dummy 1= high R&D intensity sectors
- Dummy 0= low R&D intensity sectors

Source: OECD, S&T Indicators

In all models an additional control is given by the dependent variable in absolute value at the initial year (2011).

3.4 Treatment of spatial effects

It is a standard assumption in regional economics that spillover effects decay with the distance from the source. In our case, we want to understand how the relative position of firms (closer or farther) in relation to universities has an effect on their performance.

We are in a position to estimate this effect by estimating the coefficients of distance (at various distance points) from firms to its source of knowledge, i.e. HEIs.
The procedure is as follows

(a) For each firm in the ORBIS dataset we take the georeferentiation coordinates; we do the same for all HEIs (based on the location of the headquarters) in the ETER dataset.

(b) We calculate the distance between each firm and each of the HEIs in the given country.

(c) We identify all HEIs that are at a fixed distance from the HEI and we sum their variables (e.g. number of students and graduates, number of publications and citations).

(d) We fix the distance as follows

- Less than 10 km
- Between 10 and 20 km
- Between 20 and 50 km
- Between 50 and 100 km

These areas are mutually exclusive (i.e. given a firm in a geographic location, we sum all HEIs located within 10km and define a separate variable; then we sum all HEIs within 20 km excluding the former ones, so that we identify only those that are at a distance between 10 and 20km). In this way we have “rings” of various radiiuses, that describe carefully the availability of HEIs centred around all firms. We fix the distances exogenously (i.e. imposing the distance at 10, 20, 50 and 100 km, respectively). This is considered acceptable in the literature.

The maximum distance (100 km) is defined in accordance with the prevailing literature. We do not explore the effects beyond 100 km, leaving this issue to future research.

In practice, we assume as spatial reference point the location of the firm, as made available by Orbis (GSM coordinates) or calculated by our own computation by applying a geo-referentiation software to the data containing the address of the firm.

The location of the firm is defined as the address of the headquarters. This definition creates a well-known problem in the case of large, multiplant companies.

The location of university is defined as the address of the central administration. Here the problem is less serious, given that the location of departments and schools is typically geographically limited to the urban environment. However, a number of universities have a few branches located in other cities (usually within short distance and in the same region). The ETER data does not provide the location of branches. It is known, however, that future editions of the ETER dataset (available end 2019) will offer an estimate of the distribution of students across branches, by collecting non official data on university websites. These data might be exploited in future studies.

Using this spatial approach we identify a differentiated spatial scale environment for each firm. HEIs located within 10 km from the firms are co-located in the same urban environment, allowing intense social and professional interaction and frequent face-to-face meetings. The 10-20 km ring extends this interaction to the province level, or the peripheries of large cities. On the contrary, the 20-50 ring captures the labour market mobility, or the possibility for graduates to shuttle. Finally, the location of HEIs beyond 50 km and until 100 km from the location of the firm is likely to capture the effect of the broader regional environment.

In principle we should leave the distance endogenous, adopting a full scale spatial econometric approach. This will be left for future research, given the heavy computational burden due to the huge size of the dataset. For the time being we use an approximation, by estimating the impact for a discrete number of distances from the centroid of the area.
With respect to the choice to examine spatial differences only within countries, it might be removed in further analyses allowing for cross-border interaction.
4. Main results

4.1 Models of growth of firm performance and aggregate independent variables

In the first set of models we use the three main independent variables (number of graduates, citations, academic staff) in an aggregate way at the level of universities. This means that, for each firm in the dataset, we compute the total number of these variables for all universities that are located at the various distances from the address of the firm. Aggregated variables constitute indicators of the overall supply of education and research, with no consideration for the disciplinary specialization.

We include all control variables (omitted in the table) of the specification described in the previous section. The independent variables are in standardized form, while the dependent variables are in logarithm. The dependent variables describe in log the absolute variation of the performance variable between 2011 and 2015. They are of two types:

(e) Variation 2011-2015 (in log): one measure per firm (Panel A)
(f) Annual variation (in log): four measures per firm (Panel B).

In Panel A we include only firms for which data on all four performance variables are available for 2011 and 2015. This greatly reduces the size of the sample but ensure comparability across various performance variables. In Panel B we use the same sample of firms but we compute annual variations in the interval. Due to missing observations in the year 2012, 2013 or 2014 in some cases, the total number of observations is smaller than the one in Panel A multiplied by the number of annual variations, i.e. four.

By using total and annual variation we exploit the richness of the ORBIS dataset, while at the same time we keep the set of companies observed stable across the various descriptors of performance. The slight differences in the number of observations across independent variables is the result of a few missing observations in the ETER dataset.

For the sake of simplicity, in all tables we only report statistically significant results. All complete tables are available upon request. Table 5 summarizes the main results (Table 5a and 5b).

The $R^2$ of the various models in Panel A of Table 5 are in the range 14-21%: given that the performance of firms is a construct with many antecedents and explanatory variables, this result should be considered satisfactory.

The explanatory power drops to 5-9% in the models in which the dependent variable is defined in terms of annual variation.

Control variables show the following patterns (not shown in tables):

- Variables describing the size of the firm are almost always significant. This finding is confirmed in all models described below.
- Variables related to the legal form of firms are significant across all models with the exception of intangible assets.
- Economic sectors enter with significant coefficients in almost half of the sectors and in all models, while in the model of turnover almost all sectors show significant coefficients.
- Country dummies are almost invariably not significant.
- Regional dummies follow very different patterns, with several significant coefficients, usually in a minority of cases.
Table 5a. Regression results for models of impact of higher education on growth of firms. Aggregated independent variables. Total variation 2011-2015.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of graduates (standardized variable)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>within 10 km</td>
<td>.022 ***</td>
<td>.019 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>between 11 and 20 km</td>
<td>.023 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>between 21 and 50 km</td>
<td>.029 ***</td>
<td>.011 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>between 51 and 100 km</td>
<td>.015 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>N= 393,931</td>
<td>N= 393,931</td>
<td>N= 393,931</td>
<td>N= 393,931</td>
</tr>
<tr>
<td>R²</td>
<td>0.213</td>
<td>0.145</td>
<td>0.170</td>
<td>0.145</td>
</tr>
</tbody>
</table>

Number of citations (standardized variable)

| within 10 km | -.019 *** |                                   |                                     |                                         |
| between 11 and 20 km |                                  | .010 *                          |                                     |                                         |
| between 21 and 50 km | .010 * | .008 *                          |                                     |                                         |
| between 51 and 100 km | .015 *** | .002 *                          |                                     |                                         |
| Number of observations | N= 356,732 | N= 356,732                      | N= 356,732                          | N= 356,732                              |
| R² | 0.207 | 0.144                           | 0.172                               | 0.144                                   |

Number of academic staff (standardized variable)

| within 10 km | .014 ** | .010 **                        |                                     |                                         |
| between 11 and 20 km | .017 *** |                                  |                                     |                                         |
| between 21 and 50 km | .022 *** | .008 *                          |                                     |                                         |
| between 51 and 100 km | .013 ** | .002 *                          |                                     |                                         |
| Number of observations | N= 324,077 | N= 324,077                      | N= 324,077                          | N= 324,077                              |
| R² | 0.208 | 0.144                           | 0.184                               | 0.152                                   |

Note: Tables report only statistically significant coefficients. Complete data are available upon request. *p < 0.05, **p < 0.01, ***p < 0.001.

Table 5b. Regression results for models of impact of higher education on growth of firms. Aggregated independent variables. Annual variation 2011-2015

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of graduates (standardized variable)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>within 10 km</td>
<td>.009 ***</td>
<td>.005 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>between 11 and 20 km</td>
<td>.007 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>between 21 and 50 km</td>
<td>.009 ***</td>
<td>.002 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>between 51 and 100 km</td>
<td>.004 **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>N= 1,181,252</td>
<td>N= 1,181,252</td>
<td>N= 1,181,252</td>
<td>N= 1,181,252</td>
</tr>
<tr>
<td>R²</td>
<td>0.095</td>
<td>0.054</td>
<td>0.060</td>
<td>0.060</td>
</tr>
</tbody>
</table>

Number of citations (standardized variable)

| within 10 km | -.006 *** | .004 *                          | .008 *                            |                                         |
| between 11 and 20 km |                                  | -.002 **                        |                                  |                                         |
| between 21 and 50 km | .003 * | -.002 *                         |                                  |                                         |
| between 51 and 100 km | -.003 *** | -.002 **                       | -.006 *                           |                                         |
| Number of observations | N= 1,060,162 | N= 1,060,162                    | N= 1,060,162                      | N= 1,060,162                              |
| R² | 0.093 | 0.054                           | 0.061                             | 0.058                                   |

Number of academic staff (standardized variable)

| within 10 km | .008 *** | .004 **                        |                                     |                                         |
| between 11 and 20 km | .006 *** |                                  |                                     |                                         |
| between 21 and 50 km | .008 *** | .002 *                          |                                     |                                         |
| between 51 and 100 km |                                  | -.002 *                         |                                     |                                         |
| Number of observations | N= 1,005,020 | N= 1,005,020                 | N= 1,005,020                      | N= 1,005,020                              |
| R² | 0.092 | 0.052                           | 0.065                             | 0.057                                   |

Note: Tables report only statistically significant coefficients. Complete data are available upon request. *p < 0.05, **p < 0.01, ***p < 0.001.
Coming to the coefficients, as a general finding, the supply of university-related inputs to the economy seems to have significant effects on the growth of companies.

There is a sharp difference between the dependent variables considered, however. Models of growth in turnover show positive coefficients for all three independent variables (graduates, citations, and academic staff), although the impact of graduates and staff is found across (almost) all distances, while for citations only in one case.

Models of growth in employment exhibit a different pattern: the coefficient is positive only for academic staff in the total variation model (panel A), while it is negative for citations in both models. In particular in the model of annual variation (panel B) citations enter negatively across all distances.

Models of growth in total assets show some positive coefficients across all independent variables (with the exception of citations in panel A), with only two negative coefficients in panel B for the longest distance.

Finally, there are only two variables with significant coefficient for intangible assets, one negative and one positive, in the case of citations in panel B.

Summing up, we find two models of growth of companies for which the spillover from spatially close HEIs generates a positive and significant effect: being located close to HEIs contributes to the growth of turnover and of total assets. Combining the two specifications (total variation + annual variation) we find significant coefficients for 17 variables in the model of turnover and for 11 variables in the model of total assets (negative in two cases), out of 24 potentially significant coefficients (3 independent variables * 4 distances * 2 specifications).

The other two models do not show consistent results across the independent variables. The model of growth in employment has only 6 significant coefficients, of which 5 for citations. The strong negative relation between the number of citations and the growth in employment would require further examination. On the other hand, the growth in intangible assets is left unexplained, with only 2 significant coefficients.

In terms of independent variables we also find a clear structure: the number of graduates and the number of academic staff enter positively and significantly in 25 cases, of which 24 have positive sign. More problematic is the case of citations, that enter significantly in 11 cases, of which 4 positive and 7 negative.

We therefore find a clear structure in the results: the aggregate supply of human capital (as measured by graduates) and of embodied research capabilities (as measured by academic staff), greatly benefits companies located close to HEIs in their competitiveness (as measured by growth of turnover) and structure (as measured by growth of total assets). Direct research flows (as measured by citations to publications) play a much weaker role, mainly for the growth in turnover.

In order to test the robustness of the specification, we run the regression models by taking as fixed variables the product of sector and region dummies (Table 6). This means that all firms operating in the same industry and the same region have the same categorical variable among their independent variables. The fixed variable specification largely confirms the findings from the initial specification. As in the case above, Table 6 is displayed in two panels, one for the total variation 2011-2015 (panel A), the other for the annual variation (panel B).

The inspection of Table 6 shows that all main findings in Table 5 are confirmed, in terms of sign and statistical significance. The single most important difference is that in the fixed variable specification we observe a negative impact of citations on the growth of employment (panel A) at most distances (3 cases).
Table 6a. Regression results for models of impact of higher education on growth of firms. Aggregated independent variables. Fixed variable specification. Total variation 2011-2015

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of graduates (standardized variable) within 10 km</td>
<td>.024 ***</td>
<td>-.010 **</td>
<td>.013 **</td>
<td></td>
</tr>
<tr>
<td>between 11 and 20 km</td>
<td>.023 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>between 21 and 50 km</td>
<td>.029 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>between 51 and 100 km</td>
<td>.018 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>N = 393,931</td>
<td>N = 393,931</td>
<td>N = 393,931</td>
<td>N = 393,931</td>
</tr>
<tr>
<td>R²</td>
<td>0.225</td>
<td>0.158</td>
<td>0.182</td>
<td>0.159</td>
</tr>
</tbody>
</table>

| Number of citations (standardized variable) within 10 km | | | | -.024 *** |
| between 11 and 20 km | | | | -.007 ** |
| between 21 and 50 km | .010 ** | | | |
| between 51 and 100 km | | | | -.006 * |
| Number of observations | N = 356,732 | N = 356,732 | N = 356,732 | N = 356,732 |
| R² | 0.219 | 0.156 | 0.184 | 0.158 |

| Number of academic staff (standardized variable) within 10 km | .015 ** | | | .008 * |
| between 11 and 20 km | .017 *** | | | |
| between 21 and 50 km | .021 *** | | | |
| between 51 and 100 km | .015 *** | | | |
| Number of observations | N = 323,853 | N = 323,853 | N = 323,853 | N = 323,853 |
| R² | 0.220 | 0.155 | 0.196 | 0.164 |

Note: Tables report only statistically significant coefficients. Complete data are available upon request. *p < 0.05, **p < 0.01, ***p < 0.001.

Table 6b. Regression results for models of impact of higher education on growth of firms. Aggregated independent variables. Fixed variable specification. Annual variation 2011-2015

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of graduates (standardized variable) within 10 km</td>
<td>.009 ***</td>
<td>-.003 **</td>
<td>.004 **</td>
<td></td>
</tr>
<tr>
<td>between 11 and 20 km</td>
<td>.007 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>between 21 and 50 km</td>
<td>.008 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>between 51 and 100 km</td>
<td>.004 *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>N = 1,181,252</td>
<td>N = 1,181,252</td>
<td>N = 1,181,252</td>
<td>N = 1,181,252</td>
</tr>
<tr>
<td>R²</td>
<td>0.101</td>
<td>0.058</td>
<td>0.064</td>
<td>0.063</td>
</tr>
</tbody>
</table>

| Number of citations (standardized variable) within 10 km | .006 * | -.007 *** | .003 * | .012 * |
| between 11 and 20 km | -.003 *** | | | |
| between 21 and 50 km | -.003 ** | | | |
| between 51 and 100 km | -.003 *** | -.003 ** | | |
| Number of observations | N = 1,060,110 | N = 1,060,110 | N = 1,060,110 | N = 1,060,110 |
| R² | 0.098 | 0.057 | 0.065 | 0.062 |

| Number of academic staff (standardized variable) within 10 km | .008 *** | | | .004 * |
| between 11 and 20 km | .006 *** | | | |
| between 21 and 50 km | .007 *** | | | |
| between 51 and 100 km | | | | -.002 * |
| Number of observations | N = 1,004,969 | N = 1,004,969 | N = 1,004,969 | N = 1,004,969 |
| R² | 0.098 | 0.055 | 0.069 | 0.060 |

Note: Tables report only statistically significant coefficients. Complete data are available upon request. *p < 0.05, **p < 0.01, ***p < 0.001.
4.2 Models of growth of firm performance. (a) models with control variables

We now turn to models in which the independent variables are disaggregated by discipline, in terms of Fields of Education, for graduates, and Subject Categories of publications, for citations.

4.2.1 General results

As a general observation, it can be said that regression models do not necessarily have the goal of explaining the whole variability of firm performance, which is the result of a complex array of variables, among which the availability of human capital and research is not certainly the most important one. Having said that, in presenting the results we will stick to a cautious style of description or association, never of explanation or causal demonstration. There are simply too many omitted variables. Furthermore, while the dependent variables are taken from a large dataset that makes use of official information (balance sheets), there is still room for measurement errors in reporting, or in transferring balance sheet data in the ORBIS dataset. At the same time, all models are based on very large samples of firms.

4.2.2 Graduates

The first run of models has the number of graduates (ISCED 5-7) as the independent variable, measured in units. This variable captures the flow of students that complete their degrees in any given year. Covering all degrees from ISCED 5 (short course) to ISCED 6 (bachelor) and ISCED 7 (master) gives a full representation of the supply of tertiary education.

We examine the substantive results first by discipline, second by performance variable. Table 7 shows the findings, again with panel A and panel B. We present only significant coefficients, for the sake of simplification and clarity.

**Business**

Tertiary education in Business is positively associated to higher performance in terms of growth of turnover and employment but only at a close distance (within 10 km) or at a long distance (between 51 and 100 Km) only for turnover (panel A). In panel B graduates in Business are positively associated to growth in total assets at the shortest distance, but negatively to growth in intangible assets (for the two shortest distances).

This finding is interesting, as it supports the notion that firms hire graduates with a managerial background in order to address growth processes. In addition, it is possible that recruitment of graduates educated in business trickles down to graduates in other disciplines.

**Social Sciences and Humanities**

The supply of graduates in all SSH disciplines (with the exception of Business) is associated to performance in a very small number of cases, only three in total. Of these, two are positive (total assets at distances within 10 and 11-20 km, respectively, in panel A), one negative (employment within 10 km, same panel). We conclude that the impact is negligible.

It is not easy to interpret the findings. It might be said that graduates in SSH do not target employment in firms, but rather in public administration services (education, research, public service) or in the professional field (lawyers, psychologists, consultants). Therefore the findings do not point to an argument of low impact of education in SSH in society, but rather show that the absorption of graduates from SSH in companies (with the exception of Business) is still very limited. These findings would open at least two lines of further inquiry:
— measuring the propensity of firms to hire graduates in SSH and/or examining the degree to which SSH education provides students with the skills required for an employment in companies;

— measuring the growth in the entrepreneurial sector based on SSH, or the Creative industries in Europe, in order to identify gaps in skills and in formal tertiary education.

On the latter point, it might be that the lack of impact of graduates in SSH on the business sector is due to the distinctive lack of skills in management during the university training. Integrating the education in SSH with elements of practical application might create an incentive for firms to hire these graduates. For the time being, this is a conjecture that we cannot test here.

Natural Sciences

In this area we find a positive impact on turnover and employment at closer distances (within 10 and 11-20 km, respectively, in panel B). A positive impact is confirmed for employment only within 10 km, in panel A.

Table 7a. Regression results for models of impact of higher education on growth of firms.
Total variation 2011-2015

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of graduates within 10 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td>-.029 *</td>
<td>.046 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td>.066 ***</td>
<td>.038 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td>.028 **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td>-.056 ***</td>
<td>-.029 ***</td>
<td>-.029 **</td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td>-.023 *</td>
<td>.089 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of graduates between 11 and 20 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td>.023 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td>-.033 **</td>
<td>-.018 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td>.078 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of graduates between 21 and 50 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td>.060 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of graduates between 51 and 100 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td>.060 *</td>
</tr>
<tr>
<td>Number of observations</td>
<td>N= 393,931</td>
<td>N= 393,931</td>
<td>N= 393,931</td>
<td>N= 393,931</td>
</tr>
<tr>
<td>R²</td>
<td>0.215</td>
<td>0.156</td>
<td>0.180</td>
<td>0.158</td>
</tr>
</tbody>
</table>

Note: Tables report only statistically significant coefficients. Complete data are available upon request. *p < 0.05, **p < 0.01, ***p < 0.001.

Medicine

The supply of graduates in Medicine is almost everywhere associated negatively to firm performance. This is found for the growth in turnover and in employment at distances within 10 and 11-20 km, respectively in panel A, and at distance within 10 km in panel B, where a positive coefficient is found for turnover at the highest distance.
The interpretation of results may be obtained by reflecting on the fact that medical doctors are not hired by companies, nor (in general) create their own companies. They are hired by the Human health sector, either public (public hospitals) or private (private hospitals). In general, the recruitment of medical doctors follows highly regulated long term plans, which take into account both the supply side (given that the training of medical doctors takes 10 years on average) and the demand side (driven by demographic and epidemiological factors). Medical doctors graduate in a university city with a Medical school and migrate to hospitals where their work is required. It might be possible that the supply of medical doctors generates more employment at local level, either in the same sector (i.e. university hospitals hiring more doctors and attracting patients from other regions) or in closely related sectors (i.e. private hospitals benefiting from the prestige of local university in order to establish a business and attract patients). These effects might be estimated by a dedicated analysis on the NACE category Q-Human health, for which ORBIS offer data. Consequently, there is no reason to expect a relation between the supply of medical doctors in a region or city and the performance of all firms, as in the general models.

**Table 7b. Regression results for models of impact of higher education on growth of firms. Annual variation 2011-2015.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of graduates within 10 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td>.024 ***</td>
<td>.012 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td>-.013 **</td>
<td>-.009 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td>-.012 ***</td>
<td></td>
</tr>
<tr>
<td><strong>Number of graduates between 11 and 20 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td>-.029 **</td>
</tr>
<tr>
<td>Natural sciences</td>
<td>.017 ***</td>
<td>.007 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td>-.008 **</td>
</tr>
<tr>
<td><strong>Number of graduates between 21 and 50 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td>-.008 **</td>
</tr>
<tr>
<td><strong>Number of graduates between 51 and 100 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td>-.014 **</td>
</tr>
<tr>
<td>Number of observations</td>
<td>N= 1,181,199</td>
<td>N=1,181,199</td>
<td>N= 1,181,199</td>
<td>N= 1,181,199</td>
</tr>
<tr>
<td>R²</td>
<td>0.101</td>
<td>0.058</td>
<td>0.064</td>
<td>0.063</td>
</tr>
</tbody>
</table>

Note: Tables report only statistically significant coefficients. Complete data are available upon request. *p < 0.05, **p < 0.01, ***p < 0.001.

**Technology**

In the case of graduates from Engineering schools we find a surprising dichotomy between models of performance in total assets and in intangible assets. For the latter we find positive coefficients at all but the longest distance in panel A, while for total assets...
we find negative coefficients across all distances in panel B and at short and long distance in panel A.

On the contrary, there is no impact whatsoever on the growth of employment and only a negative impact on turnover at the largest distance in both panels.

This result might be interpreted in terms of the different determinants of total assets and intangible assets. The former follow more closely the short term growth in turnover, while the latter depend on long term investment decisions.

In addition, the negative coefficients for total assets are mostly found in panel B, in which the dependent variable is the annual variation, with more volatility than the total 2011-2015 variation across firms. However, it is not easy to understand why the impact of graduates in Technology on total assets is negative.

Dimensions of firm performance

Reading the results by column -that is, by type of firm performance- sheds some light on the underlying dynamics. The growth of turnover is associated positively only to the supply of graduates in Business in panel A and of graduates in Natural sciences in panel B, while the significant coefficients of graduates in Medicine are mostly negative. As for technology, it is related negatively to turnover in both panels but only at the highest distance.

As for employment, we have some evidence of a positive correlation with the number of graduates in Business and Natural Sciences, but only at low distance, and some negative correlation with the number of graduates in Social Sciences and Humanities and Technology (panel A), and Medicine (panel B), mostly for the short distance.

The growth of total assets in panel A depends negatively on graduates in Medicine (at the lowest distance) and Technology (mostly in panel B), and positively on graduates in SSH (panel A) and in Business (panel B), in both cases at short distance.

The growth in intangible assets depends positively on graduates in Technology in panel A and negatively on graduates in Business in panel B.

An interpretation pointing to the difference between performance in turnover (which is more short term and dependent on external variables) and performance in intangibles (which is more long term and depends more strongly on internal or strategic variables) is a candidate for further exploration.

4.2.3 Citations

The total number of citations received by publications produced by HEIs in the 2011-2015 period, and measured cumulatively until 2018 (Scopus data downloaded in July 2018), try to capture the volume and quality of research produced by institutions.

We examine the results first by discipline, second by performance variable (Table 8), following the specifications in panel A and panel B.

The first striking result is that the total number of significant coefficients is much lower than in the case of graduates. We find 20 statistically significant coefficients in Table 8, against 33 in Table 7. No significant coefficients are found for the model of intangible assets, only two for turnover. This is a preliminary warning on the danger of overestimating the direct impact of scientific research on the economy, or the need to build up a more complex analysis of impact pathways. Let us now turn to disciplinary effects.

Social Sciences and Humanities

Citations to publications in SSH are significantly associated to variables describing performance only in 3 cases out of possible 32 cases (4 distances * 4 performance variables* 2 specifications for panel A and panel B). While these coefficients are all positive, they are found only at the largest distance (51-100 km). This finding reinforces
the one already mentioned, pointing to the need to consider SSH in a broader context of utilization of skills.

The fact that research in SSH has a negligible impact on firm performance does not mean that this is a sector of research that is not useful for society and the economy. It means that a diverse, perhaps more sophisticated and articulated, model of impact, must be adopted to make justice for these fields.

**Basic Sciences**

We find only two significant coefficients, all with negative sign, in the model of growth in total assets, and a negative sign for employment. We believe there is nothing that can be generalized on the role of Basic Sciences.

**Table 8a. Regression results for models of impact of scientific research on growth of firms. Total variation 2011-2015**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of citations within 10 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td>-0.077 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of citations between 11 and 20 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td>-0.026 *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td>-0.038 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td>0.022 *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of citations between 21 and 50 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td>-0.036 **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td>0.024 *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of citations between 51 and 100 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td>0.022 *</td>
<td>0.040 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td>-0.041 ***</td>
<td>-0.026 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td>-0.055 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td>0.051 ***</td>
<td>0.021 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>356,479</td>
<td>356,479</td>
<td>356,479</td>
<td>356,479</td>
</tr>
<tr>
<td>R²</td>
<td>0.218</td>
<td>0.156</td>
<td>0.183</td>
<td>0.157</td>
</tr>
</tbody>
</table>

**Note:** Tables report only statistically significant coefficients. Complete data are available upon request. *p < 0.05, **p < 0.01, ***p < 0.001.

**Natural Sciences**

Here we find a consistent pattern of negative coefficients, although all of them are concentrated in the model of growth of employment (4 in panel A, 3 in panel B).

**Medicine**

Medicine has only two significant coefficients, both positive (in panel B). Although the number is low, here we find an interesting difference with respect to the impact of graduates. While graduates have a negative impact on the economic environment surrounding the Medical schools, the volume and quality of research does have a positive impact.
It is possible to interpret these findings as a window on the knowledge spillovers generated by life science and clinical research at university level.

**Technology**

In the case of Technology we find a consistent pattern for panel A. The number of citations is associated to the growth of total assets across all but the closest distances. It is also associated positively to the growth of employment but only in the 51-100 km distance. This pattern is remarkably different from the one for graduates.

**Table 8b. Regression results for models of impact of scientific research on growth of firms. Annual variation 2011-2015**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of citations within 10 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td>.012 ***</td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of citations between 11 and 20 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td>.006 **</td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of citations between 21 and 50 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td>.009 *</td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of citations between 51 and 100 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td>.012 ***</td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td>.010 **</td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td>.014 ***</td>
</tr>
<tr>
<td><strong>Number of observations</strong></td>
<td>1,060,110</td>
<td>1,060,110</td>
<td>1,060,110</td>
<td>1,060,110</td>
</tr>
<tr>
<td>R²</td>
<td>0.098</td>
<td>0.057</td>
<td>0.065</td>
<td>0.062</td>
</tr>
</tbody>
</table>

**Note:** Tables report only statistically significant coefficients. Complete data are available upon request. *p < 0.05, **p < 0.01, ***p < 0.001.

**Dimensions of firm performance**

We try to make sense of the various results by reading the findings by type of performance examined.

As already mentioned, we have two models completely unexplained, i.e. turnover and intangible assets. It is difficult to understand how the selectivity of knowledge spillovers works in these two cases.

On the contrary, the growth of total assets is associated positively to the supply of citations in Technology (only in panel A), and negatively to the citations in Basic sciences (only in panel A).

The growth of employment, in turn, is associated negatively to the citations in Natural and Basic sciences, and positively (but only weakly) to citations in SSH at the highest distance.
Summing up, it is shown that knowledge spillover effects from scientific research are highly selective. Knowledge bases, as captured by broad scientific fields, are highly differentiated in terms of the size of the impact and the distance at which they arrive.

Overall, we find reasons to confirm the importance of scientific research for the performance of firms, although the size of the impact is clearly small with respect to other economic and financial factors. Furthermore, it seems that the pervasiveness of the impact of direct knowledge spillovers, as measured by the number of significant coefficients, is lower than the impact of skilled human capital via the supply of graduates.

4.2.4 Joint model of graduates and citations

We examine here a family of integrated models in which the independent variables include both education (number of graduates) and research (number of citations). The remaining structure of the models is left unchanged. These models try to capture the joint effect of the two main activities of HEIs. The fact that both graduates and citations enter the equations with absolute numbers is aimed at capturing the volume effect, as well as, in the case of research, the quality of output (Table 9, panel A and B).

In the analysis of findings we will focus mainly on panel A, given that in panel B the significant coefficients are very few (only 9 out of 160) and scattered across several independent variables and performance models.

Business

We find confirmation of the effects found in the simple model: the number of graduates in Business contributes positively to the performance of firms in terms of turnover and employment, at short distance (within 10 km) and long distance (between 51 and 100 km). In the case of Business we do not have disaggregated data for citations, as they are aggregated into Social Sciences and Humanities.

Social Sciences and Humanities

The joint effects of graduates and citations shows different results:

— the number of graduates enters with positive coefficients in all distances below 50 km for the model of total assets, and in the model of turnover for the distances 11-20 and 21-50 km, respectively, for a total number of statistically significant relations equal to five;
— the number of citations enters with negative coefficient in one case and with positive coefficient in another case, making the results inconclusive.

Thus in a joint model the role of SSH is predominantly positive, although the impact is channelled only through the creation of human capital. This is in contrast with the findings from the model with only graduates, in which there are no significant coefficients.

Natural Sciences

Graduates in Natural Sciences have only negative impact on the growth of turnover at large distance and no impact elsewhere. In turn, citations to publications in Natural Sciences have a negative impact on the growth of employment across all distances.

Basic Sciences

Citations to publications in Basic sciences have negative associations with the growth in total assets in three cases, and no effect elsewhere.

Medicine

We find very few significant coefficients (and only for graduates), which are negative on the growth of turnover and intangible assets.

Technology
The joint model confirms the role of human capital and knowledge spillovers from Technology: in panel A we find as many as 16 statistically significant coefficients, of which 12 with positive sign. They are positive across all distances for the impact of graduates on the growth of intangible assets, and positive across three or four distances for the impact of citations on the growth of total assets and employment, respectively. Overall, these findings point to a large impact of creation of graduates and of research activities in this field across many dimensions of performance (with the exception of turnover).

Table 9a. Regression results for models of joint impact of higher education and scientific research on growth of firms. Total variation 2011-2015

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of graduates (ISCED 5-7) within 10 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td>.077 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td>.088 ***</td>
<td>.058 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td>-.056 **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td>-.039 **</td>
<td></td>
<td>.107 *</td>
</tr>
<tr>
<td><strong>Number of graduates (ISCED 5-7) between 11 and 20 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td>.060 **</td>
<td></td>
<td></td>
<td>.045 **</td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td>-.029 *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td>.092 **</td>
</tr>
<tr>
<td><strong>Number of graduates (ISCED 5-7) between 21 and 50 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td>.076 ***</td>
<td></td>
<td></td>
<td>.049 **</td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td>-.116 *</td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td>.107 **</td>
</tr>
<tr>
<td><strong>Number of graduates (ISCED 5-7) between 51 and 100 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td>-.044 *</td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td>-.033 *</td>
<td>-.029 *</td>
</tr>
<tr>
<td><strong>Number of citations within 10 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td>-.073 **</td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td>.047 *</td>
<td></td>
</tr>
<tr>
<td><strong>Number of citations between 11 and 20 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of citations between 21 and 50 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of citations between 51 and 100 km</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Number of graduates (ISCED 5-7) within 10 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td>.019 **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td>-.014 *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of graduates (ISCED 5-7) between 11 and 20 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td>.013 **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of graduates (ISCED 5-7) between 21 and 50 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of graduates (ISCED 5-7) between 51 and 100 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Business</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of citations within 10 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td>-.019 *</td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td>.038 **</td>
</tr>
<tr>
<td>Number of citations between 11 and 20 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td>.006 *</td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of citations between 21 and 50 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social sciences and Humanities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of citations between 51 and 100 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Dimensions of performance

The performance of firms in terms of turnover is associated positively with the number of graduates in Business (two positive coefficients) and in SSH (two coefficients) and negatively with the number of graduates in Medicine (two coefficients). Interestingly, there is no impact of the number of citations across all distances and disciplines, with the only exception of Technology between 51 and 100 km.

The growth in employment is mainly positively associated to human capital in Business and research activities in Technology, while research activities in Natural and Basic Sciences have a negative impact, as does the number of graduates in Technology.

The growth of total assets and intangible assets is positively associated to the number of graduates in SSH and Technology, respectively, and to citations in Technology in the former case. However, growth in total assets is negatively associated to the number of graduates (but only to the highest distance) and to citations in Basic Sciences (in three cases out of four) and in SSH (but only at the distance of 11-20KM).
5. Summary of results

Let us try to build up a consistent summary of results. Our main goal is to identify whether the positive impact of the creation of skilled human capital and of knowledge spillovers that is found in macroeconomic models (at cross-country and cross-regional level) or in microeconomic models based on small scale surveys (usually just in one country or region) can be replicated and validated on a microlevel, with very large samples, across many European countries.

This is an exploratory exercise, aimed at integrating heterogeneous datasets and testing the potential of the available data sources for addressing this issue. The existing data sources at microlevel have a number of limitations that make it impossible to address all methodological problems simultaneously. What they can offer is large scale samples and cross-country comparability. The data constraint, however, apply here in a strong way.

Our strategy cannot realistically rely on a sound identification of a causal model. There are simply too many omitted variables. There are also likely measurement errors.

We turn, on the contrary, to an exploratory strategy in which we do not ask models to have a large \( R^2 \) and we do not claim causality relations. More modestly, we look for patterns of covariation that are sufficiently robust to be explained according to our reading of the literature. The expectation is that, given the limitations of the data, very few statistically significant coefficients can be found.

Therefore we explore a large number of models, testing different specifications for both independent and dependent variables and adding as many control variables as possible.

We believe we have found some of these patterns.

First, the covariation of firms’ performance and activities carried out by higher education institutions on firms is confirmed beyond any reasonable doubt. After controlling for many structural factors (industry, size of the firm, legal form) and location factors (country, regions) it appears that HEIs produce outputs, in terms of human capital (graduates) and knowledge (publications), that in the large majority of cases positively influence firms located in the neighbourhood. This effect is even more remarkable since it comes from flows (yearly production of graduates and publications), that is, from relatively short term phenomena, not from established, slowly created stocks.

This result holds for aggregated models in which independent variables are the total number of graduates and academic staff, at least for models of growth in turnover and total assets. The impact of citations is also positive, but is found across a lower number of distances. This result also holds for models in which independent variables are disaggregated by discipline, with the qualifications that follow.

Second, the impact is strongly dependent on specific disciplines, or bases of knowledge. This is another important result, which confirms the notion that knowledge is heterogeneous and follows different pathways to become productive in the economic system. At the same time, this finding sheds light on the necessity to consider a variety of impact models. Assuming an impact model for which the only dependent variables to be considered are those that refer to firms (whatever the dimension of performance) is a serious conceptual (and political) mistake.

Third, human capital creation and knowledge spillovers in the field of Engineering or Technology have the largest impact. The number of citations to publications is positively associated to several dimensions of firm performance (in particular, total assets), while the number of graduates is positively associated to the growth in intangible assets. In a model of joint effect between graduates and citations, Technology is the discipline with the largest number of positive associations.
Fourth, higher education in managerial disciplines (Business) is positively associated to several dimensions of performance, in particular the growth in turnover and employment. It seems that firms benefit from the education of graduates in disciplines that offer systematic background (as opposed to the intuitive, on-the-job practical knowledge) to the growth of firms. We cannot confirm a similar impact for scientific output, since the classification used for publications collapses Business and management into SSH. It should be the object of future exploration.

Fifth, basic disciplines such as Basic Sciences, Natural Sciences and, to a less extent, Medicine, have both negative and positive associations, but not robust enough to establish a general foundation for the impact. The impact is more positive for graduates than for citations, in general. It might be that their effect is more long term, hence more difficult to observe in the short term. It might also be that some of these disciplines create the bases for others (e.g., Mathematics and Physics for Engineering). Finally, it is also clear that graduates in these disciplines are less likely to find employment in companies, but more likely work in professional services or the public administration (education, health).

Sixth, we find absence of significant relations between the number of graduates and/or the number of citations to publications in Social Sciences and Humanities and the performance of firms, when they are measured in separate models. Yet a positive association is found in a joint model: the number of graduates positively impact on the growth of turnover and of total assets, while citations still do not show any impact.

These mixed results might be explained by recalling that graduates in SSH find employment mainly in professional positions outside the business sector, such as public administration. Another possible explanation for the lack of impact is that large numbers of graduates in SSH are found in less developed regions, in which employment opportunities from the private sector are more scarce and the dynamism of firms is lower. If this is the case, peripheral regions will have an excess supply of graduates in SSH with respect to the absorptive capacity of the local economic system. Another possible interpretation is that the lack of economic incentives might lead students to choose less difficult career, or less oriented towards the private labour market and more targeted to find jobs in public administration. Yet when higher education is associated to a larger number of citations in the international literature it shows a positive impact.

At a theoretical level we find indirect confirmation of macrolevel theories of the role of knowledge in advanced societies as an engine of growth, as well as of micro-based theories on the relation between human capital, knowledge spillovers, innovation and firm performance.

Several caveats and limitations should be considered.

First, the impact of human capital and research on productive activities is a long term phenomenon, not a short term one. We know from the literature that the time needed for public investment into higher education and research to generate an observable impact on economic growth is in the order of several decades. In principle, this means that the independent variables should be expressed in stocks, rather than in flows. It is the stock of skilled workers (either produced recently as new graduates, or already employed) that contributes to improving the productivity of firms. We exploit a recently created dataset (ETER), in which the time series is too short to build a significant measure of stock of graduates. It will be possible to improve this measure in the next few years, covering at least one decade. In a similar way, it is the stock of scientific publications accumulated over time by researchers that generates knowledge spillovers to companies.

Second, the impact of human capital and research on companies is the result of a complex dynamic, as it is witnessed by the literature on the relation between innovation and productivity. In principle, research spillovers may generate new ideas for product innovation, while process innovation would typically benefit more from intra-industry
relations, such as those between customers and suppliers. In turn, product innovation may result in increased turnover.

Yet the most significant impact of innovation is on productivity, which is the result of product innovation (new products have typically higher prices), process innovation (cost reducing manufacturing technologies), and organizational innovation (complementarity between technology and human capital). We do not have data available to calculate labour productivity from balance sheets (data on number of hours or days are not available), even less to estimate total factor productivity. Therefore, while this dynamic is multi-period and is subject to many intervening and mediating factors, we only have available final indicators of performance, such as turnover and intangible assets, or structural indicators such as total assets and employment.

Third, the impact of human capital on firms depends on decisions in human resources management that we cannot observe. On the one hand, the relation between innovation and productivity, and, on the other one, the recruitment process is not deterministic. The only robust evidence we may find from the literature is an average, long term, positive relationship between the educational qualification of workforce and productivity. But this long term average relation may take place following a variety of short term flows (hiring, firing, substituting personnel with automation, training existing workforce and the like), which we are not in a position to observe directly. On the other hand, hiring decisions of firms do not necessarily target workers with tertiary education and, within the pool of potential candidates of workers with tertiary education, do not necessarily target graduates in local universities. Here we miss substantive knowledge about the details of the hiring processes of the firms in the sample, although the literature suggests that a large share of workers recruited by companies graduated from universities in the vicinity.

This leads us, fourth, to another limitation of our data. Not only we do not know in detail the proportion of workers hired by companies that graduated from universities in the neighbourhood, but we also ignore the proportion of graduates of universities that migrate elsewhere. Again we find support from the studies discussed in the survey of the literature, that show large proportion of graduates who migrate from peripheral to central regions and cities, but we do not have any detail about the distribution of skilled migration at microlevel.
References


Breznitz, S.M., Feldman M.P. (2012), The engaged university. *Journal of Technology Transfer*, vol. 37,


Giuliani E., Morrison A., Pietrobelli C., Rabello R. (2010) Who are the researchers that are collaborating with industry? An analysis of the wine sectors in Chile, South Africa and Italy. Research Policy, 36(6), 748-761.


Huggins, R., Kitagawa, F. (2012). Regional policy and university knowledge transfer: Perspectives from devolved regions in the UK. Regional Studies, 46(6), 817–832.


Mariani, M. (2002), Next to Production or to Technological Clusters? The Economics and Management of R&D Location, *Journal of Management and Governance* 6, 131-152


OECD (varius years) *Education at a glance*. OECD, Paris.


Sloane P.J. (2014) Overeducation, skill mismatches, and labor market outcomes for college graduates. IZA Working paper no. 88


Sutherland, W. J., Goulson, D., Potts, S. G., & Dicks, L. V. (2011). Quantifying the impact and relevance of scientific research. [http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0027537](http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0027537)


Several estimates have been produced on the impact of various measures of production of knowledge on economic indicators such as natality of new firms, firm growth, productivity growth, or growth in GDP per capita. While not all of them deal explicitly with university-based research, it is important to review this literature in order to establish the framework for our exercise.

Among the measures of knowledge, a distinction is typically drawn between flow measures (for example, the expenditure in R&D, or a measure of input, or the number of publications, citations, patents, as measures of output) and stock measures (or measures of the accumulation of knowledge, such as cumulated R&D investment associated with a rate of depreciation that takes into account the obsolescence of knowledge). Regression models have been used, in particular, to examine the relation between R&D and productivity growth (for surveys see Mohnen & Hall, 2013; Raymond et al. 2013). Within this tradition, a number of authors have examined the impact of public R&D investment (Adams 1990; Bassanini et al. 2001; Guellec & Van Pottelsbergh, 2004; Coe et al. 2009) on productivity. Kantor and Whalley (2014) is a recent re-statement of this literature.

Underlying this representation there is an input-output model, originated in macroeconomics and in the production function tradition, associated to a spillover model, originated in the formulation of Arrow’s under-investment thesis (Arrow, 1962). The input-output model postulates that production factors (such as capital and labor) have always positive marginal products, although with diminishing returns. Knowledge is considered an additional factor of production, one that however does not contribute directly to production, but has an impact on Total Factor Productivity (TFP), whose definition approximates Solow’s residual. Investment in knowledge has an impact on the overall productivity of other factors of production. The empirical question is “how much” impact does knowledge have on the productivity of factors of production, and hence on economic growth.

Associated to this representation there is the notion that knowledge shares some of the essential features of public goods, such as non-rivalry and non-excludability. Investing in the production of knowledge, therefore, is subject to the risk of generating spillovers that give benefits to competitors, or anyway to agents who have not paid for the knowledge itself. According to this line of thinking, it is interesting to better understand the way in which knowledge produces external effects, generating economic opportunities beyond the domain of those that have invested in its production.

A large literature has investigated the issue of the impact of private R&D on firms that incur in the cost of the investment (private rate of return). Wieser (2005), Hall et al. (2010) and Mœn and Thorsen (2015) carry out surveys of existing studies. The range of rates of return reported in these surveys is as follows: from 16% to 28% in Wieser (2005), from 20% to 30% in Hall et al. (2010), while Mœn and Thorsen (2015) report a median value between 13% and 18%. Umur et al. (2016) is the most updated and comprehensive survey, being based on 1253 estimates from 65 studies that have adopted the production function approach. Their estimate for the rate of return is 14%, a lower level than previous surveys. Among the studies surveyed by Umur et al. (2016), five estimate the impact of public expenditure in business R&D, in the form of direct subsidies or fiscal credits (Bartelsman, 1990; Lichtenberg and Siegel, 1991; Mansfield,1980; Terleckyj, 1980; Wolf and Nadiri, 1993). These studies report a lower rate of return for publicly supported research. The authors of the survey offer several possible explanations (concentration of public subsidies in industries with higher spillovers (e.g. health), or in industries with high level of front-up costs (e.g. aerospace), or inefficiency in the use of public resources.
Interestingly, Umur et al. (2016) note that there is a need for further studies in the two areas mentioned above: the estimation of spillover effects (i.e. public return of private R&D expenditure), and the impact of public support to business R&D (i.e. private and public return of R&D and innovation public support programs). A general remark is that most studies estimate the impact of R&D in a simultaneous way, while it would be more appropriate to consider lagged effects in time. With respect to the rate of return of public research, we can conclude that there is a need to investigate it over several decades using a lagged effect methodology. A robust conclusion from the existing stock of empirical knowledge is that the R&D investment is, on average, quite profitable for private companies.

Another approach to the issue of rate of return of R&D is found in the economics of growth literature. Here the unit of analysis is the country. A related literature is found in regional economics, taking into account the region as unit of analysis. At both levels the main question is whether we see a causal impact of the R&D activity on the level, or rate of growth, of GDP per capita. This relation is modelled by using a Knowledge Production Function at the level of the country/region.

With respect to the operationalization of R&D, a common strategy is to use the level of expenditure, either public (GOVERD and HERD in the OECD nomenclature) or private (BERD). Another strategy is to use the stock of scientific publications or citations at country level, a variable that can be derived from commercial bibliometric sources.

The latter approach was pioneered by de Solla Price (1978) and followed by Hart and Sommerfeld (1998), Haiqi and Yiha (1997) and Pouris (2003) using citations at country level. Data on publications were instead used by De Moya-Anegón and Herrero-Solana (1999), Pouris and Pouris (2009) and Inglesi-Lotz and Pouris (2013). All these studies report significant correlation between the number of citations or publications and the level of GDP. Vinkler (2008), on the contrary, did not find correlation. One of the limits of these studies is in the lack or the under-specification of a formal model and of econometric techniques that might allow the investigation of causality relations.
APPENDIX 2

Main classifications in the ETER Handbook

Table A.1. Fields of education

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
<th>Subfields</th>
<th>ISCED 1997 FOE</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>General programmes and qualifications</td>
<td>001 Basic programmes and qualifications 002 Literacy and numeracy 003 Personal skills</td>
<td>01 Basic programmes 08 Literacy and numeracy 09 Personal development</td>
</tr>
<tr>
<td>01</td>
<td>Education</td>
<td>011 Education</td>
<td>14 Teacher training and education science</td>
</tr>
<tr>
<td>02</td>
<td>Humanities and Arts</td>
<td>021 Arts 022 Humanities 023 Languages</td>
<td>21 Arts 22 Humanities</td>
</tr>
<tr>
<td>03</td>
<td>Social sciences</td>
<td>031 Social and behavioral science 032 Journalism and information</td>
<td>31 Social and behavioral science 32 Journalism and information</td>
</tr>
<tr>
<td>04</td>
<td>Business and law</td>
<td>041 Business and administration 042 Law</td>
<td>34 Business and administration 38 Law</td>
</tr>
<tr>
<td>05</td>
<td>Natural Science, mathematics and statistics</td>
<td>051 Biological and related sciences 052 Environment 053 Physical sciences 054 Mathematics and statistics</td>
<td>42 Life sciences Part of 62 (natural parks and wildlife) 44 Physical sciences 46 Mathematics and statistics</td>
</tr>
<tr>
<td>06</td>
<td>Information and communication technologies</td>
<td>061 Information &amp; Communication Technologies</td>
<td>48 Computing</td>
</tr>
<tr>
<td>07</td>
<td>Engineering, manufacturing and construction</td>
<td>071 Engineering and engineering trades 072 Manufacturing and processing 073 Architecture and construction</td>
<td>52 Engineering and engineering trades (plus most of 85 environmental protection) 54 Manufacturing and processing 56 Architecture and building</td>
</tr>
<tr>
<td>08</td>
<td>Agriculture, forestry, fisheries and veterinary Health and welfare</td>
<td>081 Agriculture 082 Forestry 083 Fisheries 084 Veterinary 091 Health 092 Welfare</td>
<td>62 Agriculture, forestry and fishery (minus natural parks and wildlife) 64 Veterinary 72 Health 76 Social services</td>
</tr>
<tr>
<td>09</td>
<td>Services</td>
<td>101 Personal services 102 Safety services 103 Security services 104 Transport services</td>
<td>81 Personal services Part of 85 environmental protection (community sanitation and labor protection and security)</td>
</tr>
</tbody>
</table>

References:
http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-education.aspx
Table A.2. Correspondence table Fields of Education- Fields of Science

<table>
<thead>
<tr>
<th>ISCED-F 2013</th>
<th>Fields of Science FOS - 2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>00 General programmes and qualifications</td>
<td>-</td>
</tr>
<tr>
<td>01 Education</td>
<td>5.3 Educational sciences</td>
</tr>
<tr>
<td>02 Humanities and Arts</td>
<td>6. Humanities</td>
</tr>
<tr>
<td>03 Social sciences</td>
<td>5. Social sciences without 5.2, 5.3 and 5.5</td>
</tr>
<tr>
<td>04 Business and law</td>
<td>5.2 Economics and Business</td>
</tr>
<tr>
<td></td>
<td>5.5 Law</td>
</tr>
<tr>
<td>05 Natural Science, mathematics and statistics</td>
<td>1. Natural sciences without 1.2</td>
</tr>
<tr>
<td>06 Information and communication technologies</td>
<td>1.2 Computer and information sciences</td>
</tr>
<tr>
<td>07 Engineering, manufacturing and construction</td>
<td>2. Engineering and technology</td>
</tr>
<tr>
<td>08 Agriculture, forestry, fisheries and veterinary</td>
<td>4. Agricultural sciences</td>
</tr>
<tr>
<td>09 Health and welfare</td>
<td>3. Medical sciences</td>
</tr>
<tr>
<td>10 Services</td>
<td>-</td>
</tr>
</tbody>
</table>

* Includes urban planning, which is in the FOS classification included in 5. Social sciences.
### Table A.3. Levels of education

<table>
<thead>
<tr>
<th>ISCED-2011 level</th>
<th>Definition</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISCED 5 short-cycle tertiary education</td>
<td>Programmes at ISCED level 5, or short-cycle tertiary education, are often designed to provide participants with professional knowledge, skills and competencies. Typically, they are practically based, occupationally specific and prepare students to enter the labour market. However, these programmes may also provide a pathway to other tertiary education programmes. Academic tertiary education programmes below the level of a Bachelor’s programme or equivalent are also classified as ISCED level 5.</td>
<td>Duration: 2-3 years. Entry requirements: ISCED 3 or 4</td>
</tr>
<tr>
<td>ISCED 6 Bachelor’s or equivalent levels</td>
<td>Programmes at ISCED level 6, or Bachelor’s or equivalent level, are often designed to provide participants with intermediate academic and/or professional knowledge, skills and competencies, leading to a first degree or equivalent qualification. Programmes at this level are typically theoretically based but may include practical components and are informed by state of the art research and/or best professional practice. They are traditionally offered by universities and equivalent tertiary educational institutions.</td>
<td>Duration: 2-3 years. Entry requirements: ISCED 3 or 4. Usually: first degree at tertiary level</td>
</tr>
<tr>
<td>ISCED 7 Master of equivalent level</td>
<td>Programmes at ISCED level 7, or Master’s or equivalent level, are often designed to provide participants with advanced academic and/or professional knowledge, skills and competencies, leading to a second degree or equivalent qualification. Programmes at this level may have a substantial research component but do not yet lead to the award of a doctoral qualification. Typically, programmes at this level are theoretically based but may include practical components and are informed by state of the art research and/or best professional practice. They are traditionally offered by universities and other tertiary educational institutions.</td>
<td>Duration: 2-3 years. Entry requirements: ISCED 6. Usually: second degree at the tertiary level. Direct access to ISCED 8 level</td>
</tr>
<tr>
<td>ISCED 7X6 Master or equivalent level long degrees</td>
<td>Long first-degree programme at a Master’s or equivalent level with a cumulative theoretical duration (at the tertiary level) of at least five years (that does not require prior tertiary education). When possible given their different characteristics and their impact on the number of diplomas.</td>
<td>Duration: at least 5 years. Entry requirements: ISCED 3 or ISCED 4. Usually: first degree at ISCED 7X6 level</td>
</tr>
<tr>
<td>ISCED 8 Doctoral or Equivalent level</td>
<td>Programmes at ISCED level 8, or doctoral or equivalent level, are designed primarily to lead to an advanced research qualification. Programmes at this ISCED level are devoted to advanced study and original research and are typically offered only by research-oriented tertiary educational institutions such as universities. Doctoral programmes exist in both academic and professional fields.</td>
<td>Duration: at least 3 years. Entry requirements: ISCED 7. Research-based programs (not only courses).</td>
</tr>
</tbody>
</table>
APPENDIX 3
Summary statistics and Graphs of dependent variables

Figure A.1. Turnover 2011

Figure A.2. Employment 2011
Figure A.3. Assets 2011

Figure A.4. Intangibles 2011
Table A.4. Summary statistics table

<table>
<thead>
<tr>
<th>Variable</th>
<th># observations</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sector NACE</td>
<td>9,084,537</td>
<td>9.566</td>
<td>9</td>
<td>4.485</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>Standardized legal form</td>
<td>9,090,827</td>
<td>5.409</td>
<td>5</td>
<td>1.264</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Size of the company</td>
<td>9,090,838</td>
<td>1.248</td>
<td>1</td>
<td>0.545</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>NUTS-2 Region</td>
<td>9,055,089</td>
<td>151.72</td>
<td>171</td>
<td>79.89</td>
<td>1</td>
<td>267</td>
</tr>
<tr>
<td>R&amp;D intensity</td>
<td>9,090,838</td>
<td>.184</td>
<td>0</td>
<td>.387</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Institutional quality</td>
<td>9,090,838</td>
<td>.538</td>
<td>1</td>
<td>.498</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
APPENDIX 4
Development of specific hypotheses for further empirical research

The creation of the integrated dataset discussed in this Report opens the way for exploring a number of contextual factors that moderate the impact of HEIs on firms. These contextual factors are the object of a large literature, as discussed in the Report. They are as follows.

- Contextual factors moderating the impact of creation of human capital
  - complementarity between formal education and cognitive and non-cognitive skills
  - matching between the competences of graduates and the demand of firms (and associated potential mismatch, under- or over-education effects)
  - quality of education and university-work transition
  - migration of graduates.

- Contextual factors moderating the impact of research
  - quality of academic research
  - regional pattern of specialization and absorptive capacity
  - complementarity and substitution effects between the different pathways
  - complementarity between human capital accumulation and innovation

Table A5 introduces in a succinct way the specific hypotheses that refer, respectively, to the two broad categories of contextual factors discussed above.

In Table A5 we mention several sources of data, most of which are already integrated in the dataset.

### Table A.5. Development of hypotheses

<table>
<thead>
<tr>
<th>Context factor</th>
<th>Hypothesis</th>
<th>Empirical strategy and related variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complementarity between formal education and cognitive and non-cognitive skills</td>
<td>Impact of HEIs is higher in countries in which the general workforce (or general population) has a high level of skills</td>
<td>Classification of countries by using PIAAC data (OECD) on adults literacy Share of population with tertiary degree</td>
</tr>
<tr>
<td>Matching between the competences of graduates and the demand of firms</td>
<td>Impact of universities (PhD granting institutions) is higher in regions in which the sectoral composition is oriented towards science-based industries</td>
<td>OECD classification based on Pavitt taxonomy + Eurostat data</td>
</tr>
<tr>
<td></td>
<td>Impact of non-university institutions is higher in regions in which the sectoral composition is oriented towards traditional industries (supplier-dominated)</td>
<td>OECD classification based on Pavitt taxonomy + Eurostat data</td>
</tr>
<tr>
<td></td>
<td>Impact of HEIs is higher in countries with a dual institutional framework of higher education</td>
<td>Eurydice descriptive dossier on individual countries in Europe</td>
</tr>
<tr>
<td></td>
<td>Impact of HEIs is higher if there is co-specialization at regional level between Field of education and Industry classification</td>
<td>Correspondence table between FoE and NACE classification + Eurostat data</td>
</tr>
<tr>
<td></td>
<td>Impact of HEIs is higher for large firms</td>
<td>Total turnover as a size</td>
</tr>
<tr>
<td>Category</td>
<td>Impact Description</td>
<td>Variable (Orbis)</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Impact of HEIs is higher for multinational corporations</strong></td>
<td>Impact of HEIs is higher if the quality of education is higher</td>
<td>MNC dummy variable (Orbis)</td>
</tr>
<tr>
<td><strong>Quality of education and university-work transition</strong></td>
<td>Impact of HEIs is higher if the quality of education is higher</td>
<td>Academic staff/ Number of students ratio as proxy for quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total expenditure of HEI/ number of students as proxy for quality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>% student fees/ expenditure normalized at national level</td>
</tr>
<tr>
<td><strong>Migration of graduates</strong></td>
<td>In large cities the main impact of higher education is produced by high quality universities</td>
<td>Population size</td>
</tr>
<tr>
<td></td>
<td>In areas other than large cities the main impact of higher education is produced by non-high quality universities</td>
<td>Total number of students as size variable (ETER)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Leiden ranking as proxy for quality</td>
</tr>
<tr>
<td><strong>Quality of academic research</strong></td>
<td>Impact of universities is higher if the quality of their academic research is higher (respectively, is lower or does not have an impact).</td>
<td>% publications and citations in top 10% and 25% SNIP journals (GRBS)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total citations/ Academic staff (Scopus + ETER)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Leiden ranking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>% ISCED 8 students</td>
</tr>
<tr>
<td><strong>Importance of quality of academic research of universities depends on the sectoral composition</strong></td>
<td>As above with separate models using correspondence table</td>
<td>Classification of universities according to the centrality in website networks (Webometrics)</td>
</tr>
<tr>
<td></td>
<td>Impact of universities is higher if they are more central in networks of international research collaboration</td>
<td>Classification of universities according to the centrality in EU collaboration (CORDIS)</td>
</tr>
<tr>
<td><strong>Regional patterns of specialization and absorptive capacity</strong></td>
<td>Impact of academic research is higher if there is matching between scientific and industry specialization</td>
<td>Number of publications and Number of citations (GRBS + Scopus) using correspondence table</td>
</tr>
<tr>
<td></td>
<td>Impact of academic research is higher if there is absorptive capacity in the regional</td>
<td>OECD classification based on R&amp;D intensity</td>
</tr>
<tr>
<td>Complementarity and substitution effects between HEIs missions</td>
<td>Impact of HEIs is higher if education and research activities are complementary</td>
<td>Complementarity if student/staff ratios below the 3rd quartile (ETER)</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Impact of HEIs is higher if research and third mission activities are complementary</td>
<td>Complementary if included in the Leiden co-authorship academy-industry ranking</td>
<td></td>
</tr>
<tr>
<td>Impact of HEIs is higher if they have attitude to source funding from third parties</td>
<td>Third party funding/Total expenditure normalized against the national average</td>
<td></td>
</tr>
<tr>
<td>Impact of HEIs is higher if quality of education and quality of research are jointly at high level</td>
<td>Multiplicative term (interaction effect)</td>
<td></td>
</tr>
<tr>
<td>Complementarity between human capital accumulation and innovation</td>
<td>Impact of HEIs is higher if the recipient firms invest jointly in human capital and innovation</td>
<td>Initial level of Total intangible assets (Orbis) as proxy for past investment</td>
</tr>
</tbody>
</table>
List of abbreviations and definitions

HEIs  Higher Education Institutions
R&D  Research and Development
**List of tables and figures**

**List of tables**

Table 1. Selected variables from Orbis .................................................................32
Table 2. Summary statistics for relevant variables ......................................................33
Table 3. Coverage by country of selected dependent variables ......................................34
Table 4. Variables selected from ETER .........................................................................37
Table 5a. Regression results for models of impact of higher education on growth of firms. Aggregated independent variables. Total variation 2011-2015 ..................................................46
Table 5b. Regression results for models of impact of higher education on growth of firms. Aggregated independent variables. Total variation 2011-2015 ..................................................46
Table 6a. Regression results for models of impact of higher education on growth of firms. Aggregated independent variables. Fixed variable specification. Total variation 2011-2015 ..........................................................................................................................48
Table 6b. Regression results for models of impact of higher education on growth of firms. Aggregated independent variables. Fixed variable specification. Total variation 2011-2015 ..........................................................................................................................48
Table 7a. Regression results for models of impact of higher education on growth of firms. Total variation 2011-2015 ..........................................................................................................................50
Table 7b. Regression results for models of impact of higher education on growth of firms. Total variation 2011-2015 ..........................................................................................................................50
Table 8a. Regression results for models of impact of scientific research on growth of firms. Total variation 2011-2015 ..........................................................................................................................53
Table 8b. Regression results for models of impact of scientific research on growth of firms. Total variation 2011-2015 ..........................................................................................................................53
Table 9a. Regression results for models of joint impact of higher education and scientific research on growth of firms. Total variation 2011-2015 .................................................................56
Table 9b. Regression results for models of joint impact of higher education and scientific research on growth of firms. Total variation 2011-2015 .................................................................56
Table A.1. Fields of education .....................................................................................85
Table A.2. Correspondence table Fields of of Education- Fields of Science .....................86
Table A.3. Levels of education ....................................................................................87
Table A.4. Summary statistics table .............................................................................90
Table A.5. Development of hypotheses .....................................................................91

**List of figures**

Figure 1. Country average of firm performance in 2013 .............................................35
Figure A.1. Turnover 2011 .....................................................................................88
Figure A.2. Employment 2011 ................................................................................88
Figure A.3. Assets 2011 .........................................................................................89
Figure A.4. Intangibles 2011 ..................................................................................89
GETTING IN TOUCH WITH THE EU

In person
All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: http://europea.eu/contact

On the phone or by email
Europe Direct is a service that answers your questions about the European Union. You can contact this service:
- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696, or
- by electronic mail via: http://europa.eu/contact

FINDING INFORMATION ABOUT THE EU

Online
Information about the European Union in all the official languages of the EU is available on the Europa website at: http://europa.eu

EU publications
You can download or order free and priced EU publications from EU Bookshop at: http://bookshop.europa.eu. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see http://europa.eu/contact).
JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre’s mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub
EU Science Hub - Joint Research Centre
Joint Research Centre
EU Science Hub