Bolli, Thomas; Renold, Ursula; Wörter, Martin

Working Paper

Vertical educational diversity and innovation performance

KOF Working Papers, No. 395

Provided in Cooperation with:
KOF Swiss Economic Institute, ETH Zurich

Suggested Citation: Bolli, Thomas; Renold, Ursula; Wörter, Martin (2015) : Vertical educational diversity and innovation performance, KOF Working Papers, No. 395, ETH Zurich, KOF Swiss Economic Institute, Zurich,
https://doi.org/10.3929/ethz-a-010692089

This Version is available at:
https://hdl.handle.net/10419/126237

KOF

KOF Working Papers

Vertical Educational Diversity and Innovation Performance

Thomas Bolli, Ursula Renold and Martin Wörter

No. 395
October 2015

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

KOF

KOF Swiss Economic Institute
ETH Zurich
LEE G 116
Leonhardstrasse 21
8092 Zurich
Switzerland

Phone +41 44 632 42 39
Fax +41 44 632 12 18
www.kof.ethz.ch
kof@kof.ethz.ch

Vertical Educational Diversity and Innovation Performance

Thomas Bolli[1], Ursula Renold[2] and Martin Wörter[3]

Abstract

This paper uses panel data of Swiss firms to analyze the impact of education-level diversity in the workforce on innovation performance, addressing endogeneity by exploiting within-firm variation as well as variation in labor supply across regions. We find that vertical educational diversity increases the extensive margin of R&D and product innovation, particularly new product innovation. However, the relationship with process innovation, R&D intensity, and product innovation intensity is insignificant or even negative. These results are in line with the idea that vertical educational diversity enhances the creative moment of the invention phase, while it might affect the commercialization phase negatively due to the dominance of coordination and communication costs relative to the gains in creativity.

JEL: O3

Keywords: Vertical educational diversity, innovation performance, R&D, product innovation, process innovation

Acknowledgements: We thank the participants of the 15th ISS Conference, the KOF-UZH Seminar and the KOF Brown Bag Seminar for helpful comments and suggestions.

[1] ETH Zurich, KOF Swiss Economic Institute, Leonhardstrasse 21, 8092 Zurich, Switzerland, +41 44 632 61 08, bolli@kof.ethz.ch

[2] ETH Zurich, KOF Swiss Economic Institute, Leonhardstrasse 21, 8092 Zurich, Switzerland, +41 44 632 53 29, ursula.renold@kof.ethz.ch

[3] ETH Zurich, KOF Swiss Economic Institute, Leonhardstrasse 21, 8092 Zurich, Switzerland, +41 44 632 51 51, woerter@kof.ethz.ch

1. Introduction

It is often said that a nation's competitiveness is determined by the proportion of students at universities. The following quote from the WEF Competitiveness Report stands as an example: "While Switzerland demonstrates much competitive strength, maintaining its innovative capacity will require boosting the university enrollment rate of 49.4 percent, which continues to lag behind that of many other high-innovation countries." (Sala-i-Martin et al. 2011, p. 11). However, other studies show an imbalance between the demands on the labor market and the skills profiles of the workforce (Allen, et al. 2011; Livingstone, 2009). We hypothesize in this article that optimizing skill-grade-mix is a more appropriate aim than maximizing a single indicator of educational attainment. To this end, we analyze the impact of vertical educational diversity, i.e. diversity in formal education degrees, on the innovation performance of firms.

The relevance of analyzing the optimal mix is supported by the fact that studies on different aspects of the diversity of the workforce emphasize the creative power on one hand and the costs on the other hand of a heterogeneous workforce—whether that diversity be ethic, religious (see, e.g., Alesina and La Ferrara 2005), or educational (see, e.g., Østergaard et al. 2011). Neither theory nor empirical investigations are clear on the influence of vertical educational diversity on the innovation performance of firms. While, for example, Østergaard et al. (2011) identify a positive relationship between educational diversity and the propensity to innovate, McGuirk and Jordan (2012) cannot confirm the positive relationship for process innovations, and Parrotta et al. (2014) find little evidence that educational diversity affects the patenting propensity or the patenting intensity of firms. The econometric challenges, the specificities of the countries analyzed, the comprehensiveness of the data, or the fact that the innovation process is largely collapsed into a single stage, might be reasons for the inconclusive results on the diversity-innovation relationship. Hence, these circumstances ask for further empirical investigations.

In looking at the cost and benefits aspects of a workforce diverse in education, the paper at hand tries to address some of the above-mentioned issues. First, our data allows to use a wide range of innovation measures. Hence, we do not collapse innovation performance into a single stage, but differentiate between measures related to invention and to commercialization. We distinguish between innovation input (R&D expenditures) and three types of innovation output: incremental product innovations, more drastic product innovations, and process innovations. Consequently, we can fine-grain the effect of diversity on the innovation activities of firms and emphasize the different impacts of diversity on invention and commercialization (March 1991). Aside from differentiating between these measures, we further provide evidence regarding potential differences between the extensive and intensive margins, thereby extending the evidence for the impact of educational diversity on the invention (exploration) and commercialization (exploitation) activities of firms.

Second, our models have a comprehensive control vector comprising firm size, the appropriability of innovation results, the technological potential of a firm, the development of firm demand, and incoming spillovers. In addition, we address the potential endogeneity of vertical educational diversity by exploiting within-firm variation as well as instrumenting vertical educational diversity by the labor supply available in the region.

Third, we look at vertical educational diversity—the diversity of education levels such as compulsory, upper secondary and tertiary education—while most of the literature focuses on horizontal educational diversity—diversity in terms of study field on the tertiary level. This directly links to education policy issues in general, and is less related to the mix of academics from different study fields on which most of the literature focuses.

Fourth, we provide the first evidence for Switzerland, thereby extending the scant existing empirical literature that analyzes Danish (Østergaard et al., 2011, and Parrotta et al. 2014) and Irish firms (McGuirk and Jordan, 2012). Switzerland is a particularly interesting case due to its strong vocational and professional education system (OECD 2010, Hoffman and Schwartz 2015). Furthermore, its industry structure is different compared to Denmark or Ireland, with a heavy emphasis on specific high-tech industries like pharmaceuticals, machinery, electrical equipment, and electronics/optical products. Switzerland is a technologically very innovative country, which presumably requires a higher degree of technological diversity[4] and a high level of technological skills and formal education. Given these circumstances, it is not clear how educational diversity affects the innovation performance of such a country.

Based on Swiss firm-level panel data comprising five waves of the Swiss Innovation Survey covering the period 1999 to 2011 we find that vertical educational diversity is significantly positively related to the probability to conduct R&D and to launch new or improved products. This indicates that vertical educational diversity increases the firm's capability and incentives to explore new knowledge or to develop a new product. We also find that vertical educational diversity is unrelated or even negatively related to process innovation as well as the commercial success of new or improved products. This means that vertical educational diversity tends to be negatively related with the

[4] Leten et al. (2007), Garcia-Vega (2006), or Bolli and Woerter (2013) found that technological diversity is positively related with new patent applications.

commercial exploitation of the results of R&D or innovative activities. This is a relatively new finding and it pioneers improved theorizing on the diversity-innovation relationship.

The reminder of the paper is organized as follows. Section two introduces the conceptual framework and provides a literature review. Section three describes the data used and section four informs about the econometric models applied. Section five presents the results and section six concludes.

2. Conceptual Framework and Literature Review

From a theoretical point of view it is an unresolved question whether diversity in formal education levels should foster innovativeness and increase the innovation output of a firm. Like other types of knowledge diversity (see, e.g., Laurson, 2012), educational diversity has two opposite effects on innovation ability referring to the cost and benefits of the collaboration process.

Costs and benefits of vertical educational diversity

On the one hand, vertical educational diversity might increase innovation performance. Different types of education might provide alternative bodies of knowledge (see, e.g., Jacobsson and Oskarsson 1995, Hong and Page 1998, Carlile 2002, Faems and Subramanian 2013) which can be combined on the firm level and improve decision making. Collaboration of employees with different educational backgrounds (i.e. academic or vocational education) along with different experiences, insights, or interests might cause different interpretations of problems, enhance problem awareness, and increase the spectrum of problem solutions. This is not least because diversity is likely to improve the absorptive capacity of a firm (Cohen and Levinthal, 1989, 1990; Quintana-Garcia and Benavides-Velasco 2008). This makes it easier to identify valuable knowledge surging from the research activities of other firms and institutions, and at the same time less likely that promising new ideas or technologies will pass by unnoticed by the firm. In cases where a firm pursues different research projects simultaneously, cross-fertilizing spillovers across the projects might arise (Weitzmann, 1998). Hence, decisions are improved if different perspectives are involved in the decision making process (Alesina and La Ferrara 2005), which can improve, for example, the capacity to develop a new prototype into a marketable innovation: this requires not only scientific views and expertise, but also professional competences, work experience, and last but not least soft skills (Bolli and Renold, 2015) in many different fields of knowledge.

On the other hand, vertical educational diversity might decrease innovation performance, since vertical educational diversity can increase the level of conflicts, mistrust, and misunderstandings due to high cognitive distances. As a consequence, vertical educational diversity is likely to increase the communication and coordination costs of integrating available knowledge or coordinating the innovation process (see, e.g., Wittenbaum and Stasser, 1996, Stasser and Titus, 1985, Dahlin et al., 2005). According to social identity theory, such coordination and communication costs might arise because individuals value members of their own social identity more highly. This indicates a potential for competitive behavior and conflict due to vertical educational diversity (see, e.g., Joshi and Jackson, 2003). Consequently, the risk of failure might increase. An additional opportunity cost of vertical educational diversity stems from economies of scale in the knowledge production process. These arise in cases where a concentration of workers with similar education level and similar knowledge base are necessary for an efficient production process, such as when a sufficient number of academics is required to run a laboratory.

These two opposing forces might explain the mixed empirical results provided by the literature concerning the impact of educational diversity on the innovation performance of firms (for an overview see Table 1). Østergaard et al. (2011) match data from the Danish innovation survey to employee data and find that horizontal educational diversity of employees with tertiary education—diversity in terms of thematic background at the same education level—improves the probability of introducing an innovation. However, they claim that this positive relationship might decrease for higher levels of horizontal diversity. Also based on Danish employer-employee data but merged this time with patent data, Markus and Kongsted (2012) find that hiring R&D workers distant from one another in the educational space (Jaffe 1986), improves exploratory patent applications. Their focus on R&D workers suggests that their measure of diversity is mainly driven by horizontal diversity. Furthermore, the benefit of hiring distant workers decreases with rising diversity in the existing workforce. Parrotta et al. (2014) use the same dataset as Parrotta et al. (2012) to analyze the effect of horizontal educational diversity on innovation performance. In addition to instrumenting educational diversity by exploiting regional variation, they use pre-sample information to account for unobserved firm characteristics. They also include measures of knowledge spillovers based on geographic and technological distances to account for external knowledge. They find little evidence that horizontal educational diversity affects patenting propensity, patenting intensity, or patenting diversity. The different meta-analyses of the literature on the relationship between team-member diversity and innovation performance provide mixed results as Williams and O'Reilly (1998), Horwitz (2005), Horwitz and Horwitz (2007) and Hülsheger et al. (2009) suggest a positive relationship, while the more recent paper by van Dijk et al. (2012) finds no relationship.

Soellner (2010) uses a linked employer-employee panel dataset of German firms in order to investigate the relationship between an entropy index of occupational diversity and the propensity of a firm to innovate. He finds a positive relationship between diversity and the introduction of a completely new product for the manufacturing industry, and a significant relationship between diversity and the introduction of an innovative product (improved products, new to the business, or new to the market) for the service industries.

There are only a few papers that focus on vertical educational diversity across education levels. McGuirk and Jordan (2012) use Irish firm data to estimate the impact of educational diversity in Irish counties on the propensity to introduce product and process innovation. Calculating a Blau Diversity index for each Irish county based on six categories (primary school, lower secondary school, upper secondary school, third-level non-degree, and third-level degree or higher), they find that educational diversity improves product innovation but not process innovation. They further find that educational diversity on the labor market acts as a substitute for absorptive capacity, measured as internal tertiary education share.

Subramanian et al. (2015) base their educational diversity analysis on data from the national R&D survey in Singapore and use patents as a measure of innovation performance. They analyze vertical educational diversity, measured by one minus the Herfindahl concentration index, but focus on education levels within tertiary-educated employees only. They show mixed results. In their baseline estimation they do not detect any significant differences in innovation between similar and diverse educational level populations in the workforce of research scientists and engineers. However, when they interact educational diversity levels with the heterogeneity of technological domains, they find a positive moderating effect for similar educational levels, such that educational similarity (not diversity) shows a positive effect on innovation for heterogeneous technological domains.

Similarly, Faems and Subramanian (2013) could not find any significant relationship between diversity in terms of different types of educational degrees (PhD, master, bachelor, postgrad, or no academic degree) among R&D manpower and their technological performance. Hence, they analyze vertical educational diversity but focus on tertiary degrees.

On the innovation system level, Meuer et al. (2015) applied a vertical educational diversity measure in order to characterize different types of innovation systems. They found that specialization is the key characteristic of the “autarkic” innovation system, which shows an equal propensity to generate radical, technological, and organizational innovations. However, the paper doesn’t address potential endogeneity of vertical educational diversity.

Table 1: Overview of firm-level studies examining the educational diversity-innovation relationship

Author(s)	**Dependent Variable**	**Sample**	**Treatment of Endogeneity**	**Horizontal vs Vertical Diversity**	**Diversity Measure**	**Results**
Williams and O'Reilly (1998), Horwitz (2005), Horwitz and Horwitz (2007), Hülsheger et al.(2009)	Innovation	Meta-analysis of Team Literature	Mostly Correlation	Mostly horizontal	Various	Diversity increases innovation
van Dijk et al. (2012)	Innovation	Meta-analysis of Team Literature	Mostly Correlation	Mostly horizontal	Various	No effect of diversity on innovation
Meuer et al. (2015)	Intensity of New and Improved Product Innovation and Organization Innovation	Swiss firm data	Correlation	Vertical diversity 5 categories: -Apprentices -Lower Secondary -Upper Secondary -Vocational Tertiary -Academic Tertiary	Blau Diversity Index	Specialization characterizes "autarkic" innovation system, which is unrelated to innovation
Østergaard et al. (2011)	Product innovation 0/1	Matched Danish innovation survey and employee data	Correlation	Horizontal diversity: 16 categories of tertiary education orientation: social sciences, humanities, food/health care, engineering, natural sciences, high school teacher, military officer	Entropy-Index	Positive effect of diversity on innovation
Markus and Kongsted (2013)	Exploratory innovation (Number of new citation/total citations)	Danish employer-employee matched panel data, patent data from the EPO	Lagged explanatory variables	Horizontal diversity: Incumbent R&D workers' educational background	HHI	Educational diversity among incumbent R&D workers negatively moderates the relationship between cognitive distance and exploratory search.
Faems and Subramanian (2013)	Number of patents	Firm survey in Singapore	Correlation	Vertical diversity: Education level of R&D workforce (i.e. Ph.D., master, bachelor, post-grad, or no academic degree)	HHI	No effect of diversity on innovation
Parrotta et al. (2014)	Patenting 0/1 and number of patents	Matched employer-employee dataset for Denmark	Fixed effect model with serial dependence	Combination of horizontal and vertical diversity: -4 categories of tertiary education: engineering, humanities, natural sciences, social sciences -Secondary education -Compulsory education	HHI	No effect of diversity on innovation
McGuirk and Jordan, 2012	Product and process innovation 0/1	Firm survey in Ireland	Irish counties as instrument	Vertical diversity: 6 categories: -Primary school -Lower secondary school -Upper secondary school, -Third-level non degree -Third-level degree -Higher	Blau Diversity Index	Diversity increases product innovation but not process innovation.
Soellner, 2010	Product innovation 0/1	German firms	Correlation (random effects)	Horizontal diversity: Employment shares of occupations (three-digit level)	Entropy index	Positive effect of diversity on innovation
Subramanian et al., 2015	Number of patents	Firms in Singapore	Correlation	Vertical Diversity: Education levels of researchers and engineers: - Ph.D. - Master - Bachelor - Associated degree levels	HHI	No average correlation. Positive (negative) moderating effect for similar (diverse) educational levels in the case of heterogeneous technological domains.

Innovation stages

The benefits of stimulating creativity and the costs of coordination and communication might have different weights at different stages of the innovation process. Hence, one potential reason for the divergences observed in the empirical literature is that they collapse the innovation process into a single stage, stopping on the level of patents as a measure for innovation and neglecting the complexity of the innovation process. In a nutshell, the innovation process of a firm can be seen as a continuum[5], starting with a creative process that leads to a new product or process idea, which will be further pursued involving more routinized work processes and expertise from different departments and employees. In the end of this process we should see a new, commercially successful product or a more efficient production process. The creative process involves the generation of new knowledge—frequently through a trial-and-error process—which requires a certain level of risk-taking, experimentation, flexibility, and a good command of existing knowledge in the field.

A number of papers suggest that the creativity benefits are relatively more relevant in the early stages, while coordination and communication costs are more important for later stages in the innovation process (see, e.g. Mansfield and Wagner, 1975, Atuahene-Gima and Evangelista, 2000, Danzon et al., 2005, Arora et al. 2009, Plotnikova, 2009, Bolli and Wörter, 2013). These findings suggest that the strength of diversity lies more in exploring new ideas and research paths than in executing more routinized working processes like merchandizing a new product using existing distribution channels and marketing methods. Hence, it is plausible to assume that vertical educational diversity is positively related to early innovation activities and unrelated or even negatively related to later stages of the innovation process.

Hypotheses

Our measures of innovation activities allow us to model the different innovation stages in two dimensions. First, we differentiate between R&D activities, new or radical product innovations, improved or incremental product innovations, and process innovations. Second, we differentiate between the extensive margin and the intensive margin of each innovation stage respectively. We argue that the extensive margin is more closely related to earlier innovation stages than the intensive margin, which depends substantially on the process of commercialization in the case of product

[5] We refer to organizational steps and do not consider feedback mechanisms, which are very important but not relevant for the point we want to make.

innovations and the process of implementation in the case of process innovations. As depicted in Figure 1, we develop four pairs of hypothesis regarding these variables.

R&D activities:

H1a: Vertical educational diversity increases the extensive margin of R&D activities

H1b: Vertical educational diversity increases the intensive margin of R&D activities

New product innovation:

H2a: Vertical educational diversity increases the extensive margin of new product innovation

H2b: Vertical educational diversity increases the intensive margin of new product innovation

Incremental product innovation:

H3a: Vertical educational diversity increases the extensive margin of improved product innovation

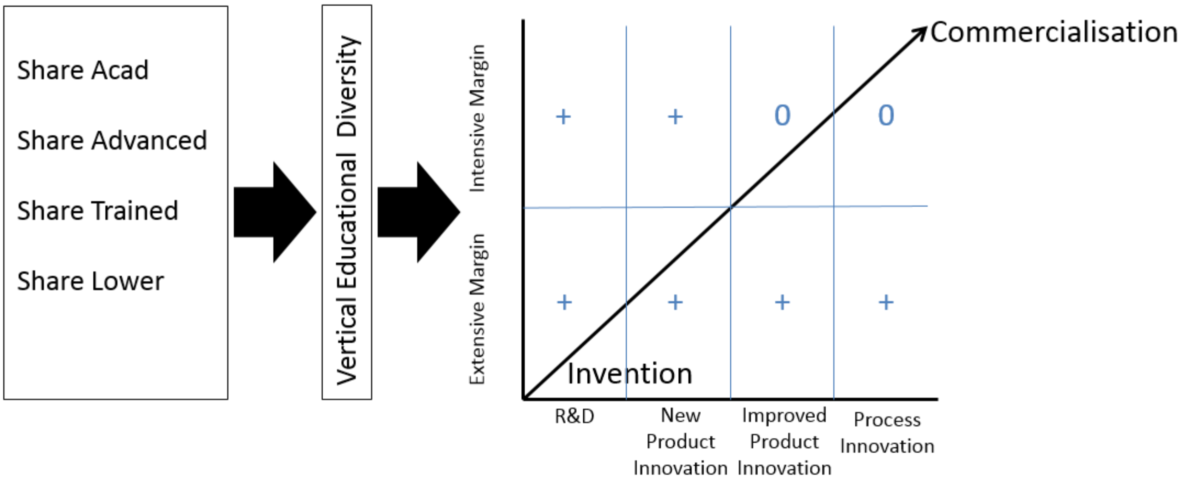
H3b: Vertical educational diversity has no effect on the intensive margin of improved product innovation

Process innovation:

H4a: Vertical educational diversity increases the extensive margin of process innovation

H4b: Vertical educational diversity has no effect the intensive margin of process innovation

Figure 1: Expected relationship between vertical educational diversity and innovation measures



3. Data

The panel data we employ stems from six waves of the Swiss innovation survey conducted by the KOF Swiss Economic Institute (www.kof.ethz.ch) in the years t={1996, 1999, 2002, 2005, 2008, and 2011}, where t denotes the time period. The surveys are based on a disproportionately stratified random sample of firms with more than five full-time equivalent (FTE) employees covering the most important industries of the manufacturing, construction, and service sectors. Stratification takes place on the industry level and in three firm-size classes. Responses were received from 1748 firms (32.5%), 2172 firms (33.8%), 2583 firms (39.6%), 2555 firms (38.7%), 2141 firms (36.1%), and 2363 firms (35.9%) for the years 1996, 1999, 2002, 2005, 2008, and 2011 respectively. In sum, we can make use of 13143 observations. However, depending on the specification of our model the number of observations we use fluctuates significantly.

3.1. Measurement issues

In this section we summarize the measurement of the main variables as defined in Table 2, which are measures for vertical educational diversity, R&D activity, new product innovation, improved product innovation, process innovation, and control variables.

Vertical educational diversity

Science has developed a number of different measures of diversity (see, e.g, Stirling, 1998, or Dawson 2011). However there is no consensus about how to measure it. In theoretical terms there are two main concepts: the first concept refers to self-categorization through social identity theory, and the second concept refers to the information/decision making perspective (Faems and Subramanian, 2013, Dawson, 2011). The latter concept is relevant for the investigation at hand since it is related to the variance in group composition in terms of things like skill levels and abilities (Dawson, 2011, Faems and Subramanian, 2013). Conceptually we define diversity as the distribution of differences among the staff of a firm with respect to a common attribute (Harrison and Klein, 2007). The common attribute in the investigation at hand is the level of education. The predominant methods of measuring the diversity of a common attribute in the literature on relationships between innovation and diversity in technologies or education are the "Herfindahl-Hirschman-Index" (HHI) and an entropy index (see, e.g., Østergaard et al. 2011). The HHI shows weak properties on the lower and upper ends of the scale. Bolli and Woerter (2013) inserted additional dummy variables to address the weakness of the upper end of the scale, while the lower end (highly diverse firms) was not relevant for the data used here. More sophisticated measures including multi-attribute approaches were applied by Nehring and Puppe (2002), Stirling (2004, 2007), and Woerter (2009). However, a multi-attribute approach is not the first choice for the paper at hand since we refer only to one attribute—the level of education—when measuring diversity. Since calculating an entropy index yields qualitatively the same results, we follow the literature that is more related to the investigation at hand and apply an HHI index.

In order to calculate the vertical educational diversity measure (Diversity Index), we refer to four categories of educational degrees: j={Lower (apprentices[6] and untrained), Trained (upper secondary education), Advanced (professional tertiary education, incl. university of applied sciences), and Acad (conventional university tertiary education)}. Following Garcia-Vega (2006), we define the vertical educational diversity (Diversity $Index_{it}$) of firm i in period t as one minus the HHI, which is calculated as the sum of squared formal education shares within a firm:

$$DiversityIndex_{it} = 1 - \left(\sum_{j} \left(N_{jit} / N_{it} \right)^2 \right) \quad (1)$$

N_{jit} denotes the number of full-time equivalent employees in education category j of firm i in period t. N_{it} refers to the number of full-time equivalent employees of firm i in period t.

[6] Apprenticeship status refers to an ongoing education rather than to an educational degree.

R&D activities

We apply two different measures for the R&D activities of a firm. First, based on the survey data we apply a dummy variable for the R&D activities of a firm (R&D 0/1). If a firm pursues R&D, our proxy for R&D activities indicates 1, otherwise 0. Second, we have data on the sales share of R&D expenditures over a three year period, which we use to represent R&D Intensity.

Innovation activities (product and process)

Given a clear definition and examples of what product and process innovations are, we measure innovation in the last three years based on the survey results. We distinguish three binary measures for innovation activities, namely new product innovation (New Product Innovation 0/1), improved product innovation (Improved Product Innovation 0/1) and process innovation (Process Innovation 0/1). Correspondingly, the quantitative measures of each innovation activity refer to the sales shares of new products (New Product Innovation Intensity) and improved products (Improved Product Innovation Intensity), and the cost reduction due to process innovations over sales (Process Innovation Intensity).

Control variables

Control variables include firm size in terms of FTE employees (Size) and firm size squared ($Size^2$). They also include the technological potential outside the firm (Technological Potential), the appropriability of innovation activities—captured by whether ease of copying innovations represents an innovation-hampering factor (Appropriability)—and the importance of incoming spillovers measured by the relevance of external sources for innovation activities (Incoming Spillovers). Furthermore, they include a variable capturing the development of demand (Deltasales).

We control for the share of FTE employees of the education-level categories Share Trained, Share Advanced, and Share Acad, with Share Lower as the baseline category. Controlling for these shares is important to capture the direct effect of education level on innovation performance, allowing us to identify the effect of vertical educational diversity itself. In addition, we include a dummy variable indicating whether vertical educational diversity takes the value of 0 to capture the extreme cases of firms in which all employees fall into a single education-level category.

Table 2: Variable descriptions

Dependent Variables	
R&D 0/1	Dummy variable that takes the value 1 if a firm has R&D expenditures and 0 otherwise.
R&D Intensity	R&D expenditures over total sales
New Product Innovation 0/1	Dummy variable that takes the value 1 if a firm has new product innovations and 0 otherwise.
New Product Innovation Intensity	Sales stemming from new product innovations over total sales
Improved Product Innovation 0/1	Dummy variable that takes the value 1 if a firm has improved product innovations and 0 otherwise.
Improved Product Innovation Intensity	Sales stemming from improved product innovations over total sales
Process Innovation 0/1	Dummy variable that takes the value 1 if a firm has process innovations and 0 otherwise.
Process Innovation Intensity	Cost reduction due to process innovation over total sales
Explanatory Variables	
Share Lower	Share of untrained employees and apprentices
Share Trained	Share of employees with an upper secondary education
Share Advanced	Share of employers with a professional tertiary education (incl. university of applied sciences)
Share Acad	Share of employers with a conventional university (academic) tertiary education
Diversity Index	One minus the Herfindahl-Hirschman concentration index, calculated as the sum of squared education shares
Diversity Index=0	Dummy variable that takes the value 1 if the Diversity Index is 0, and 0 otherwise.
Instrumental Variables	
Region Share Lower	Mobility region share of untrained employees or apprentices
Region Share Trained	Mobility region share of employees with an upper secondary education
Region Share Advanced	Mobility region share of employers with a professional tertiary education (incl. university of applied sciences)
Region Share Acad	Mobility region share of employers with a conventional university tertiary education
Region Diversity Index	One minus the mobility region Herfindahl-Hirschman concentration index (calculated as the sum of squared education shares)
Control Variables	
Size	Number of full time-equivalent employees
Appropriability	Dummy variable that takes the value 1 if innovations are difficult to copy. Scores 1 if they select 1 or 2 on a 5-point Likert scale capturing the extent to which ease of copying represents an innovation hindrance, and 0 otherwise
Incoming Spillovers	Sum of the relevance of 3 external knowledge sources (Other companies, Institutions/Consulting, Generally available information) on a 5-point Likert scale
Technological Potential	Technological potential outside the firm on a 5-point Likert scale
Deltasales	Difference of demand on the main sales market in the next 3 years and the past 3 years on a 5-point Likert scale (1=strong reduction, 5=strong increase)

Notes: All variables except dummy variables and Deltasales enter in logs, where 0s are replaced by 0.1.

3.2. Descriptive information

Table 3 presents descriptive information about our dependent and independent variables. We have between 2035 and 13143 observations depending on the variables used in our econometric models. Variables stemming from questions related to the intensive margin of R&D and innovation activities have a significantly lower number of observations, since only firms with R&D and innovation activities can answer these questions. Particularly low are the number of observations for process intensity; these figures are only available for firms with process innovations and significant production cost reductions due to new processes introduced into the firm.

Table 3: Summary statistics of dependent and explanatory variables

Variable	Obs	Mean	Std. Dev.	Min	Max
R&D 0/1	13143	0.401887	0.490298	0	1
New Product Innovation 0/1	13143	0.478125	0.49954	0	1
Improved Product Innovation 0/1	13143	0.476756	0.499478	0	1
Process Innovation 0/1	13143	0.437039	0.496039	0	1
R&D Intensity*	5097	3.252473	6.507942	0.00009	98.72554
New Product Innovation Intensity*	6284	34.83085	27.18615	1	100
Improved Product Innovation Intensity*	6266	34.83126	27.16087	1	100
Process Innovation Intensity*	2035	10.2286	7.242676	1	60
Share Lower*	13143	0.327925	0.240771	0	1
Share Trained*	13143	0.472663	0.230823	0	1
Share Advanced*	13143	0.143538	0.144874	0	1
Share Acad*	13143	0.055865	0.119482	0	1
Diversity Index=0	13143	0.021532	0.145156	0	1
Diversity Index*	13143	0.498828	0.151103	0	0.75
Technological Potential*	13143	2.678536	1.136104	1	5
Incoming Spillovers*	13143	7.109979	1.939901	3	15
Deltasales	13143	-0.00053	0.543672	-2	2
Appropriability	13143	0.580689	0.493465	0	1
Size*	13143	239.2996	1574.276	1	60000

Notes: *Summary statistics display values before taking logs

We observe R&D activities in 40% of the sample (R&D 0/1). 48% of firms have launched a new product (New Product Innovation 0/1), and also around 48% of firms have launched essentially improved products (Improved Product Innovation 0/1). Process innovations (Process Innovation 0/1) have been introduced by 44% of firms. Referring to the intensity measures, we see that firms spent on average 3.25% of their turnover on R&D (R&D Intensity). New or essentially improved products (New Product Innovation Intensity, Improved Product Innovation Intensity) yielded on average 34% of the total turnover, and process innovation (Process Innovation Intensity) resulted in cost reductions of 10% on average. About 33% of the employees in a firm are untrained or have the status of an apprentice (Share Lower), 47% are trained and have an upper secondary education (Share Trained), 14% of employees have completed a professional tertiary education (Share Advanced), and only 6% have conventional university tertiary educations (Share Acad). The “Diversity Index” shows a rather diverse composition of the workforce in Swiss firms, with an index value of 0.5 on average where the highest value is 0.75. The Technological Potential of the sample firms amounts to a value of 2.68 on average (maximum 5), Incoming Spillovers yield an average level of 7.11 (maximum 15), and the Appropriability of research results is considered by 58% of firms to hinder innovation. Due to extreme values, the average number of employees of our sample firms amounts to 239.

4. Econometric methods

In order to investigate the relationship between vertical educational diversity and innovation performance, we apply two different types of models—one each for the extensive and intensive margins. First, we apply a logit estimator in order to estimate the extensive margins for R&D, new product innovation, improved product innovation, and process innovation, respectively.

$$\begin{cases} Inno\ 0/1_{it} = 1, & if \quad Inno\ 0/1^{*}_{it} > 0 \\ Inno\ 0/1_{it} = 0, & otherwise \end{cases} \tag{2}$$

with:

$$Inno\ 0/1_{it} = \gamma_1 Diversity\ Index_{it} + \gamma_2 Education\ Shares_{it} + x'_{it}\beta + \varepsilon_{it} \tag{3}$$

where $Inno\ 0/1_{it}$ represents dummy variables indicating whether firm i at time t has conducted R&D activities, introduced new or improved product innovations, or has introduced process innovations (extensive margin). The vector x'_{it} includes a comprehensive set of control variables for firm i at time t that might affect the propensity to conduct R&D or to have successful product or process innovations. x'_{it} comprises important determinants for innovation activities (Cohen 2010). Concretely, it includes technological potential, appropriability of innovation activities, importance of incoming spillovers, changes in demand, firm size, sector fixed effects, and time fixed effects (see Table 2 for exact definitions of the variables). The error term ε_{it} is clustered on the firm level.

Most importantly, our models also include measures of the education mix employed by the firm. Specifically, we analyse the impact of shares of employees for trained staff (Share Trained), staff with professional tertiary education (Share Advanced), and staff with conventional university tertiary education (Share Acad), where the share of apprentices and untrained staff (Share Lower) serves as the reference category. While education shares represent standard elements of an innovation equation and hence do not need to be further explained, the inclusion of single education-level items deserves more explanation in equations that also include vertical educational diversity (*Diversity Index*). It is important to account for these shares in order to capture the direct effect of education-level shares on innovation performance. Hence, if we do not explicitly control for the formal education shares within a firm, the estimated diversity-innovation relationship might be biased. For similar reasons, Østergaard et al. (2011) introduced a binary variable representing whether a firm has tertiary-level employees in addition to their diversity measure. Likewise, Parrotta et al. (2014)

introduced controls into their model on workforce composition in terms of formal education in order to improve the precision of the estimates on the diversity index.

Second, we estimate an OLS estimation for our quantitative dependent variables that captures the intensive margin of innovation activities.

$$Inno\ Intensity_{it} = \gamma_1 Diversity\ Index_{it} + \gamma_2 Education\ Shares_{it} + x'_{it}\beta + \varepsilon_{it} \qquad (4)$$

$Inno\ Intensity_{it}$ identifies the intensity (intensive margin) of R&D activities, new product innovations, improved product innovations, or the quantitative success of process innovations. In order to account for potential selection, x'_{it} in the intensive margin equation entails the predicted Mills ratio of the corresponding extensive margin equation in addition to the variables mentioned above. Since we estimate the selection equation and innovation performance separately, we block-bootstrap the standard errors on the firm level with 1000 repetitions.

Identification strategy

It is possible that the relationship we observe between vertical educational diversity and the innovation performance of a firm is influenced by other, unobserved variables. For instance, it is possible that the quality of management or geographical proximity to universities influences both the vertical educational diversity of a firm and its innovation performance. Furthermore, the possibility of reverse causality exists: higher innovation performance might induce a higher vertical educational diversity. Consequently, the effect of vertical educational diversity on innovation performance observed here is very likely to be endogenous.

We address potential endogeneity in two ways. First, we estimate fixed effects models. By relying on within-firm variation across time, these models account for unobserved time-invariant heterogeneity.

Second, we address potential endogeneity due to unobserved heterogeneity and reverse causality by pursuing an instrumental variable approach. More specifically, following Parrotta et al. (2012) and McGuirk and Jordan (2012), we instrument vertical educational diversity of the firm with the vertical educational diversity of the working age population in a mobility region. There are 106 mobility regions in Switzerland as defined by Swiss Statistics.[7] Since the current location of a firm is not

[7] See http://www.bfs.admin.ch/bfs/portal/de/index/regionen/11/geo/analyse_regionen/03.html for more information

randomly chosen, we use the historical composition of formal education in the region a firm is settled. In this case, we calculate the vertical educational diversity in the region in 1990.

The rationale behind this instrument is that pre-existing vertical educational diversity at the regional level is unlikely to be correlated with the innovation performance of a firm many years later, since firms' innovation performances fluctuate more strongly than the educational composition of their regions. Furthermore, the instrument is positively correlated with the diversity of formal education in a firm located in the respective region, since it is likely that the availability of skills and competences is a necessary condition for formal educational diversity within a firm. Moreover it is clearly beyond the influence of a single firm to determine the workforce composition of a region. Therefore, reverse causality does not affect the instrument. Hence, we think that historical diversity on the regional level is a good instrument for firm-level diversity.

Once we move our estimations to the level of regions, we also measure the education shares on the levels of the region, in order to increase the precision of the diversity measure. Since we estimate the instrumental variable approach to the extensive margin manually by estimating a logit model of the first stage and including the predicted values in the main equation, we account for non-simultaneity by block-bootstrapping the standard errors on the level of mobility regions with 1000 repetitions. We estimate the IV of the intensive margin estimations by 2SLS. However, as discussed above, we include the inverse Mills ratio based on a non-simultaneously estimated selection equation. Hence, we block-bootstrap the standard errors of the intensive margin equations on the level of mobility regions with 1000 repetitions to account for non-simultaneity of the instrumental variable estimation with the prediction of the inverse Mills ratio.

The main concern regarding the validity of the instrument is that there is unobserved heterogeneity across regions that is related to both vertical educational diversity and innovation performance. We address this issue with a number of robustness checks reported in Tables A2 and A3 in the appendix. First, the industry structure across regions is different and this might determine the innovation performance of a firm. Hence, we replace sector dummies by 2-digit industry dummies in order to equalize differences in industry structures (see Within-Industry estimates). Second, we additionally inserted area dummies. Following Swiss Statistics, we can divide Switzerland into seven geographical areas: "Lake of Geneva," "Espace Mittelland," "North-East-Switzerland," "Zurich," "East-Switzerland," "Central-Switzerland," and "Tessin." It is likely that there are uncontrolled differences across regions, like different research infrastructure or geographical proximity to an international airport. Inserting area dummies suggests that the identification relies on within-region variation only, thereby equalizing such area differences that might be related to diversity on the regional level and the innovation performance of firms (see Within-Area estimates). Third, since there are firms in our sample that were founded after 1990, their location decision could be affected by the vertical educational diversity of the working-age population in the region. Consequently, we conduct a robustness test excluding those relatively young firms (see Only Old Firms estimates). Fourth, we include the average value of the dependent variable in the corresponding region in the estimation to control for the average innovation performance in the region (see DepVar Average estimates).

Table A1 in the Appendix presents summary statistics for the main instrumental variable by canton. The table shows the means of education level shares and vertical educational diversity in the working-age population across mobility regions in each of the 26 cantons. "Area" in Table A1 refers to the 7 greater areas as defined by Swiss Statistics. The abbreviations for the Cantons follow the denotations of Swiss Statistics[8]. There are significant differences across the Swiss cantons in terms of the average shares of formal education among employees as well as in terms of the "Diversity Index." The average share of untrained people or apprentices is in particularly high in the canton of Jura (JU) and in particularly low in the cantons of Zürich (ZH) and Zug (ZG). The share of trained population is less diverse across Cantons. Bern (BE), Schaffhausen (SH), and Zürich (ZH) show relatively high shares of trained people, and Jura (JU) a particularly low share. The share of people with professional tertiary education is high in Aargau (AG), Geneva (GE), and Zürich (ZH); and relatively low in Appenzell I. Rh. (AI) and Jura (JU). Not surprisingly, the share of people with university tertiary education is very high in urban Geneva (GE) and extremely low in more remote areas like Appenzell I. Rh. (AI) and Uri (UR). The most diverse populations in terms of formal education (Diversity Index) can be found in Geneva (GE), Vaud (VD), and Zug (ZG). Hence, vertical educational diversity seems to be positively correlated with the share of conventional university tertiary education degrees in a canton. That is not surprising, given that the share of academics is relatively low compared to the other educational groups in the respective cantons so higher shares of academics balances the imbalance among education groups. As a result, vertical educational diversity increases.

[8] http://www.bfs.admin.ch/bfs/portal/de/index/dienstleistungen/premiere_visite/03/03_02.html

5. Results

Table 4a presents the results for the extensive margin estimations and Table 4b for the intensive margin estimations. We show three estimations for the effects of the Diversity Index on the probability to conduct R&D, to successfully launch a new product, to successfully launch an improved product, and to successfully implement a process innovation. The first estimation (Logit) is a simple logit regression, followed by a fixed effects logit estimation (FE Logit), and the instrumental variable estimation that instruments education level diversity by the vertical educational diversity of employed population in the mobility region (IV). The components of the Diversity Index—education level shares—are measured on the same level of aggregation as the instrument, which is the regional level. In addition, Tables A2 and A3 in the Appendix show the robustness checks of the IV estimates discussed above.

The estimations in Table 4a support hypothesis 1a, suggesting that vertical educational diversity increases the propensity to conduct R&D significantly. This finding holds in the simple logit estimation as well as in the estimations accounting for endogeneity by exploiting within-firm variation and instrumenting vertical educational diversity. Table A3 in the appendix shows that the IV estimates are robust across specifications.

Similarly, Table 4a supports hypotheses 2a and 3a. Vertical educational diversity increases the propensity to introduce both new and improved products in the logit, fixed effects logit, and IV estimates. It should be noted that the statistical significance of the IV estimates are somewhat less robust, particularly in the case of improved product innovation. However, the estimated marginal effects remain unaffected by the robustness checks. Hence, we conclude that hypotheses 2a and 3a hold, particularly since our preferred specification—the fixed effects logit estimate—support these hypotheses.

The results regarding the introduction of process innovation, however, fail to support hypothesis 4a. While the simple logit estimates indicate a positive relationship between vertical educational diversity and process innovation, the fixed effects logit estimates turn insignificant. Furthermore, though the baseline IV estimates are significantly positive, the robustness checks shown in Table A3 in the appendix suggest that the relationship between vertical educational diversity and process innovation propensity is insignificant.

Looking at the control variables, we see that—as expected—technological potential is significantly and positively related to the propensity to conduct R&D as well as to introduce new product innovations, improved product innovations, and process innovations. Also, our proxy for incoming spillovers is significant and positively correlated with the extensive margin of innovation performance measures, though it becomes insignificant in the fixed effects and instrumental variable estimations. Appropriability is negatively related with the dependent variables in Table 4a except for the case of process innovations. This indicates that high appropriability increases barriers to entry into the markets for innovative products and processes, since leakage of information or technologies from incumbents is unlikely, which increases the costs for potential entrants. The effect of firm size on the probability to conduct R&D or to be innovative is inverted-U-shaped: the firms most likely to conduct R&D or be innovative are mid-sized firms.

Usually one would expect that all education shares included in the estimation increase innovation performance compared to the reference share of employees without upper secondary education (Share Lower). This is particularly true regarding the expectation that the share of academics is positively related to innovation activities (see, e.g., McGuirk and Jordan, 2012). From this perspective, it is surprising that the estimates for the share of workers with professional tertiary education (Share Advanced) and conventional university tertiary education (Share Acad) remain insignificant in the fixed effects logit estimations, though their estimated marginal effects remain positive. The only exception is the positively significant effect of Share Acad on R&D propensity. The regional Share Acad even becomes negative in the IV estimates. However, this finding reflects a multicollinearity problem between Share Acad and Share Advanced.[9] Nevertheless, it should be noted that we find a robust positive and significant relationship between the Share Trained and innovation performance.

[9] Preliminary analysis suggest that including only Share Trained and the combined Share Advanced and Share Acad in the estimation yields a significantly positive effect of the combined share on innovation performance. Since our analysis focuses on the impact of vertical educational diversity, we do not show these effects.

Table 4a: Impact of Vertical Educational Diversity on Extensive Margin of Innovation

Dependent Variable	R&D 0/1			New Product Innovation 0/1			Improved Product Innovation 0/1			Process Innovation 0/1		
Estimation	Logit	FE Logit	IV	OLS	FE Logit	IV	OLS	FE Logit	IV	OLS	FE Logit	IV
Diversity Index	0.274***	1.478***	4.182*	0.278***	1.673***	4.994*	0.287***	1.732***	4.835*	0.167***	0.564	3.403*
	(0.051)	(0.564)	(2.202)	(0.054)	(0.524)	(2.749)	(0.054)	(0.528)	(2.667)	(0.055)	(0.479)	(2.013)
Diversity Index=0	0.131***	0.474	1.651*	0.121***	0.836**	1.958*	0.126***	0.938**	1.897*	-0.016	-0.222	1.241
	(0.036)	(0.459)	(0.864)	(0.039)	(0.379)	(1.077)	(0.039)	(0.385)	(1.044)	(0.044)	(0.383)	(0.785)
Technological Potential	0.092***	0.187**	0.100***	0.077***	0.214***	0.058	0.078***	0.215***	0.061*	0.070***	0.131*	0.053*
	(0.009)	(0.090)	(0.030)	(0.010)	(0.080)	(0.036)	(0.010)	(0.080)	(0.035)	(0.010)	(0.074)	(0.028)
Incoming Spillovers	0.064***	0.028	-0.036	0.036**	-0.191	-0.101	0.037**	-0.165	-0.096	0.071***	0.141	-0.029
	(0.016)	(0.163)	(0.066)	(0.016)	(0.151)	(0.084)	(0.016)	(0.151)	(0.081)	(0.017)	(0.135)	(0.062)
Deltasales	0.008	-0.007	0.007	0.019**	0.065	0.019**	0.021***	0.064	0.021**	-0.012	-0.076	-0.015*
	(0.007)	(0.069)	(0.008)	(0.008)	(0.062)	(0.010)	(0.008)	(0.062)	(0.010)	(0.008)	(0.057)	(0.009)
Appropriability	-0.039***	-0.128*	-0.038***	-0.025***	-0.045	-0.028***	-0.026***	-0.035	-0.028***	0.006	0.073	0.003
	(0.008)	(0.076)	(0.008)	(0.008)	(0.072)	(0.010)	(0.008)	(0.072)	(0.010)	(0.009)	(0.065)	(0.008)
Size	0.081***	0.244	0.086***	0.078***	0.418	0.092***	0.073***	0.461*	0.085***	0.059***	0.017	0.072***
	(0.013)	(0.290)	(0.020)	(0.014)	(0.270)	(0.023)	(0.015)	(0.274)	(0.024)	(0.015)	(0.251)	(0.018)
$Size^2$	-0.002*	0.000	-0.006*	-0.002	-0.016	-0.008*	-0.002	-0.019	-0.007	0.001	0.023	-0.003
	(0.001)	(0.033)	(0.004)	(0.002)	(0.032)	(0.005)	(0.002)	(0.032)	(0.005)	(0.002)	(0.029)	(0.003)
Share Trained	-0.028	0.961**		0.035	0.530		0.042	0.590*		-0.040	0.694**	
	(0.031)	(0.384)		(0.033)	(0.357)		(0.033)	(0.358)		(0.033)	(0.324)	
Share Advanced	0.118***	-0.323		0.156***	-0.029		0.156***	-0.013		0.032	0.761*	
	(0.044)	(0.532)		(0.047)	(0.491)		(0.047)	(0.491)		(0.049)	(0.456)	
Share Acad	0.372***	1.771**		0.219***	0.594		0.205***	0.503		0.084	1.071	
	(0.053)	(0.765)		(0.059)	(0.714)		(0.059)	(0.713)		(0.057)	(0.664)	
Region Share Trained			0.151			0.489***			0.478***			0.477***
			(0.129)			(0.148)			(0.145)			(0.131)
Region Share Advanced			0.017			-0.036			-0.028			-0.150*
			(0.075)			(0.090)			(0.087)			(0.077)
Region Share Acad			-0.027**			-0.039***			-0.043***			0.002
			(0.013)			(0.014)			(0.014)			(0.019)
N	13142	3890	13143	13136	4610	13143	13136	4605	13143	13142	5639	13143
Kleibergen			14.958			14.958			14.958			14.958

Notes: The table displays marginal effects of logit, fixed effects logit and IV estimates that instrument vertical educational diversity by the regional average of vertical educational diversity in 1990. Standard errors in parantheses are clustered at firm level in the Logit and FE Logit estimates and block bootstrapped at the region level in the IV estimates. All continuous variables except Deltasales enter in logs. All regressions include time dummies. Logit estimates include 2-digit industry dummies and FE Logit and IV estimates include sector dummies. Testing for potential weak instruments, *Kleibergen* refers to the Kleibergen-Paap rk Wald F statistic, which has critical value of 16.38 for 10% maximal IV size. *,** and *** denote significances at the 10%, 5% and 1% level. See Table A2 in the Appendix for robustness checks of the IV estimates.

Table 4b: Impact of Vertical Educational Diversity on Innovation Intensity

Dependent Variable	R&D Intensity			New Product Innovation Intensity			Improved Product Innovation Intensity			Process Innovation Intensity		
Estimation	OLS	FE	Region IV	OLS	FE	Region IV	OLS	FE	Region IV	OLS	FE	Region IV
Diversity Index	0.875***	0.121	4.603	-0.096	-0.625	2.757	-0.130	-0.787	3.637	-0.547**	0.368	6.585
	(0.324)	(0.666)	(10.924)	(0.314)	(0.653)	(15.407)	(0.314)	(0.672)	(23.343)	(0.228)	(0.854)	(43.750)
Diversity Index=0	0.279	-0.026	1.994	-0.037	0.072	1.115	0.184	0.004	1.690	-0.428**	0.159	2.593
	(0.248)	(0.559)	(4.446)	(0.285)	(0.669)	(6.280)	(0.309)	(0.750)	(9.499)	(0.188)	(0.689)	(18.884)
Technological Potential	0.379***	0.034	0.745***	0.368***	0.227**	0.519	0.361***	0.196*	0.504	0.097**	0.102	-0.027
	(0.058)	(0.099)	(0.233)	(0.057)	(0.100)	(0.320)	(0.057)	(0.109)	(0.428)	(0.043)	(0.144)	(1.309)
Incoming Spillovers	0.169*	0.143	0.309	0.012	-0.047	0.047	0.012	-0.120	0.021	0.058	-0.077	-0.119
	(0.091)	(0.154)	(0.354)	(0.097)	(0.181)	(0.603)	(0.097)	(0.177)	(0.691)	(0.071)	(0.191)	(1.546)
Deltasales	0.072*	0.050	0.092	-0.028	-0.077	-0.021	-0.020	-0.049	-0.018	0.050*	0.047	0.042
	(0.038)	(0.054)	(0.056)	(0.037)	(0.059)	(0.094)	(0.038)	(0.062)	(0.161)	(0.030)	(0.073)	(0.182)
Appropriability	0.103**	0.020	0.246***	0.065	-0.006	0.111*	0.057	-0.060	0.104*	0.037	0.091	0.074
	(0.041)	(0.060)	(0.057)	(0.043)	(0.070)	(0.061)	(0.043)	(0.071)	(0.058)	(0.035)	(0.089)	(0.213)
Size	-0.275***	-0.278	-0.540***	-0.129*	-0.316	-0.236***	-0.072	-0.256	-0.198*	-0.163***	-0.138	-0.270
	(0.079)	(0.347)	(0.081)	(0.075)	(0.408)	(0.085)	(0.073)	(0.406)	(0.115)	(0.063)	(0.504)	(0.366)
$Size^2$	0.020**	0.024	0.037***	0.011	0.028	0.015	0.003	0.021	0.008	0.009*	0.001	0.013
	(0.008)	(0.036)	(0.013)	(0.007)	(0.039)	(0.015)	(0.007)	(0.039)	(0.016)	(0.006)	(0.045)	(0.022)
Share Trained	-0.029	-0.162		0.383**	0.039		0.520***	0.081		-0.014	0.391	
	(0.170)	(0.397)		(0.174)	(0.415)		(0.177)	(0.387)		(0.132)	(0.486)	
Share Advanced	0.990***	0.122		1.669***	0.207		1.882***	0.399		0.225	0.044	
	(0.237)	(0.492)		(0.268)	(0.572)		(0.265)	(0.553)		(0.196)	(0.633)	
Share Acad	2.324***	0.117		1.446***	0.658		1.377***	0.755		0.692***	0.245	
	(0.287)	(0.601)		(0.335)	(0.789)		(0.336)	(0.797)		(0.230)	(1.010)	
Inverse Mills Ratio	0.039***	0.011	0.025	0.119***	0.144*	0.099	0.102***	0.117	0.076	0.044*	-0.082	-0.014
	(0.011)	(0.030)	(0.021)	(0.028)	(0.084)	(0.108)	(0.027)	(0.088)	(0.107)	(0.025)	(0.183)	(0.338)
Region Share Trained			-0.398			-0.294			-0.121			-0.008
			(0.631)			(0.603)			(1.313)			(1.832)
Region Share Advanced			0.415			0.326			0.323			-0.142
			(0.430)			(0.423)			(0.920)			(1.302)
Region Share Acad			0.073			-0.100			-0.121			-0.011
			(0.121)			(0.131)			(0.103)			(0.487)
N	5097	5097	5097	6278	6278	6284	6260	6260	6266	2035	2035	2035
Kleibergen			13.144			24.877			23.967			4.997

Notes: The table displays coefficients of OLS, fixed effects OLS and IV estimates that instrument vertical educational diversity by the regional average of vertical educational diversity in 1990. Standard errors in parantheses are bootstrapped in the OLS estimates, bootstrapped clustered at the firm level in the FE OLS estimates and block bootstrapped at the region level in the IV estimates. All continuous variables except Deltasales enter in logs. All regressions include time dummies. OLS estimates include 2-digit industry dummies and FE and IV estimates include sector dummies. Testing for potential weak instruments, *Kleibergen* refers to the Kleibergen-Paap rk Wald F statistic, which has critical value of 16.38 for 10% maximal IV size. *,** and *** denote significances at the 10%, 5% and 1% level. See Table A3 in the Appendix for robustness checks of the IV estimates.

Table 4b presents the estimation results for the intensive margins. Again, we present three types of estimations (OLS, FE, IV) each for R&D intensity, new product innovation intensity, improved product innovation intensity, and process innovation intensity. The results provide little evidence that vertical educational diversity increases the intensive margin of innovation performance. Hence, the findings do not support hypothesis 1b and hypothesis 2b, which suggest that vertical educational diversity increases R&D intensity and the intensity of new product innovation. Since hypotheses 3b and 4b claim an insignificant relationship, we confirm these two hypotheses.

The effects of the control variables in the intensive margin equation are generally as expected, though they often remain insignificant, possibly due to a lack of variation across time. Concretely, Technological Potential is positive for innovation performance in all estimates but remains insignificant in the fixed effects estimations for R&D intensity and process innovation intensity. While Appropriability becomes insignificant in the fixed effects estimates, it has a positive relationship with innovation performance in the OLS and IV estimations. If we remember the significant negative sign for Appropriability for the extensive margin, we see exactly the opposite for the intensive margin. This confirms the idea that high Appropriability constitutes an entry barrier to the innovation market, and once entered it is positively related with the innovation performance of the incumbents, since their innovation results are well protected. Incoming Spillovers are positively related with the intensive margin of the dependent variables and we observe a non-linear relationship between firm size and the intensive margins. Education shares generally have the expected positive sign, though not all estimates display significance.

The overall picture of our results is in line with the theoretical framework presented in this paper. Concretely, the framework assumes that commercialization character becomes more prevalent when moving from the extensive margin to the intensive margin and when moving from R&D to new product innovation to improved product innovation to process innovation. The empirical results draw a picture that is coherent with this framework as the value of vertical educational diversity decreases along both of these dimensions.

In sum the robustness tests confirm the positive relationship between vertical educational diversity and the extensive margins of R&D, New Product Innovation and Improved Product Innovation. However, the corresponding intensive margins as well as process innovation remain unaffected.

6. Policy implications

Our results create quite a challenging situation for education policy makers. On one hand, a highly qualified workforce tends to have a positive impact on the innovation performance of a firm. On the other hand, an education system that produces a workforce with similar educational backgrounds—too many people with a conventional university tertiary education—it will ultimately decreases the diversity of the workforce and the creative potential of the economy is likely to decrease. Consequently, an education system should be flexible enough to generate a workforce that provides a sufficient amount of people with different educational backgrounds. Skills gap analysis of different institutions gives us some evidence that this is indeed a challenge in many countries (i.e. Deloitte and the Manufacturing Institute, 2011).

The results at hand are based on Swiss data and should be interpreted against the background of the Swiss education system, which has some unique features that contribute to the vertical educational diversity of the country. This enables firms to have a diverse workforce, which increases the propensity of innovation activities. And indeed, the fraction of Swiss firms with innovative activities is one of the highest in the world.

The following paragraphs discuss features of the Swiss education system that contribute to the positive effect of vertical educational diversity in Switzerland. First and foremost, the principle of horizontal and vertical permeability (transition mechanisms) in the whole education system is the basis for lifelong learning where no dead-ends exist throughout a person's entire lifespan—no matter their educational starting points or objectives. This is particularly important in an economy confronted with substantial technological change where employees may need to retool their skills several times. The high proportion of employees belonging to our Share Trained means that Switzerland's work force depends heavily on the upper secondary education level, which is itself strongly linked to the needs of the labor market. In each cohort, around two out of three youngsters chose an apprenticeship as their first entry into the labor market instead of following the purely academic path to university education.

Second, employers are in the driver's seat when it comes to the definition of the vocational framework curriculum. Training companies are members of professional organizations, which are heavily involved in curriculum design as well as providing training material for companies. The labor market relevance of curricula not only benefits the students themselves, but contributes to technological spillovers from firms at the technological frontier to firms operating below the technological frontier (Rupietta and Backes-Gellner, 2012).

Third, regarding the application of that curriculum, students' doing a high percentage of their training at the workplace is a big advantage because apprentices or students at the tertiary education level learn on the best-available technologies used in firms while picking up all the softs and behavioral skills required for being members of high-performing teams (see, e.g., Bolli and Renold, 2015). In addition, a high percentage of workplace learning guarantees that knowledge and skills are encultured and encoded (Lewis, 2005), and that young people are exposed to unfamiliar and unexpected situations for a substantial part of their education time. They are often embedded in working teams that are actually developing innovations. This improves their problem-solving skills and their creativity, which are prerequisites for innovation performance.

Furthermore, Swiss firms that train apprentices make a profit on average from training, which might explain their high motivation to invest in the human capital of their employees. In Switzerland, the benefits of apprenticeship training outweigh the corresponding costs mostly because the wages of apprentices are low compared to what a skilled worker would earn, and because companies make sure their apprentices are productive by the second or third year of their apprenticeship by using occupation-specific syllabi from the professional associations (Dionisius et al., 2009, Wolter et al., 2006). This effect is also beneficial to the government and reduces public expenditures for vocational education.

In sum, one has to understand the particularity of the Swiss vocational and professional education system, and the specificities behind the employer-driven approach of the Swiss system in order to understand why and under which premises vertical educational diversity provides such a comparative advantage for firms in Switzerland.

7. Conclusion

This paper provides an in-depth investigation of the relationship between the vertical educational diversity of employees and the innovation performance of firms. This is a relevant question since it emphasizes the importance not only of the mix of employees' educational backgrounds, but also of the combination of different qualifications for innovation. Based on a comprehensive dataset covering the period from 1999 – 2011 for a representative sample of Swiss firms and the application of econometric methods, we determine that there is a significant relationship between vertical educational diversity and innovation performance during different stages of the innovation process. Vertical educational diversity is significantly and positively related to the extensive margins of R&D, new product innovation, and improved product innovation. However, vertical educational diversity is insignificantly related with process innovation and the intensive margins of R&D and product innovations. We conducted a number of robustness tests addressing the potential endogeneity of vertical educational diversity as observed in firms.

These results point to the favorable effect of vertical educational diversity for the generation of new knowledge or the invention of new products. The results also show that vertical educational diversity does not positively affect the commercial success of innovation or R&D activities. Actually, it tends to hinder firms in their efforts to efficiently exploit their innovative findings for commercial purposes. This poses challenges not only on the level of firms, but also to the education system of a country.

On the firm level, these results clearly challenge management systems. They should be aware of the likely adverse effects of increased vertical educational diversity on the efficiency of turning R&D results into commercially successful products. That is clearly a very difficult task, since it also requires them to consider the diversity implications of a change in the education level of employees. An increase in overall education level usually is positively related with the market success of innovative products. However, if it also increases the vertical educational diversity, an indirect negative effect on innovation success can be expected. This requires a change in perspective away from a focus on single educational categories towards a perspective considering the "mix of degrees" and a more combinatorial view of educational outcomes.

There are also implications for the education system of a country. It should provide an "optimal" mix of educational backgrounds to increase the flexibility of firms to create their "optimal degree mix." This requires an attractive institutional environment and a positive attitude of the population as well as of companies towards education activities including vocational education. Moreover, job market prospects should be sufficiently good in order to increases peoples' readiness to pay for formal education. To investigate the institutional setting of the education system in Switzerland is not part of the study at hand. However, it would be an interesting topic for future research to analyze the interactions between the output of the education system and the demand for a diverse workforce by industry. We also leave in-depth investigation of the complementarity of different formal education levels and the meaning of commuting between regions in Switzerland or migration from other countries for future research activities. The influence of international labor mobility on the diversity of the national workforce could also be part of future investigations.

Literature

Allen, J., and R. K. W. van der Velden (2011): The Flexible Professional in the Knowledge Society New Challenges for Higher Education. Dordrecht; New York: Springer.

Alesina, A. and E. La Ferrara (2005): Ethnic Diversity and Economics Performance*, Journal of Economic Literature*, XLIII, 762-780.

Arora, A., Gambardella, A., Magazzini, L. and F. Pammolli (2009): A breath of fresh air? Firm type, scale, scope, and selection effects in drug development. *Management Science*, 55(10), pp. 1638-1653.

Atuahene-Gima, K. and F. Evangelista (2000): Cross-Functional influence in new product development: an exploratory study of marketing and R&D perspectives, *Management Science*, 46, 1269-1284.

Bolli, T. and U. Renold (2015): "Comparative Advantages of School and Workplace Environment in Competence Acquisition - Empirical Evidence From a Survey Among Professional Tertiary Education and Training Students in Switzerland", *KOF Working Paper Series*, 389.

Bolli, T. and M. Woerter (2013): Technological Diversification and Innovation Performance, *KOF-Working Paper* No. 336, Zurich.

Carlile, P.R. (2002): A Pragmatic View of Knowledge and Boundaries; Boundary Objects in New Product Development, *Organization Science*, 13, 442-455.

Cohen, W.M. (2010): Fifty Years of Empirical Studies on Innovative Activity and Performance, in: Hall, B.H., Rosenberg, N. (Eds.) Handbook of the Economics of Innovation, pp. 129-213, Elsevier

Cohen, W.M. and D.A. Levinthal (1989): Innovation and Learning: The Two Faces of R&D, *Economic Journal*, 99, 569-596.

Cohen, W.M. and D.A. Levinthal (1990): Absorptive Capacity: A New Perspective on Learning and Innovation, *Administrative Science Quarterly*, 35, 128-152.

Dahlin, K.B., L.R. Weingart and P.J. Hinds (2005): Team Diversity and Information Use, *Academy of Management Journal*, 48, 1107-1123.

Danzon, P.M., Nicholson, S. and N.S. Pereira (2005): Productivity in pharmaceutical–biotechnology R&D: the role of experience and alliances, *Journal of Health Economics,* 24, 317–339.

Dawson, J. (2011): Measurement of work group diversity, Doctoral thesis, Aston University.

Deloitte and the Manufacturing Institute (2011): Boiling point? The skills gap in U.S. http://www.themanufacturinginstitute.org/~/media/A07730B2A798437D98501E798C2E13AA.ashx.

Dionisius, R., Muehlemann, S., Pfeifer, H., Walden, G., Wenzelmann, F., and S.C. Wolter (2008): Cost and Benefit of Apprenticeship Training-A Comparison of Germany and Switzerland.

Faems, D., and A.M. Subramanian (2013): R&D Manpower and Technological Performance: The Impact of Demographic and Task-Related Diversity. *Research Policy*, 42(9), 1624–33

Garcia-Vega, M. (2006): Does Technological Diversification Promote Innovation? An Empirical Analysis for European Firms, *Research Policy*, 35, 230-246.

Harrison, D., and K. Klein (2007): What's the difference? Diversity constructs as separation, variety, or disparity in organizations, *The Academy of Management Review*, 32 (4), 1199–1228.

Hoffman, N. and R. Schwartz (2015): "Gold Standard: The Swiss Vocational Education and Training System". Washington, DC: National Center on Education and the Economy.

Hong, L., and S.E. Page (1998): Diversity and Optimality, Santa Fe Institute Working Paper 98-08-077.

Horwitz, S. (2005): The compositional impact of team diversity on performance: theoretical considerations, *Human Resource Development Review*, 4, 219–245.

Horwitz, S. and I. Horwitz, (2007): The effects ofteamdiversityonteamoutcomes: a metaanalytic review of team demography, *Journal of Management*, 6, 987–1015.

Hülsheger, U.R., N. Anderson, and J.F. Salgado. (2009): "Team-level predictors of innovation at work: a comprehensive meta-analysis spanning three decades of research." *Journal of Applied Psychology,* 94.5, 1128.

Jacobsson, S. and C. Oskarsson (1995): Educational Statistics as an Indicator of Technological Activity, *Research Policy*, 24, 137-136.

Jaffe, A. B. (1986): Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value. *The American Economic Review*, 76(5): 984–1001.

Joshi, A. and Jackson, S. (2003): Managing workforce diversity to enhance cooperation in organizations. In: West, M., Tjosvold, D., Smith, K. (Eds.), International Handbook of Organizational Teamwork and Cooperative Working. John Willey & Sons Ltd.

Laurson, K. (2012): Keep Searching and you'll Find: What Do We Know About Variety Creation Through Firms' Search Activities For Innovation, *Industrial and Corporate Change*, 21, 1181-1220.

Leten, B., R. Belderbos and B.V. Looy (2007): Technological Diversification, Coherence and Performance of Firms, *Journal of Product Innovation Management*, 24, 567-579.

Livingstone, D. W. (2009): Education & Jobs : Exploring the Gaps. Toronto: University of Toronto Press.

Lewis, T. (2005): At the Interface of School and Work. *Journal of Philosophy of Education*, 39, 421-441.

Mansfield, E. and S. Wagner (1975): Organizational and Strategic Factors Associated with Probabilities of Success in Industrial R&D, *Journal of Business*, 48, 179-198.

March, J.G. (1991): Exploration and Exploitation in Organizational Learning, *Organization Science*, Vol. 2(1), 71-87.

Markus, A. and H.C. Kongsted (2012): How Does R&D Worker Recruitment Affect Technological Exploration? A Longitudinal Study of R&D Active Firms in Denmark, DRUID,

Markus, A., and Kongsted, H. C. (2013): It All Starts with Education: R&D Worker Hiring, Educational Background and Firm Exploration. In *Academy of Management Proceedings* (Vol. 2013, No. 1, p. 14296). Academy of Management.

McGuirk, H. and D. Jordan (2012): Local Labour Market Diversity and Business Innovation: Evidence from Irish Manufacturing Businesses, *European Planning Studies*, 20, 1945-1960.

Meuer, J., Rupietta, C. and U. Backes-Gellner (2015): Layers of Co-existing Innovation Systems, *Research Policy*, 44, 888-910.

Nehring, K., and C. Puppe (2002): A Theory of Diversity. *Econometrica* 70 (3), 1155-1198

OECD (2010): Learning for Jobs. Paris: OECD Publishing.

Østergaard, C.R., B. Timmermans and K. Kristinsson (2011): Does a Different View Create Something New? The Effect of Employee Diversity on Innovation, *Research Policy*, 40, 500-509.

Parrotta, P., D. Pozzoli and M. Pytlikova (2012): Does Labor Diversity Affect Firm Productivity?, *IZA Discussion Paper*, No. 6973.

Parrotta, P., D. Pozzoli and M. Pytlikova (2014): The Nexus between Labor Diversity and Firm's Innovation, *Journal of Population Economics,* 27, 303-364.

Plotnikova, T. (2010): Success in Pharmaceutical Research: The Changing Role of Scale and Scope Economies, Spillovers and Competition, *Jena Economic Research Papers*, 45.

Quintana-García, C., and C.A. Benavides-Velasco (2008): Innovative competence, exploration and exploitation: The influence of technological diversification. *Research Policy*, *37*(3), 492-507.

Rupietta, C. and U. Backes-Gellner, (2012): High quality workplace training and innovation in highly developed countries. *Leading House Working Paper*, 74.

Subramanian, A.M., Choi, Y.R., Lee, S-H., and C-C. Hang (2015): Linking Technological and Educational Level Diversities to Innovation Performance, *Journal of Technology Transfer,* online-version, 1-23.

Sala-i-Martin, X., K. Schwab, and World Economic Forum (2011): The Global Competitiveness Report 2011-2012. Geneva: World Economic Forum.

Soellner, R. (2010) Human Capital Diversity and Product Innovations: A Micro-Level Analysis, *Jena Economic Research Papers*, No. 2010,027.

Stasser, G., and W. Titus (1985): Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of personality and social psychology*, *48*(6), 1467.

Stirling, A. (1998): On the Economics and Analysis of Diversity. SPRU Electronic Working Paper Series, Paper No. 28, October

Stirling, A. (2004): Diverse Design: Fostering Technological Diversity in Innovation for Sustainability. Paper Submitted to Colloquium DEA, Durham Business School and Cranfield School of Management on: The Role of Diversity in Social Systems and its Relation to Innovation. Bologna, 12-13 July

Stirling, A. (2007): A General Framework for Analysing Diversity in Science, Technology and Society. *Journal of the Royal Society Interface*, 4: 707-719

van Dijk, H., van Engen, M. L., and D. van Knippenberg (2012): Defying conventional wisdom: A meta-analytical examination of the differences between demographic and job-related diversity relationships with performance. *Organizational Behavior and Human Decision Processes*, *119*(1), 38-53.

Weitzman, M.L. (1998): Recombinant Growth, *Quarterly Journal of Economics*, 113, 331-360.

Williams, K. and C. O'Reilly (1998): Demography and diversity in organizations: a review of 40 years of research. In: Staw, B., Cummings, L. (Eds.), Research in Organizational Behavior. JAI Press, Greenwich, pp. 77–140.

Wittenbaum, G.M. and G. Stasser (1996): Management of Information in Small Groups, In J.L. Nye and A.M. Brower (Eds.), What's Social About Social Cognition? Research on Socially Shared Cognition in Small Groups, 3-28, Thousand Oaks, CA: Sage.

Woerter, M. (2009): Industry Diversity and its Impact on the Innovation Performance of Firms, *Journal of Evolutionary Economics*, 19, 675-700.

Appendix

Table A1: Summary Statistics of Regional Instrumental Variable by Canton

			Share Lower				Share Trained				Share Advanced				Share Acad				Diversity Index			
Canton	N	Area	Mean	Std.Dev	Min	Max	Mean	Std.Dev	Min	Max	Mean	Std.Dev	Min	Max	Mean	Std.Dev	Min	Max	Mean	Std. Dev	Min	Max
AG	974	3	0.33	0.04	0.26	0.42	0.54	0.02	0.48	0.59	0.09	0.01	0.07	0.11	0.04	0.01	0.02	0.06	0.58	0.01	0.57	0.60
AI	33	5	0.48	0.04	0.33	0.49	0.45	0.02	0.44	0.53	0.05	0.01	0.05	0.07	0.02	0.01	0.02	0.07	0.56	0.01	0.56	0.60
AR	93	5	0.40	0.01	0.35	0.41	0.49	0.01	0.49	0.54	0.07	0.00	0.07	0.08	0.03	0.00	0.02	0.04	0.59	0.00	0.57	0.59
BE	1598	2	0.32	0.06	0.25	0.48	0.56	0.03	0.43	0.59	0.08	0.01	0.05	0.10	0.04	0.02	0.01	0.07	0.57	0.01	0.55	0.62
BL	447	3	0.32	0.04	0.26	0.42	0.55	0.02	0.49	0.59	0.08	0.00	0.05	0.10	0.05	0.01	0.01	0.07	0.59	0.00	0.56	0.61
BS	372	3	0.37	0.02	0.25	0.52	0.49	0.02	0.41	0.58	0.07	0.01	0.05	0.09	0.07	0.01	0.02	0.07	0.61	0.01	0.56	0.61
FR	360	2	0.46	0.06	0.26	0.55	0.45	0.04	0.38	0.59	0.06	0.01	0.04	0.10	0.03	0.01	0.02	0.07	0.58	0.02	0.55	0.62
GE	520	1	0.38	0.01	0.29	0.43	0.42	0.01	0.42	0.56	0.09	0.00	0.06	0.11	0.12	0.01	0.04	0.12	0.66	0.01	0.59	0.66
GL	78	5	0.43	0.03	0.26	0.47	0.48	0.03	0.45	0.59	0.07	0.00	0.06	0.10	0.02	0.00	0.02	0.04	0.57	0.00	0.57	0.57
GR	350	5	0.37	0.05	0.28	0.53	0.54	0.04	0.42	0.63	0.06	0.01	0.03	0.09	0.03	0.01	0.02	0.07	0.57	0.01	0.52	0.61
JU	156	2	0.52	0.01	0.44	0.52	0.41	0.01	0.41	0.47	0.05	0.00	0.05	0.06	0.03	0.00	0.03	0.03	0.56	0.00	0.56	0.58
LU	635	6	0.37	0.05	0.31	0.53	0.52	0.03	0.42	0.55	0.08	0.01	0.04	0.10	0.03	0.01	0.01	0.07	0.58	0.01	0.54	0.60
NE	301	2	0.43	0.04	0.38	0.51	0.45	0.02	0.42	0.46	0.07	0.01	0.05	0.09	0.05	0.02	0.03	0.12	0.60	0.02	0.56	0.66
NW	85	6	0.34	0.03	0.25	0.38	0.54	0.01	0.53	0.58	0.08	0.00	0.07	0.09	0.03	0.01	0.03	0.07	0.58	0.00	0.57	0.58
OW	53	6	0.42	0.03	0.31	0.43	0.49	0.02	0.49	0.54	0.07	0.01	0.06	0.10	0.03	0.01	0.02	0.05	0.58	0.00	0.58	0.60
SG	1066	5	0.38	0.04	0.26	0.47	0.52	0.03	0.45	0.59	0.07	0.01	0.06	0.10	0.03	0.01	0.02	0.07	0.58	0.00	0.57	0.60
SH	164	5	0.32	0.01	0.26	0.43	0.56	0.01	0.47	0.59	0.09	0.00	0.06	0.10	0.04	0.00	0.03	0.07	0.58	0.01	0.57	0.60
SO	489	2	0.37	0.02	0.25	0.42	0.53	0.01	0.49	0.58	0.07	0.01	0.05	0.09	0.03	0.01	0.01	0.07	0.58	0.01	0.56	0.61
SZ	222	6	0.42	0.04	0.27	0.48	0.49	0.02	0.45	0.56	0.07	0.01	0.05	0.11	0.03	0.00	0.02	0.06	0.58	0.01	0.57	0.60
TG	458	5	0.36	0.02	0.26	0.38	0.53	0.01	0.52	0.59	0.08	0.00	0.07	0.10	0.03	0.01	0.02	0.05	0.58	0.00	0.57	0.60
TI	692	7	0.42	0.04	0.25	0.54	0.47	0.02	0.42	0.58	0.06	0.01	0.03	0.09	0.05	0.01	0.02	0.07	0.59	0.02	0.54	0.60
UR	45	6	0.48	0.02	0.35	0.49	0.44	0.02	0.44	0.54	0.06	0.00	0.06	0.08	0.02	0.00	0.02	0.04	0.57	0.00	0.57	0.58
VD	766	1	0.39	0.05	0.25	0.54	0.46	0.02	0.39	0.59	0.08	0.01	0.05	0.11	0.07	0.03	0.02	0.12	0.62	0.02	0.56	0.66
VS	398	1	0.47	0.03	0.35	0.52	0.44	0.03	0.41	0.54	0.06	0.01	0.04	0.09	0.03	0.01	0.02	0.12	0.58	0.01	0.55	0.66
ZG	246	6	0.31	0.01	0.27	0.44	0.54	0.01	0.48	0.57	0.10	0.00	0.06	0.10	0.05	0.00	0.02	0.06	0.60	0.00	0.57	0.60
ZH	2542	4	0.29	0.03	0.23	0.49	0.56	0.03	0.42	0.59	0.09	0.01	0.06	0.11	0.05	0.02	0.02	0.12	0.58	0.02	0.57	0.66

Table A2: Robustness of IV Estimates: Extensive Margin

Dependent Variable	R&D 0/1					New Product Innovation 0/1				
	Baseline	Within-Industry	Within-Area	Only Old Firms	DepVar Average	Baseline	Within-Industry	Within-Area	Only Old Firms	DepVar Average
Diversity Index	4.182*	4.404*	4.336*	4.538**	2.480*	4.994*	4.997*	4.850*	5.191	3.105
	(2.202)	(2.274)	(2.430)	(2.310)	(1.424)	(2.749)	(2.974)	(2.935)	(3.274)	(1.903)
Diversity Index=0	1.651*	1.718*	1.712*	1.782**	0.983*	1.958*	1.942*	1.899*	2.071	1.217
	(0.864)	(0.877)	(0.950)	(0.906)	(0.557)	(1.077)	(1.150)	(1.149)	(1.281)	(0.745)
Technological Potential	0.100***	0.072***	0.096***	0.091***	0.119***	0.058	0.051**	0.059	0.047	0.081***
	(0.030)	(0.019)	(0.033)	(0.031)	(0.020)	(0.036)	(0.024)	(0.039)	(0.042)	(0.026)
Incoming Spillovers	-0.036	-0.036	-0.041	-0.063	0.016	-0.101	-0.087	-0.098	-0.118	-0.045
	(0.066)	(0.058)	(0.073)	(0.075)	(0.046)	(0.084)	(0.078)	(0.090)	(0.105)	(0.059)
Deltasales	0.007	0.004	0.007	0.003	0.009	0.019**	0.014	0.019*	0.017	0.021**
	(0.008)	(0.008)	(0.009)	(0.009)	(0.008)	(0.010)	(0.010)	(0.010)	(0.011)	(0.009)
Appropriability	-0.038***	-0.035***	-0.038***	-0.038***	-0.036***	-0.028***	-0.023**	-0.028***	-0.025**	-0.027***
	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)	(0.010)	(0.010)	(0.010)	(0.011)	(0.010)
Size	0.086***	0.090***	0.086***	0.122***	0.074***	0.092***	0.092***	0.090***	0.131***	0.080***
	(0.020)	(0.017)	(0.021)	(0.028)	(0.017)	(0.023)	(0.022)	(0.023)	(0.035)	(0.018)
$Size^2$	-0.006*	-0.007**	-0.006	-0.010**	-0.004	-0.008*	-0.008	-0.007	-0.012*	-0.005
	(0.004)	(0.003)	(0.004)	(0.005)	(0.003)	(0.005)	(0.005)	(0.005)	(0.006)	(0.003)
Region Share Trained	0.151	0.242	0.065	0.153	0.093	0.489***	0.582***	0.317*	0.491***	0.310***
	(0.129)	(0.169)	(0.157)	(0.136)	(0.089)	(0.148)	(0.219)	(0.175)	(0.164)	(0.117)
Region Share Advanced	0.017	-0.035	0.079	0.021	-0.008	-0.036	-0.084	0.032	-0.021	-0.050
	(0.075)	(0.076)	(0.076)	(0.079)	(0.051)	(0.090)	(0.104)	(0.079)	(0.100)	(0.058)
Region Share Acad	-0.027**	-0.008	-0.045**	-0.031**	0.004	-0.039***	-0.013	-0.050**	-0.046***	-0.007
	(0.013)	(0.015)	(0.021)	(0.013)	(0.011)	(0.014)	(0.020)	(0.022)	(0.015)	(0.014)
Region Average DepVar					0.340***					0.366***
					(0.053)					(0.087)
N	13143	13142	13143	11307	13143	13143	13136	13143	11307	13143
Kleibergen	14.958	10.673	12.305	9.960	14.532	14.958	10.673	12.305	9.960	14.532

Dependent Variable	Improved Product Innovation 0/1					Process Innovation 0/1				
Estimation	Baseline	Within-Industry	Within-Area	Only Old Firms	DepVar Average	Baseline	Within-Industry	Within-Area	Only Old Firms	DepVar Average
Diversity Index	4.835*	4.888	4.767	5.038	3.032	3.403*	2.934	3.222	2.423	2.173
	(2.667)	(3.219)	(2.942)	(3.189)	(1.874)	(2.013)	(2.059)	(2.370)	(2.170)	(1.348)
Diversity Index=0	1.897*	1.902	1.868	2.014	1.190	1.241	1.054	1.167	0.883	0.757
	(1.044)	(1.243)	(1.153)	(1.248)	(0.732)	(0.785)	(0.791)	(0.921)	(0.851)	(0.527)
Technological Potential	0.061*	0.053**	0.061	0.049	0.083***	0.053*	0.054***	0.054*	0.060**	0.068***
	(0.035)	(0.026)	(0.037)	(0.040)	(0.025)	(0.028)	(0.019)	(0.032)	(0.030)	(0.021)
Incoming Spillovers	-0.096	-0.083	-0.095	-0.113	-0.042	-0.029	-0.005	-0.026	-0.008	0.007
	(0.081)	(0.083)	(0.090)	(0.102)	(0.057)	(0.062)	(0.055)	(0.075)	(0.068)	(0.044)
Deltasales	0.021**	0.017	0.020**	0.018	0.022**	-0.015*	-0.015*	-0.015*	-0.018**	-0.014*
	(0.010)	(0.010)	(0.010)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Appropriability	-0.028***	-0.023**	-0.029***	-0.025**	-0.028***	0.003	0.006	0.003	0.006	0.003
	(0.010)	(0.010)	(0.010)	(0.011)	(0.010)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)
Size	0.085***	0.086***	0.084***	0.124***	0.074***	0.072***	0.069***	0.069***	0.091***	0.064***
	(0.024)	(0.023)	(0.023)	(0.035)	(0.019)	(0.018)	(0.017)	(0.020)	(0.024)	(0.015)
$Size^2$	-0.007	-0.007	-0.007	-0.011*	-0.004	-0.003	-0.003	-0.003	-0.004	-0.002
	(0.005)	(0.005)	(0.005)	(0.006)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.002)
Region Share Trained	0.478***	0.571**	0.306*	0.485***	0.306***	0.477***	0.502***	0.273**	0.417***	0.284***
	(0.145)	(0.232)	(0.179)	(0.156)	(0.113)	(0.131)	(0.168)	(0.127)	(0.127)	(0.091)
Region Share Advanced	-0.028	-0.075	0.037	-0.019	-0.045	-0.150*	-0.148*	-0.087	-0.112	-0.108**
	(0.087)	(0.107)	(0.082)	(0.096)	(0.056)	(0.077)	(0.082)	(0.066)	(0.075)	(0.051)
Region Share Acad	-0.043***	-0.018	-0.053**	-0.049***	-0.010	0.002	0.004	-0.001	-0.000	0.008
	(0.014)	(0.021)	(0.021)	(0.015)	(0.014)	(0.019)	(0.019)	(0.025)	(0.017)	(0.013)
Region Average DepVar					0.362***					0.361***
					(0.086)					(0.067)
N	13143	13142	13143	11307	13143	13143	13136	13143	11307	13143
Kleibergen	14.958	10.673	12.305	9.960	14.532	14.958	10.673	12.305	9.960	14.532

Notes: Each block of the table shows the results for the extensive margin of one of the four innovation measures, i.e. R&D, new product innovation, improved product innovation and process innovation. The table displays marginal effects of IV estimates that instrument vertical educational diversity by the regional average of vertical educational diversity in 1990. Standard errors in parantheses are block bootstrapped at the region level. All regressions include time and sector dummies. *Within-Industry* and *Within-Area* estimations entail dummies for 2-digit industries and 7 large areas, respectively. *Only Old Firms* estimations exclude firms founded after 1990 from the sample. *DepVar Average* estimates control for the regional average of the innovation measure. Testing for potential weak instruments, *Kleibergen* refers to the Kleibergen-Paap rk Wald F statistic, which has critical value of 16.38 for 10% maximal IV size. *,** and *** denote significances at the 10%, 5% and 1% level.

Table A3: Robustness of IV Estimates: Intensive Margin

Dependent Variable	R&D Intensity					New Product Innovation Intensity				
	Baseline	Within-Industry	Within-Area	Only Old Firms	DepVar Average	Baseline	Within-Industry	Within-Area	Only Old Firms	DepVar Average
Diversity Index	4.603	4.391	7.417	2.069	2.712	2.757	2.645	3.226	-0.891	3.609
	(10.924)	(38.693)	(10.007)	(35.265)	(6.626)	(15.407)	(14.101)	(8.360)	(15.408)	(19.198)
Diversity Index=0	1.994	1.703	3.151	1.073	1.166	1.115	0.994	1.303	-0.366	1.459
	(4.446)	(15.925)	(4.054)	(14.728)	(2.705)	(6.280)	(5.623)	(3.356)	(6.199)	(7.820)
Technological Potential	0.745***	0.387	0.682***	0.810	0.743***	0.519	0.370*	0.511***	0.580*	0.484
	(0.233)	(0.500)	(0.206)	(0.677)	(0.155)	(0.320)	(0.210)	(0.198)	(0.334)	(0.376)
Incoming Spillovers	0.309	0.166	0.223	0.359	0.361*	0.047	0.009	0.030	0.173	0.030
	(0.354)	(0.908)	(0.346)	(0.906)	(0.215)	(0.603)	(0.372)	(0.302)	(0.562)	(0.687)
Deltasales	0.092	0.074	0.082	0.090	0.099**	-0.021	-0.033	-0.023	-0.015	-0.027
	(0.056)	(0.184)	(0.052)	(0.110)	(0.044)	(0.094)	(0.072)	(0.066)	(0.117)	(0.119)
Appropriability	0.246***	0.130	0.246***	0.252***	0.245***	0.111*	0.086	0.110**	0.097	0.116**
	(0.057)	(0.160)	(0.050)	(0.070)	(0.050)	(0.061)	(0.069)	(0.054)	(0.065)	(0.059)
Size	-0.540***	-0.410	-0.536***	-0.488***	-0.553***	-0.236***	-0.244***	-0.233***	-0.098	-0.251***
	(0.081)	(0.258)	(0.081)	(0.142)	(0.072)	(0.085)	(0.086)	(0.088)	(0.096)	(0.081)
$Size^2$	0.037***	0.029	0.035***	0.036	0.040***	0.015	0.019	0.015	0.007	0.016
	(0.013)	(0.050)	(0.011)	(0.035)	(0.009)	(0.015)	(0.015)	(0.010)	(0.019)	(0.021)
Inverse Mills Ratio	0.025	0.026	0.022	0.024	0.022	0.099	0.084	0.100	0.115	0.092
	(0.021)	(0.062)	(0.019)	(0.037)	(0.015)	(0.108)	(0.080)	(0.064)	(0.103)	(0.137)
Region Share Trained	-0.398	-0.461	-0.970	-0.605	0.251	-0.294	-0.138	-0.251	-0.822	0.472
	(0.631)	(2.558)	(0.711)	(0.868)	(0.328)	(0.603)	(0.848)	(0.708)	(0.615)	(0.469)
Region Share Advanced	0.415	0.303	0.535	0.488	-0.121	0.326	0.167	0.305	0.640	-0.134
	(0.430)	(1.084)	(0.410)	(1.016)	(0.198)	(0.423)	(0.407)	(0.324)	(0.476)	(0.415)
Region Share Acad	0.073	-0.046	0.049	0.090	0.166**	-0.100	-0.060	-0.087	-0.140	0.060
	(0.121)	(0.134)	(0.159)	(0.481)	(0.075)	(0.131)	(0.071)	(0.137)	(0.103)	(0.168)
Region Average DepVar					0.626***					0.621***
					(0.090)					(0.153)
N	5097	5097	5097	4422	5097	6284	6278	6284	5443	6284
Kleibergen	13.144	11.376	13.502	11.106	12.643	24.877	21.955	23.076	21.067	23.800
Dependent Variable	Improved Product Innovation Intensity					Process Innovation Intensity				
Estimation	Baseline	Within-Industry	Within-Area	Only Old Firms	DepVar Average	Baseline	Within-Industry	Within-Area	Only Old Firms	DepVar Average
Diversity Index	3.637	3.259	4.922	0.755	3.917	6.585	6.141	7.004	3.060	1.728
	(23.343)	(24.253)	(14.825)	(28.196)	(8.662)	(43.750)	(44.834)	(45.700)	(77.842)	(13.087)
Diversity Index=0	1.690	1.469	2.208	0.653	1.797	2.593	2.390	2.747	1.167	0.457
	(9.499)	(9.652)	(6.010)	(11.230)	(3.528)	(18.884)	(19.581)	(19.698)	(33.913)	(5.692)
Technological Potential	0.504	0.349	0.475*	0.552	0.482***	-0.027	-0.041	-0.034	0.050	0.124
	(0.428)	(0.382)	(0.255)	(0.495)	(0.169)	(1.309)	(1.157)	(1.348)	(2.541)	(0.419)
Incoming Spillovers	0.021	-0.010	-0.021	0.132	0.018	-0.119	-0.102	-0.130	-0.001	0.042
	(0.691)	(0.699)	(0.479)	(1.002)	(0.278)	(1.546)	(1.529)	(1.671)	(3.420)	(0.504)
Deltasales	-0.018	-0.030	-0.027	-0.036	-0.020	0.042	0.044	0.038	0.033	0.054
	(0.161)	(0.129)	(0.107)	(0.137)	(0.068)	(0.182)	(0.136)	(0.188)	(0.437)	(0.053)
Appropriability	0.104*	0.081	0.104*	0.097	0.109**	0.074	0.067	0.077	0.073	0.060
	(0.058)	(0.083)	(0.056)	(0.164)	(0.049)	(0.213)	(0.217)	(0.256)	(0.516)	(0.089)
Size	-0.198*	-0.204**	-0.198*	-0.037	-0.208**	-0.270	-0.251	-0.274	-0.198	-0.217
	(0.115)	(0.098)	(0.105)	(0.131)	(0.086)	(0.366)	(0.490)	(0.399)	(0.873)	(0.155)
$Size^2$	0.008	0.012	0.007	-0.002	0.009	0.013	0.012	0.013	0.011	0.012
	(0.016)	(0.018)	(0.018)	(0.025)	(0.009)	(0.022)	(0.026)	(0.027)	(0.036)	(0.008)
Inverse Mills Ratio	0.076	0.053	0.072	0.094	0.075	-0.014	-0.006	-0.011	0.017	0.026
	(0.107)	(0.154)	(0.079)	(0.233)	(0.046)	(0.338)	(0.399)	(0.346)	(0.583)	(0.136)
Region Share Trained	-0.121	-0.051	-0.315	-0.599	0.495	-0.008	0.003	-0.041	0.051	0.461
	(1.313)	(1.389)	(0.737)	(1.452)	(0.521)	(1.832)	(2.729)	(2.653)	(3.445)	(0.546)
Region Share Advanced	0.323	0.177	0.337	0.566	-0.121	-0.142	-0.133	-0.113	-0.111	-0.161
	(0.920)	(0.759)	(0.404)	(1.443)	(0.364)	(1.302)	(1.226)	(1.178)	(3.010)	(0.256)
Region Share Acad	-0.121	-0.081	-0.109	-0.142	0.040	-0.011	-0.011	-0.087	0.047	0.063
	(0.103)	(0.116)	(0.149)	(0.164)	(0.054)	(0.487)	(0.461)	(1.177)	(0.785)	(0.166)
Region Average DepVar					0.594***					0.809***
					(0.063)					(0.190)
N	6266	6260	6266	5420	6266	2035	2035	2035	1730	2035
Kleibergen	23.967	20.776	22.105	20.668	23.016	4.997	3.568	4.501	5.137	4.021

Notes: Each table block shows the intensive margin results of the four innovation measures, i.e. R&D, new product innovation,and improved product innovation and process innovation. The table displays coefficients of IV estimates that instrument vertical educational diversity by the regional average of vertical educational diversity in 1990. Standard errors in parantheses are block bootstrapped at the region level. All regressions include time and sector dummies. *Within-Industry* and *Within-Area* estimations entail dummies for 2-digit industries and 7 large areas, respectively. *Only Old Firms* estimations exclude firms founded after 1990 from the sample. *DepVar Average* estimates control for the regional average of the innovation measure. Testing for potential weak instruments, *Kleibergen* refers to the Kleibergen-Paap rk Wald F statistic, which has critical value of 16.38 for 10% maximal IV size. *,** and *** denote significances at the 10%, 5% and 1% level.