Paper to be presented at

DRUID15, Rome, June 15-17, 2015

(Coorganized with LUISS)

Collective knowledge processing and innovation

Christian Rupietta
University of Zurich
Department of Business Administration
christian.rupietta@uzh.ch

Uschi Backes-gellner
University of Zurich
Department of Business Administration
backes-gellner@business.uzh.ch

Abstract
The innovation output of countries like Germany and Switzerland has been remarkably high during the last decade. Both countries rely on a national education systems that contains academic and vocational education. According to latest studies on the relation of education type (academic vs. vocational) and innovation and growth, academic education provides workers with up-to-date knowledge that is required for innovation while vocational education is associated with very specific and outdated knowledge. Countries that achieve top positions in innovation rankings and heavily rely on vocational education are thus puzzling according to these studies. In this paper we investigate the function of institutions for knowledge processing in vocational education that is collectively organized such as in Germany and Switzerland. We identify institutions that collect, process and transfer innovation-relevant knowledge. Our model shows that collectively-organized vocational education couples two ideal types of innovation and learning: science, technology and innovation (STI) and doing, using and interacting (DUI) We test the implications of our model empirically and find that firms that participate in vocational education have a higher innovation output than firms not doing so. Our paper concludes with implications of our model for the design of national (vocational) education systems.

Jelcodes:I20,O31

Collective knowledge processing and innovation

I. Introduction

Germany and Switzerland have achieved top positions in innovation ranking over the past decade (Weissenberger-Eibl et al. 2011) and have both heavily relied on vocational education. While academic education is typically associated with innovation, vocational education is not (Aghion 2008; Aghion and Vandenbusche 2006). Krueger and Kumar (2004a,b) attribute a high specialization and a focus on old and established technology to vocational education and argue that it should not be positively associated with innovation and growth. This way of reasoning contradicts the finding that countries like Germany and Switzerland, which rely heavily on vocational education, are top innovators.

While researchers detected some underlying mechanisms that explain a negative impact of vocational education on innovation (Krueger and Kumar, 2004a,b), these explanations do neither focus on different forms of vocational education (e.g. collectively organized vs. company organized) nor on a comprehensive description of knowledge collection, processing and transfer within a vocational education system. As both, academic and vocational education provides employees with knowledge, understanding how new knowledge enters the system is an integral step to understand the impact of education on innovation. For academic education researchers often (implicitly) assume that this form of education contains institutions that provide students with innovation-relevant and up-to date knowledge. However, a systematic classification of such institutions has received little attention. A description of institutions within both academic and vocational education and the identification of institutions that are functional equivalents in terms of knowledge collection, processing and transfer across both educational types provides a more nuanced understanding on the contribution of both academic and vocational education to innovation.

In this paper we theoretically analyze institutions in academic education and collectively-organized vocational education such as in Germany and Switzerland and their implications for knowledge collection, processing and transfer. To identify institutions that affect the knowledge flow in both educational types, we use Jensen et al's (2007) model of learning. This model distinguishes two ideal-type learning modes, the science technology and innovation (STI) mode and the doing, using and interacting (DUI) mode. The STI mode mainly relies on the use and production of generalized and codified knowledge whereas the DUI mode relies on learning-by-doing, informal knowledge processing and localized knowledge. Although this model has its focus on the company as a level of analysis, it has strong implications for institutions such as those contained in education systems. After describing and categorizing knowledge flows in vocational education, we compare our categorization to knowledge flows in academic education and identify institutions that are functional equivalents. Moreover, we test the implications of our model empirically and provide first evidence on the effect of collectively organized vocational education on innovation in companies.

Collectively-organized vocational education like in Germany, Austria and Switzerland consist of a complex network of institutions that regulate knowledge flow. Vocational training is regulated by nation-wide accepted training curricula that ensure quality standards. These curricula regulate training in companies and vocational schools and are frequently updated in a collective effort by employer, employee and governmental organizations. Within this system the learning modes are located at different levels. The STI mode is located at the firm level and the intra-firm level, such as employer organizations. It describes best the updating process of training curricula in which firms codify their tacit knowledge on technology and best practices to transfer it to employer organizations. Employer organizations process and synthesize this knowledge, thereby selecting knowledge that is relevant for multiple firms and future technologies. This selection process is a collective effort of representatives of multiple firms. The DUI mode is located at the firm level and describes best the application of knowledge that training curricula contain and its absorption by firms. Close interactions of apprentices and instructors enable firms to absorb knowledge on new technologies that is part of the updated training curricula. These interactions allows apprentices to apply their knowledge from vocational schools and to generate new tacit knowledge in the company-specific context.

Institutions in academic education are functionally equivalent to those in collective-vocational education. Although the knowledge collection, processing and transfer are similar, differences occur when companies apply knowledge from academic education. In academic education curricula are a product of collective knowledge collection, processing and transfer. This curriculum development is decentralized in the sense that every university can develop its own curriculum. Curriculum development in vocational education on the contrary is centralized in the sense that only one curriculum exists for an occupation. Moreover curricula in vocational education also regulate training in companies. In academic education a strict separation between knowledge transfer to students and knowledge application by students exists. Thus the DUI mode of learning is not part of a training curriculum like in vocational education. Despite these differences, the identification of several institutions in academic and vocational education that are functional equivalents leads us to the hypothesis that collectively-organized vocational education has a positive effect on a company's innovation output.

For our empirical analyses we use the KOF (Swiss Economic Institute) Innovation Survey. This panel data set comprises two to three thousand observations per wave and is representative for the construction, manufacturing, and service sectors in Switzerland. It contains information on various innovation outcomes such as patent applications, process and product innovations, and general innovation, a joint measure that combines the two latter measures. To analyze the effect of vocational education on a company's innovation output we require these gradual innovation measures, because curricula contain knowledge that is not patentable. The availability of these innovation measures allows us to analyze which impact vocational education has on innovations with various degrees of novelty.

Our results show that vocational education is positively associated with all innovation measures. Results from an estimation that takes the endogeneity of a company's decision to train into account, show a positive and highly robust impact of vocational education on general innovation. These results have several theoretical and practical implications. By showing that vocational education has a positive impact on a company's innovation output, we provide a theoretical and empirical explanation that not only tertiary and academic qualifications positively influence innovation, but also do secondary vocational qualifications. Our model distinguishes collectively-organized vocational education from types that firms organize on their own. Our findings are consistent with our model. Lucas (2009) and Staley (2011) argue that a rapid diffusion of knowledge within a country results in higher frequency of innovations and thus higher growth. Due to the participation in vocational education a training company benefits from the knowledge that is diffused within the vocational education system and can be more innovative than companies not participating in this system.

II. National education systems an collective learning

National education systems diffuse knowledge to students that they later transfer to their employers. As knowledge is a key ingredient of innovation, National education systems are a main factor for a company' innovation output. A common classification distinguishes between academic and vocational education and associates distinct types of knowledge with these education types. Recent studies consider academic education as a main provider of up-to-date and innovation relevant knowledge. Vocational education on the contrary is associated with outdated and specific knowledge that helps single firms to maintain its productivity but not to improve its innovation output. These two classifications originate from a view that consideres vocational education to be mainly organized by a single company and neglects other forms of vocational education especially those that are organized on an intra-company level.

The difference between vocational education being organized by a single company and multiple companies (collectively-organized vocational education) is the potential to induce knowledge diffusion between companies. Knowledge diffusion between companies allows all companies in the system to exchange their knowledge and to generate new knowledge form these exchanges. In academic education knowledge exchange (from universities to companies) is a major reason for its strong and positive impact on innovation and growth. A vocational education system that enables companies to exchange their knowledge and establishes institutions that ensure the diffusion of up-to-date and innovation relevant knowledge might have a similar potential to improve innovation of companies and growth of countries.

A more nuanced understanding of the implications of collectively-organized vocational for innovation of companies requires an anlysis of mechanisms of knowledge exchange within the system. As knowledge in collectively-organized vocational education is processed at different level of aggregation (company and intra-company) an explanation of the internal mechanism must consider different forms of knowledge and learning. Jensen et. al (2007) introduce a model that distinguishes different types of knowledge (generalized and localized) and different modes of learning (DUI and STI). These distinctions are highly relevant for the description of knowledge exchanges in a vocational education system that is operated by multiple actors.

A. Institutional Background

Vocational education combines vocational schooling with workplace training in 3-4-year programs[2]. Similar to academic education, vocational bases on curricula. These curricula regulate both education in vocational schools and training at the workplace. Curricula are

[2] For the duration of the training program, apprentices have a fixed-term contract with a firm. During their training apprentices are fully integrated in the production process and perform productive tasks that are similar or identical to those of qualified workers.

nation-wide accepted and ensure the transferability of skills between companies (Wolter & Ryan, 2011). Thus the curricula contain general skills that are valuable for all companies.

As technology qualification requirements change over time, keeping training curricula up to date is a major task in such an education system. In collectively-organized vocational education employers' associations together with employee and governmental representatives (social partnership) decide on the contents of the curricula (Bundesversammlung der Schweizerischen Eidgenossenschaft, 2002; Wolter & Ryan, 2011). To ensure that the contents of the curricula meet actual skill demands, mainly the employer organization initiate revisions of curricula and integrate innovations in dual-track VET (Bauder & Osterwalder, 2008; Bundesversammlung der Schweizerischen Eidgenossenschaft, 2002; Der Schweizerische Bundesrat, 2003)[3]. Thereby, the social partners integrate best practices into the revision of curricula (Bauder & Osterwalder, 2008). The resulting curricula base on skill requirements multiple firms have and have a strong focus on competences that a qualified worker in a certain occupation must have (Der Schweizerische Bundesrat, 2003; Pedró, Burns, Ananiadou, & de Navacelle, 2009). In sum such a vocational education system regulates and updates curricula for 200-300 occupations.

B. Modes of innovation and learning

Jensen et al. (2007) developed a model for innovation that contains two ideal-type modes of innovation and learning: the Science, Technology and Innovation (STI) mode and the Doing, Using and Interacting (DUI) mode. These two modes differ with respect to the types of knowledge (Explicit vs. implicit knowledge and global vs. local knowledge) that is processed and the way how employees share these different types of knowledge (codification vs. personal interaction. The distinction between different types of knowledge and knowledge transfer makes this model useful for the description of knowledge flows in collectively-organized vocational education. Although this model has the company as the level of analysis, the concepts are also valid for higher aggregation levels such as industries or nations (Jensen et al., 2007). Knowledge in collectively-organized vocational education is processed at different levels of aggregation. Therefore the categorization of the two learning modes are especially helpful for the explanation of the underlying mechanisms of knowledge processing in this environment.

[3] Future-orientation and thus innovation is a key component of the Swiss VET system Bundesversammlung der Schweizerischen Eidgenossenschaft (2002) For an overview of innovation-fostering institutions in the Swiss VET system see Pedró, Burns, Ananiadou, and de Navacelle (2009) and Rauner (2008).

The use of global and explicit knowledge characterizes the STI mode. Expressing knowledge in a written from for example, in scientific articles, manuals, books of instruction or reports, is a prerequisite for knowledge transfer. The knowledge in the STI mode focusses on the explanation of why something works and less on how it works in a specific context. Typically scientists in universities and R&D departments of companies use and produce such knowledge. The application of codified knowledge to local problems is still part of the STI mode as long as the scientist transfer knowledge in an explicit and codified form (Jensen et al., 2007).

The use of local and implicit knowledge are central elements of the DUI mode. Employees share knowledge by closely interacting with colleagues. The understanding of how something works is more important than an explicit explanation of why something works. The DUI mode also contains learning and interactions with colleagues from different departments and different educational backgrounds. Practices that organize such interactions for example team work and job rotation are also part of this mode. Communication in these interactions are mainly informal and do not required to codify knowledge.

C. Vocational education and collective learning

We use the two ideal-type learning modes to describe collectively-organized vocational education as a system that consists of knowledge collection, collective knowledge processing and aggregation, knowledge transfer, and knowledge application. In such a system knowledge changes its state of aggregation in the sense that it local in the beginning, highly aggregated in between and local in the end.

We illustrate this knowledge aggregation with recent examples from the German automotive industry. The German automotive is a suitable example, because the introduction of electric cars calls for new qualifications of car mechanics. Electric cars operate with high-voltage engines that workers without a specific qualification cannot maintain. Although the old (2007) training curriculum for car mechanics contains building, maintaining and repairing electronics in cars with a combustion engines, this knowledge cannot easily transferred to high-voltage engines (Bundesministrium für Wirtschaft und Technologie, 2007). The new (2013) curriculum therefore prepares car mechanics also to work with high-voltage engines (Bundesministrium für Wirtschaft und Technologie, 2013). As electromobility is currently not widespread in Germany, the new curriculum introduces an emerging technology and fosters its diffusion. Another example is the introduction of computer-numerical-controlled (CNC) machines in the curricula of machining metal operators in Germany in the end of 1980s.

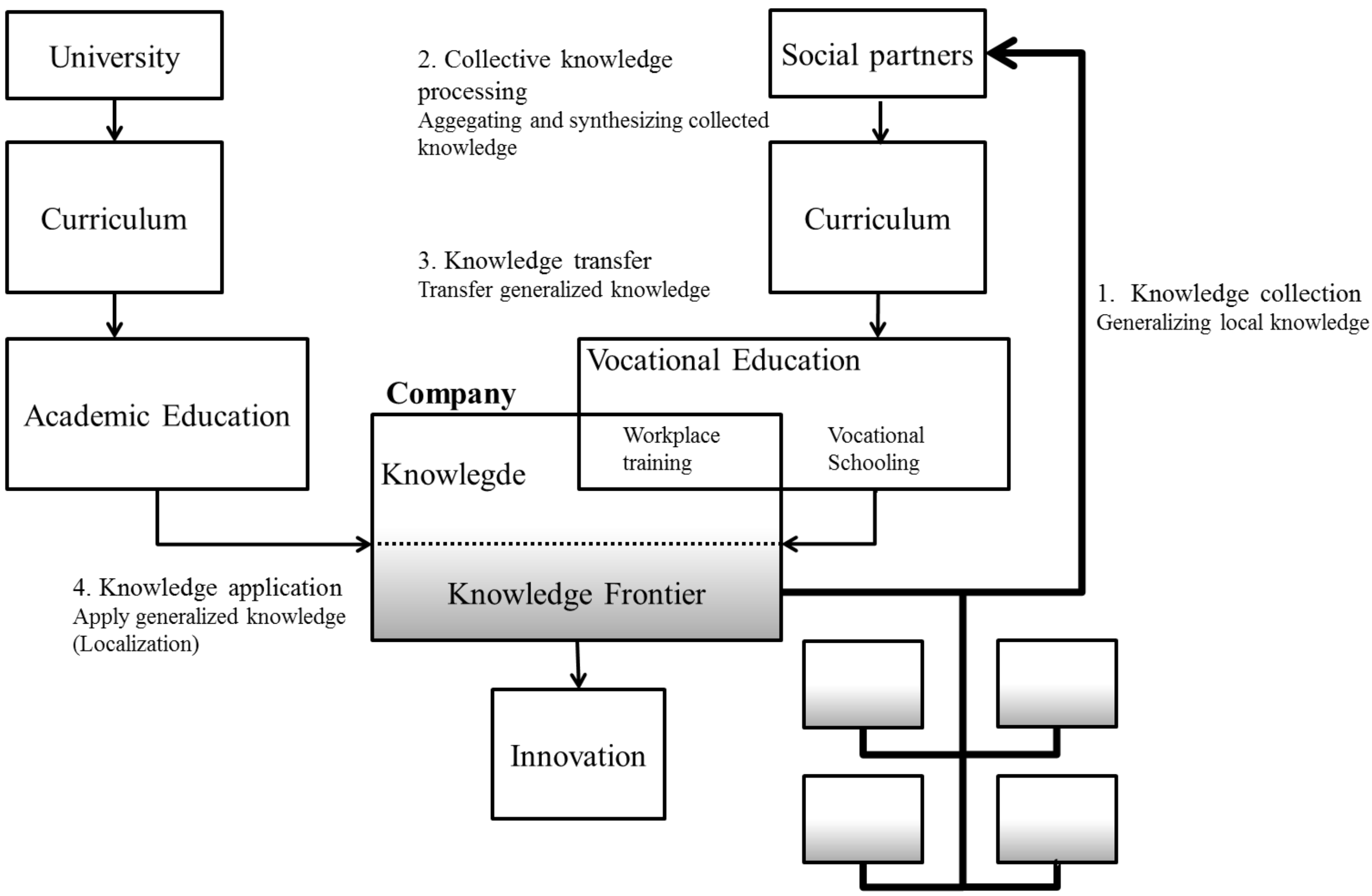

Figure 1 Knowledge processing in collectively organized vocational education.

1. Knowledge collection

The left side of figure 1 depicts knowledge flows during the curriculum updating process. To integrate knowledge on latest technologies and best practices in the curriculum the social partners (employer, employee and governmental organizations) collect knowledge from companies that participate in the system. Knowledge on latest technologies and best practices is often highly localized. To communicate this knowledge firms have generalize it. In this process firms are autarkic.

Training curricula contain information on what an apprentice should know, how certain tasks should be performed and why something works. The revision of a curriculum therefore requires a company not only to communicate codified knowledge but also localized knowledge. Detailed descriptions of best practices require explanations on how a certain task is performed, how workers are organized, how workers communicate and how the task is embedded in other processes. Therefore companies generalize their localized knowledge to also communicate best practiced for the revision of curricula.

To illustrate the generalization of localized knowledge we take an example from the development of training curricula for car mechanics. AUDI, a large German manufacturer of cars with combustion engines but also electric engines, analyzed company-specific processes in the production of electric cars and launched a project for the development of training standards for building, maintaining and repairing high-voltage engines in cooperation with the f-bb (Research Institute for Vocational Education and Training). To generalize its local knowledge AUDI conducted structured interviews with employees (car mechanics, car electricians) who work with high-voltage technology. AUDI uses the responses to develop a competence profile and minimal requirements for working with high-voltage electric cars (Müller and Kohl, 2014; Elektromobilität verbindet, 2015a). This competence profile covers expert knowledge on electrical engineering, measurement engineering, occupational safety and the design of high-voltage systems. Moreover it covers knowledge how tasks, for example, how to install high-voltage parts in cars, and communication competences and self-competences (Müller and Kohl, 2014). AUDI publishes this competence profile and the developed learning materials to update curricula of all car-related occupations.

This example shows that companies not only communicate general knowledge but also generalize local knowledge on processes and on competences. Companies generalize knowledge that is typically part of the DUI mode of innovation and learning. They thus communicate knowledge from both modes. By linking scientific knowledge to competences and social interactions, this example shows that STI and DUI mode work in concert. Transferring only knowledge that is easily transferable (codified knowledge) might not unfold its effectiveness and usefulness if not combines with processes and competences. The joint communication of knowledge form the STI and the DUI mode is in line with Jensen et al. (2007) who argue for a stronger effect on innovation if both modes are combined.

2. Collective knowledge processing

Collective knowledge processing summarizes the aggregation and synthesis of knowledge that originates from the first step. While the knowledge that comes firm the first step still might be specific to a single company, because the company generalized it only for the sake of communication, it might not be applicable as such in all companies. The social partners therefore aggregate and standardize the collected knowledge to integrate it in training curricula.[4] In this step firm-specific knowledge is generalized to occupation-specific knowledge. That vocational education should contain occupation-specific knowledge instead

[4] An example for revised curricula for commercial employees and knowledge aggregation provides Pedró, Burns, Ananiadou, and de Navacelle (2009).

of firm-specific knowledge only is part of the legislation in vocational education (e.g. for Switzerland: Bundesversammlung der Schweizerischen Eidgenossenschaft 2002).

A further example form the German automotive industry shows the development of standardized knowledge and competences. In a joint project 4 companies developed a qualification program in the field of electromobility for several occupations (Electrical engineering, mechanics and business administration). They separated commercial form technical occupations and developed a competence profile and training courses for each. The aim of the training program is to qualify as much apprentices as possible independent form the company they work for (Elektromobilität verbindet. 2015b). The program contains 6 courses for commercial occupations (e.g., Sustainability, Financing, Marketing, Technical Basics) and 4 courses for technical occupations (e.g., high-voltage components, installing high-voltage components, basics of automotive engineering). All courses have been integrated in the respective training curricula.

This example shows that companies aggregate knowledge to generate knowledge that can be applied in every company. Thereby the companies do not separate scientific knowledge from practices. In terms of STI and DUI mode this aggregation reflects a stronger codification of knowledge to filter out irrelevant company-specific parts but at the same time to connect best practices with aggregated scientific knowledge.

3. Knowledge transfer

In a third step updated training curricula transfer new standardized knowledge to all firms that train apprentices. The training curricula regulate training in vocational schools and companies (Wolter & Ryan, 2011). To ensure that instructors in vocational schools and companies train according to the updated curricula, instructors attend preparatory courses. The knowledge form the training curricula is thus transferred to the instructors first.

A further example from the automotive industry are train-the-trainer qualifications that convey new standardized knowledge to instructors (Elektromobilität verbindet. 2015c). In these courses instructors learn to work with high-voltage systems of multiple manufacturers. They also learn how to install, maintain and repair high-voltage systems in e-bikes. Thus the application of this new and highly standardize knowledge is not specific to cars but also applicable to other electric vehicles.

The knowledge that instructors transfer to apprentices is highly generalized. This knowledge transfer can be compared with knowledge that is transferred in universities and is therefore strongly linked to the STI mode of learning. As vocational education does not only consist of vocational schools, the DUI part is added in the company where apprentices learn to apply their new knowledge. Contrary to academic education the DUI and the STI mode are not separated here. The codified knowledge that is transferred by the training curricula is connected with best practices. This knowledge is the result of what multiple companies consider to be most relevant in terms of scientific knowledge and practices.

4. Knowledge application

In the last step companies apply the knowledge that is transferred to them by the training curricula. Companies benefit from both the new qualifications that their instructors have and the new qualifications that the apprentices has. By having qualified instructors, firms can absorb the new knowledge that apprentices learn in vocational schools faster.

By applying highly standardized knowledge to local problems and knowing best practices to perform a certain task, the new general knowledge is localized in companies. Companies thereby augment their local and global knowledge that they use to introduce new products and processes. In terms of the learning modes a transfer of highly generalized knowledge adds to a company's STI mode of learning. By applying this knowledge with new practices and combining it with local knowledge also the DUI mode of learning plays a major role in the application of new knowledge from curricula.

In sum collectively-organized vocational education collects, processes, transfers and applies knowledge. It thereby generalizes highly localized knowledge and processes it jointly with codified scientific knowledge. Both types of knowledge remain integrated throughout the entire aggregation process. The only reason for decoupling both types of knowledge is to replace either type with the state of the art and then to integrate both parts again. STI and DUI mode work in concert over the entire cycle form knowledge collection to application of new knowledge.

D. Collective knowledge processing in academic education

Knowledge processes in academic education are highly similar to those in collectively organized vocational education. Its main differences, however, is that DUI and STI mode are mainly decoupled in the curriculum[5] and that revisions of curricula are decentralized (each

[5] Abstracting from university-industry cooperations and also that a student does not pursue an academic career. If we would assume that a student pursues an academic career, university education would also contain DUI components.

university revises its own curricula) whereas it is centralized in vocational education (on curriculum per occupation).

Similar to collectively-organized vocational education, academic education has a knowledge collection process. Researchers codify localized knowledge of their research group in scientific papers. Lecturers and text book editors collect this knowledge and aggregate it. By giving lectures lecturers transfer the knowledge of the curriculum to their students. Students than transfer the knowledge to companies where they can apply it.

E. Hypothesis

Our model of knowledge processing in collectively-organized vocational education has several implications. We will however focus on its main implication that the positive effect of collectively organized vocational education on innovation. Following our argumentation we expect that firms that participate in collectively-organized vocational education have a higher innovation output that firms not doing so. We therefore formulate our hypothesis:

H1: Firms that participate in collectively-organized vocational education become more innovative than firms that do not doing so.

III. Data

A. Sample and Descriptive Statistics

For our empirical analysis we use the Innovation Survey of the Swiss Economic Institute (KOF). The KOF collects the data triennially since 1990 and includes several innovation indicators such as process and product innovation, and patent applications in the questionnaire. These three innovation indicators are binary and take the value of one if a firm innovated or applied for a patent during a three-year period. We use the process and product innovation indicators to construct a fourth innovation measure. This binary measure takes the value of one if either product or process innovation takes the value of one and thus is an indicator for general innovation activity. Thus we title this constructed measure general innovation.

Table 1
Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Dependent variables					
General innovation	2936	0.725	0.446	0	1
Product innovation	2936	0.633	0.482	0	1
Process innovation	2936	0.520	0.500	0	1
Patent applications	2936	0.192	0.394	0	1
Explanatory variables					
Training firm	2936	0.723	0.447	0	1
Firm size (Number of workers)	2936	174.507	515.450	1	8371
Share of workers with university degree	2936	0.059	0.118	0	1
Share of workers with degree higher than VOC	2936	0.159	0.153	0	1
Share of workers with apprenticeship degree	2936	0.509	0.239	0	1
Share of workers with degree lower than VOC	2936	0.273	0.249	0	1
Non-price-competition	2936	0.420	0.494	0	1
Price-competition	2936	0.720	0.449	0	1
Increase in estimated demand for next 3 years	2936	0.426	0.495	0	1
Foreign firm	2936	0.134	0.341	0	1
Lack of skilled workers	2936	0.166	0.372	0	1
Sector					
Manufacturing	2936	0.582	0.493	0	1
Construction	2936	0.066	0.248	0	1
Service	2936	0.351	0.478	0	1
Year					
year 1999	2936	0.197	0.397	0	1
year 2002	2936	0.286	0.452	0	1
year 2005	2936	0.300	0.458	0	1
year 2008	2936	0.218	0.413	0	1
Region					
Lake Geneva Region	2936	0.013	0.115	0	1
Espace Mittelland	2936	0.213	0.409	0	1
Northwestern Switzerland	2936	0.173	0.378	0	1
Zurich	2936	0.238	0.426	0	1
Estern Switzerland	2936	0.248	0.432	0	1
Central Switzerland	2936	0.115	0.320	0	1
Instruments					
Firm age	2936	59.966	42.295	1	351
German-Speaking firm	2936	0.985	0.120	0	1

The data set contains detailed information on the educational composition of a firm's workforce. We use this information to construct our main explanatory variable: Firms' participation in vocational education. The KOF classifies workers into 5 categories: workers with university degrees, workers with degrees higher than vocational education, workers with vocational education degrees, workers with degrees lower than vocational education, and apprentices. We generate a binary variable entitled “training firm” that takes the value of one if a firm employs at least one apprentice and zero otherwise.

To investigate the influence of vocational education on firms’ innovativeness, we exclude observations from waves 1990 to 1996. These waves do not provide the necessary information either for the innovation measures or for the main explanatory variables. We furthermore restrict our sample to German-speaking regions as vocational education is more widespread in these regions.

Table 1 contains descriptive statistics for our estimation sample. More than 50% of the firms in our sample report either a process innovation or a product innovation. Compared to the number of firms that applied for a patent (19.2%), these number show that we measure innovations with different significance and novelty. While process and product innovations might include innovations that are new to the firm, innovations that are patentable might be either new to the world or new to the industry.

Firms in our sample are on average larger than the average firm in Switzerland. The KOF survey applies a sampling scheme that oversamples large and medium-size firms. Moreover, as most of the firms that report missing values are small, our remaining sample represents large and medium-size firms stronger. The stronger representation of large and medium-size firms does not threaten the validity of our results as we focus on the effect of vocational education on firms’ innovativeness in general not on the estimation of a representative effect for Switzerland.

B. Instrumental Variables

The data set includes valuable information for the construction of instrumental variables for firms’ training decision such as information on firms’ age, the location of the firm, and the language spoken by the employees. Vocational education in Switzerland has a long-lasting tradition and gained its popularity in the mid of the 20th century (Knutti, 2007). Due to the adoption of vocational education and its integration in firms’ organizational structure, we expect that older firms developed a training tradition that relies on vocational education. Although vocational education is still highly recognized and wide-spread in Switzerland, we expect younger companies not to have such a highly established training tradition compared to older companies. Therefore we choose firm age as an instrument for firms training participation.

As we cannot assess the validity of the instrument in the just-identified case, we choose an additional instrument to conduct robustness checks of our IV specification. The additional instrument we use is also related to firms training tradition. Firms' training traditions differ by the three linguistic regions (French, German and Italian) in Switzerland. Firms located in the German-speaking part of Switzerland typically are more likely to offer vocational education programs than firms located in the French- or Italian-speaking part (Gonon & Maurer, 2012; OPET, 2010). As the data set contains linguistic information at the postcode level and at the firm level, we combine both measures to construct an instrumental variable.

The KOF makes an initial assignment of the questionnaires to firms based on a linguistic categorization provided by the Swiss Post. If a firm is unable to fill in a questionnaire due to linguistic difficulties it has the possibility to request a different linguistic version of the questionnaire. Having this information, we can construct a binary variable that indicates whether firms' language is German or not. As our sample contains only firms that are located in German-speaking regions, we expect all firms that returned a German questionnaire to have a stronger training tradition than the remaining firms in our sample.

IV. Estimation Strategy

For our empirical analysis we augment the knowledge production function proposed by Pakes and Griliches (1984). Their original model contains of two equations: The first equation explains knowledge growth by a firm-specific productivity shifter, a time trend and past period's research expenditures. The second equation explains the generation of patents with time trend and knowledge growth. We augment this framework by introducing the knowledge that the training curricula of vocational education contain as an additional factor that enhances the knowledge growth of a firm. Following Pakes and Griliches (1984) we include firms' knowledge growth in the patent equation and allow for the occurrence of other innovation outcomes that summarize non-patentable knowledge. Thus product and process innovation are considered as outcomes of firms knowledge growth.

We adapt this equation to our available variables. We operationalize firms' application of training curricula by its participation in vocational education.[6] We estimate a linear probability model, because we have binary innovation measures taking the value of one if the firm innovated successfully during the observation period and zero otherwise. The advantage of the use of the linear probability over probit models is the direct calculation of the marginal

[6] Due to missing observations for firms' R&D expenditures, we cannot include their lags into our model.

effects of the training decision on innovation. If we take endogenous training decision into account this feature allows the comparison of the results between the IV specification and the OLS specification. To compare the results from the linear probability models with more suitable models for binary choice variables we estimate probit models.

$$i_{jt} = \gamma_0 + \gamma_1 tr_{jt} + \sum_{k=2}^{K} \gamma_k x_{kjt} + e_{jt} \tag{1}$$

Equation 1 is our basic estimation equation. We use 4 different measures for innovation (general innovation, product innovation, process innovation and patent applications). The variable of interest is the binary indicator for training participation of firm j at time t: tr_{jt}. We subsequently include a set of control variables. These control variables contain firm size, educational composition of a firms' workforce, competition measures (price and non-price competition)(Aghion, Bloom, Blundell, Griffith, & Howitt, 2005), lack of qualified workers, an indicator for foreign firm, and sector, year ,and regional dummies.

Endogenous training decision might be a potential source of bias in equation 1. This bias occurs if unobservable decisions influence both the training decision and the innovation output. Strategic management decisions for example might aim to foster innovation and simultaneously introduce the training of apprentices. To take this endogenous training decision into account we apply an instrumental variable strategy (Angrist & Krueger, 2001). As instrumental variable, we use the firm age and firms' language to measure training tradition.

Due to the panel structure of our data set we risk getting biased standard errors if do not correct for clustering at the firm level (Moulton, 1990). Therefore we use cluster-corrected standard errors for the basic equation and the instrumental variable equation.

V. Results

A. OLS and Probit Estimates

To test our hypothesis, we estimate equation 1 and use general innovation, product innovation, process innovation, and patent applications as dependent variables. According to our hypothesis we expect a positive impact on firm innovativeness of the firms' training participation. Table 2 shows the results for the estimation equation that includes the full set of control variables.

Column 1 of table 2 shows the analysis of the influence on general innovation of a firms' decision to train apprentices. The results indicate that firms that participate in training have a 7.8 percentage points higher probability to innovate than firms not participating in vocational education. Columns 2 and 3 show the estimation results for product innovation and process innovation, respectively. Again, the decision to train apprentices is positively associated with product and process innovation. These coefficients are statistically significant at the 1 percent and the 5 percent level, respectively. The marginal effect of training is 7.2 percentage points for product innovation and 4.8 percentage points for process innovation. We also obtain a positive association of firms training participation and its patent applications. The marginal effect of training on patent applications is 6.7 percentage points.

Table 2
Linear Probability Model

Dependent variable	General innovation	Product innovation	Process innovation	Patent applications
Independent variable	Coef.	Coef.	Coef.	Coef.
Training firm	0.0782***	0.0715***	0.0475**	0.0666***
	(0.0220)	(0.0236)	(0.0240)	(0.0191)
Workforce education controls	yes			
Sector controls	yes			
Year controls	yes			
Region controls	yes			
Firm controls	yes			
Observations	2936			
R-squared	0.1192	0.1232	0.0719	0.1730

Note: Cluster robust standard errors in parenthesis (Cluster level: Firm).
* Statistically significant at the 0.1 level; ** at the 0.05 level; *** at the 0.01 level.

We re-estimate equation 1 using a Probit model to compare the estimation results from the linear probability model to a model that is more suitable for binary choice data. The marginal effects for the Probit estimates appear in table A1 in the appendix. Compared to the results we report in table 2, the marginal effects from the Probit model are similar in magnitude and direction. The estimates from the Probit model are highly significant for all innovation outcomes like those estimated in the linear probability model.

B. Instrumental Variable Estimates

As equation 1 does not account for potential endogenous training decision, we use an instrumental variable approach. Table 3 shows the estimation results of the first and second stage of a TSLS estimation.[7] In the first stage we regress the dummy that indicates firms' training participation on firm age and a set of control variables. Firm age is statistically significant at the 1 percent level (indicating a strong instrument) and shows a positive coefficient as expected. This result supports our expectation that older firms are more likely to offer dual-track VET compared to younger firms.

Table 3
Linear probability model, IV estimation (TSLS), instrument firm age

	First Stage	Second Stage			
Dependent Variable	Training firm	General Innovation	Product Innovation	Process Innovation	Patent Applications
Independent Variable	Coef.	Coef.	Coef.	Coef.	Coef.
Training firm		0.1946*	0.0794	0.1072	0.1727
		(0.1039)	(0.1194)	(0.1198)	(0.1091)
Firm Age	0.0021***				
	(0.0003)				
Workforce Education Controls	yes	yes			
Sector Controls	yes	yes			
Year Controls	yes	yes			
Region Controls	yes	yes			
Firm Controls	yes	yes			
Observations	2936	2936			
Centered R-squared	0.1090	0.1066	0.1232	0.0693	0.1595

Note: Cluster robust standard errors in parenthesis (Cluster level: Firm).
* Statistically significant at the 0.1 level; ** at the 0.05 level; *** at the 0.01 level.

Columns 2-5 of table 3 show the second stage results for the four innovation measures. These results deviate from the OLS and Probit estimates in the sense that their level of statistical significance is lower and the estimated coefficients are larger. Similar to the OLS and Probit estimates the IV estimates indicate a positive effect of vocational education on the innovativeness of firms. This effect is statistically significant at the 10 percent level only for general innovation. For product and process innovation and for patents we do not obtain a statistically significant effect at conventional significance levels. The statistically significant estimate for general innovation is more than twice the size of the OLS estimate. The increase in the value of the coefficients might occur due to the importance of vocational education for the innovativeness of highly traditional firms. These firms might rely more heavily on the

[7] Because a firm's training decision does not change considerably over time, we refrain from estimating a fixed effects model.

recruitment of workers with vocational education and qualifications that build upon vocational education than on firms that do not have a strong training tradition.

C. Robustness of Instrumental Variable Estimates

As we cannot asses the validity of the instrument in the just-identified case presented in table 3, we extend our IV approach by including instruments measuring training tradition based on firm-level language measures. Table 4 shows the estimation results of the first and second stage of a GMM estimation.[8] The first stage contains two instruments: firm age and a dummy that indicates German as firms' major language (German-speaking firm).

Table 4 shows statistically significant estimates at the 1 percent and 5 percent level, respectively and the expected signs for all of our instruments in the first stage. Again, as in table 3, we find a positive coefficient for firm age. Furthermore, we find a positive coefficient for German-speaking firms. The positive coefficient indicates that German-speaking firms have a higher probability to participate in vocational education than French-speaking firms (the reference). To assess the strength of both instruments we test for the joint significance of both instruments in the first stage. Table 4 shows the F-statistic and the corresponding p-value. Both instruments are jointly significant at the 1 percent level, indicating strong instruments. Moreover a comparison of the F-statistic with the critical values reported in Stock and Yogo (2002) shows that the instruments are below the 10 percent maximal size threshold.[9]

A specification with two instruments and one endogenous variable allows us to test for overidentifying restrictions. As we use a GMM estimator that is efficient for clustering and heteroscedasticity, the Hansen J statistic and its corresponding p-value are appropriate to test the validity of our instruments. For the second stage specification shown in columns 2-4 of table 4 we find p-values above 0.33. Thus the hypothesis that the instruments are valid cannot

[8] We present the results of a GMM estimation in this section as GMM is an efficient estimator if heteroscedasticity and clustering occurs. For the sake of comparability we include the TSLS estimates in the appendix (Table A3). We also report estimates from a limited information maximum likelihood (LIML) estimation (Table A4). LIML is more robust to weak instruments than the procedures mentioned above (Stock, Wight, & Yogo, 2002). The estimates of the LIML estimation are similar to those we report in Table 4.

[9] We follow Baum, Shaffer, and Stillman (2007) by comparing the F-statistic of the joint significance test for both instruments to the critical values reported in Stock and Yogo (2002) for the Cragg-Donald statistic. As we adjust the standard errors for clustering at the firm level, the i.i.d. assumption of the Cragg-Donald statistic is violated and the Kleibergen-Paap statistic is appropriate. In the single equation case, the Kleibergen-Paap rk F-statistic reduces to a F-statistic for the joint significance of the instruments in the first stage (Kleibergen & Schaffer 2010). The comparison of the F-statistic to the critical values reported in Stock and Yogo (2002) shows that a 5 percent bias hypothesis is rejected in less than 10 percent of the time. This test result supports the strength of our instruments.

be rejected in these cases. Only in the specification shown in column 5 we can reject this hypothesis at the 10 percent level.

Table 4
Linear probability model, IV estimation (GMM), all Instruments

	First Stage	Second Stage			
Dependent Variable	Training firm	General Innovation	Product Innovation	Process Innovation	Patent Applications
Independent Variable	Coef.	Coef.	Coef.	Coef.	Coef.
Training firm		0.1923**	0.1171	0.0698	0.0923
		(0.0978)	(0.1118)	(0.1135)	(0.0986)
Firm Age	0.0021***				
	(0.0003)				
German-Speaking Firm	0.2507**				
	(0.0979)				
Workforce Education Controls	yes	yes			
Sector Controls	yes	yes			
Year Controls	yes	yes			
Region Controls	yes	yes			
Firm Controls	yes	yes			
Observations	2936	2936			
F-Statistic (joint significance of instruments)	34.222				
p-value	0.000				
Hansen J Statistic		0.004	0.875	0.919	2.961
p-value		0.9476	0.3495	0.3376	0.0853
Centered R-squared	0.1132	0.1246	0.1216	0.0715	0.1721

Note: Cluster robust standard errors in parenthesis (Cluster level: Firm).
* Statistically significant at the 0.1 level; ** at the 0.05 level; *** at the 0.01 level.

The estimates of the second stage in table 4 are in line with those in table 3. Similarly to table 3, the results in table 4 show a positive impact of firms' participation in training on its innovativeness. This effect is statistically significant at the 5 percent level. For the other innovation outcomes we do not find statistically significant effects. The results in tables 3 and 4 indicate that our hypothesis cannot be rejected. With different sets of instruments we find support for a positive impact of participating in vocational education on the innovativeness of firms.

The use of firm age and firm language as instruments has its limitations. The population of older firms might be different from the population of younger firms. Older firms in the sample have undergone a selection process by market forces. Only those firms that survived this selection process are included in the sample. The population of younger firms in contrast might include firms that will not survive in the market in the next decade. As we do not have information on firms that exited the market, we cannot properly account for selectivity in this sample. Nevertheless, we make an attempt to run the IV estimation with the full set of instruments in a more homogeneous sample by excluding firms that are younger than 10. This restriction allows the exclusion of firms that are most likely to exit the market.

Table 5
Linear probability model for firms aged 10 and older, IV estimation (GMM), all Instruments

	First Stage	Second Stage			
Dependent Variable	Training firm	General Innovation	Product Innovation	Process Innovation	Patent Applications
Independent Variable	Coef.	Coef.	Coef.	Coef.	Coef.
Training firm		0.1861*	0.0873	0.1081	0.1027
		(0.1110)	(0.1285)	(0.1269)	(0.1122)
Firm Age	0.0020***				
	(0.0003)				
German-Speaking Firm	0.2233**				
	(0.1019)				
Workforce Education Controls	yes	yes			
Sector Controls	yes	yes			
Year Controls	yes	yes			
Region Controls	yes	yes			
Firm Controls	yes	yes			
Observations	2764	2764			
F-Statistic (joint significance of instruments)	26.545				
p-value	0.000				
Hansen J Statistic		0.068	0.745	2.138	2.779
p-value		0.7936	0.3881	0.1437	0.0955
Centered R-squared	0.1105	0.1055	0.1223	0.0657	0.1643

Note: Cluster robust standard errors in parenthesis (Cluster level: Firm).
* Statistically significant at the 0.1 level; ** at the 0.05 level; *** at the 0.01 level.

Table 5 shows the GMM estimates of a specification that includes firm age and the German-speaking firm dummy as instruments in a sample that contains firms that are 10 or older. This table shows similar results compared to table 4 where we use the entire sample. The first stage estimates in table 5 show a positive effect of all instruments on firms' training participation at a similar statistical significance level reported in table 5. Furthermore, the results in table 5 support the second stage results in table 3 and 4. We find a positive and statistically significant (at the 10 percent level) effect of firms participation in training on general innovation. Thus the estimation in a more homogeneous sample supports our hypothesis that vocational education has a positive impact on firms' innovativeness.

We repeat the estimation of the specification that uses the full set of instruments in samples that are less homogeneous. Therefore we generate an additional sample that includes firms that are 5 or older. Table A5 in the appendix show the estimation results. The results from this table do not contradict our hypothesis. In table A5 we again find a positive effect of vocational education on general innovation and thus support for our hypothesis.

D. Effects of Training Intensity

The impact of vocational education on innovativeness might not only depend on the participation in the vocational education system but also on the number of apprentices employed. Apprentices are one channel that contributes to the knowledge diffusion within the vocational education system. Firms that employ apprentices have to decide on the optimal number. The employment of apprentices comes along with their benefits for innovation but might also generate obstacles. Firms that employ an excessively large number of apprentices might lack sufficient resources to develop apprentices' skills trough workplace training. In this case apprentices do not have the opportunity to apply the knowledge they gained from training curricula to real-world problems. Consequently, these firms can integrate the knowledge from the curricula at a lower rate than firms that foster the application of this knowledge during workplace training. We therefore expect that the relation between innovation and training intensity is inverted u-shaped.

To measure the association of training intensity with innovativeness, we include the number of apprentices employed to our estimation equation. As we expect an inverted u-shaped functional form, we include the squared number of apprentices in our specification. Table 6 shows the estimation results for specifications that include the full set of control variables. Vocational education is positively associated with all innovation measures. Furthermore, the coefficients for the training intensity measures show an inverted u-shaped relationship between the number of apprentices and innovation outcomes. This relation is statistically significant for general innovation, product innovation and patent applications. A calculation of the turning points for all specification yields values that exceed 400. These large values should be interpreted with caution as the estimates might be driven by large firms in the sample. We refrain from an exact calculation of the turning points for each industry as we are more interested in the general relation between training intensity and innovation[10].

Table 6
Linear Probability Model Including Training Intensity Measures

Dependent Variable	General Innovation	Product Innovation	Process Innovation	Patent Applications
Independent Variable	Coef.	Coef.	Coef.	Coef.
Training firm	0.0699***	0.0603***	0.0414	0.0501***
	(0.0224)	(0.0243)	(0.0245)	(0.0195)
Training Intensity	0.0014**	0.0019**	0.0010	0.0027***
	(0.0006)	(0.0008)	(0.0007)	(0.0009)
Training intensity squared	-0.0017*	-0.0021**	-0.0017	-0.0034***
	(0.0009)	(0.0011)	(0.0013)	(0.0010)
Workforce Education Controls	yes	yes	yes	yes
Sector Controls	yes	yes	yes	yes
Year Controls	yes	yes	yes	yes
Region Controls	yes	yes	yes	yes
Firm Controls	yes	yes	yes	yes
Observations	2936	2936	2936	2936
R-squared	0.1208	0.1255	0.0726	0.1810

Note: Cluster robust standard errors in parenthesis (Cluster level: Firm).
* Statistically significant at the 0.1 level; ** at the 0.05 level; *** at the 0.01 level.
Coefficients and Standard Errors for the variable Training intensity squared are multiplied by 1000.

The results in table 6 show that after controlling for training intensity a positive association of training participation remains. Compared to training intensity, the coefficient of training participation is up to 50 times larger (in the case of general innovation) than the coefficient for training intensity. This result shows that firms experience an innovation premium from participating in training regardless their training intensity. Due to the lack of a sufficient number of instrumental variables, we cannot show that this relationship is a causal one. Nevertheless, the results are in line with our theoretical expectations and support our hypothesis.

[10] The large turning point of 400 should not be interpreted in the way that firms should hire more apprentices to achieve optimal innovativeness. Instead it should show in a general manner that firms might enhance their innovativeness by increasing training intensity.

E. Effects of Firm Size

The size of a firm reflects the availability of resources that can be devoted to R&D and thus lead to innovation (Acs & Audretsch, 1988). Larger firms are expected to have more resources that can be used for R&D than smaller firms. Thus, we expect firms of different sizes to benefit differently from the knowledge diffusion of vocational education. In terms of innovativeness we expect larger firms to benefit less from vocational education as they foster the inclusion of their latest knowledge in training curricula. Smaller firms might not have access to the latest technologies and therefore benefit strongly from knowledge the curricula contains.

To obtain estimates for the relation of firm size and firms' training participation, we augment our estimation equation by including the interaction of firm size and training participation. Table 7 shows the results of this estimation including the full set of instruments. Participating in vocational education is still positively associated with all innovation measures. The interaction of vocational education and firm size is negatively associated with innovation. These estimates are not statistically significant at conventional levels. Only for patent applications we obtain a statistically significant estimate at the 5 percent level.

Table 7
Linear Probability Model Including Firm Size-Interaction with Training Status

Dependent Variable	General Innovation	Product Innovation	Process Innovation	Patent Applications
Independent Variable	Coef.	Coef.	Coef.	Coef.
Training firm	0.0835***	0.0760***	0.0574**	0.0802***
	(0.0233)	(0.0247)	(0.0256)	(0.0198)
Firm size	0.0001	0.0001	0.0003*	0.0003***
	(0.0001)	(0.0001)	(0.0002)	(0.0001)
Training firm*Firm size	-0.0001	-0.0001	-0.0002	-0.0003**
	(0.0001)	(0.0001)	(0.0002)	(0.0001)
Workforce Education Controls	yes	yes	yes	yes
Sector Controls	yes	yes	yes	yes
Year Controls	yes	yes	yes	yes
Region Controls	yes	yes	yes	yes
Firm Controls	yes	yes	yes	yes
Observations	2936	2936	2936	2936
R-squared	0.1194	0.1234	0.0726	0.1749

Note: Cluster robust standard errors in parenthesis (Cluster level: Firm).
* Statistically significant at the 0.1 level; ** at the 0.05 level; *** at the 0.01 level.

This robustness check shows that the positive association of vocational education and innovation remains if we extend the analysis by including a vocational education and firm size interaction. By taking the magnitude of this association into account, larger firms do not gain as much from participating in vocational education as smaller firms do.

VI. Discussion

In this paper we analyze the impact of vocational education on innovation at the firm level. We argue that institutions in secondary education function as a knowledge diffusion mechanism. We identify the training curricula of vocational education as a diffusion channel. Thus we hypothesize that firms participating in vocational education are more innovative than firms not participating in such training. We use data from a large Swiss innovation survey to test our hypothesis. Our estimation results support that vocational education has a positive impact on innovation at the firm level. We find a positive and statistically significant association of vocational education with general innovation, product innovation, process innovation and patent applications. This positive association remains after controlling for firms' training intensity. If we take the endogeneity of the training decision into account, the results show a positive and statistically significant effect of vocational education on general innovation.

Our empirical results are in line with our hypothesis. Due to the high involvement of employer association, employee association and the government in the creation and revision of curricula, new knowledge enters the vocational education. Apprentices learn this knowledge in vocational schools and apply it during workplace training. Firms participate in this system by providing vocational education and thus have access to an additional knowledge source. This knowledge source accelerates the adaptation of new technologies and results in a higher innovative output compared to firms that do not participate in the system.

We find support for our hypothesized effect throughout several empirical approaches that treat firms' participation vocational education as either exogenous or endogenous. We use firms' training tradition as an instrument for training participation and take two approaches to operationalize training tradition. In the first approach we argue that firms' age influences its training tradition. Older firms had more time to establish their training tradition than younger firms. As vocational education experienced a boom in the mid of the last century, we expect that it became part of an older firms' training tradition. In the second approach we use a firms' language to operationalize training tradition. In Switzerland firms training participation differs by linguistic region. In the German-speaking part of Switzerland firms offer vocational education more frequently than in the French or Italian-speaking part. We argue that besides the regional differences of training participation, language reflects firms' training tradition. Therefore, we focus our analysis on the German-speaking part and identify firms that have a different language. We expect these firms to have a lower probability to participate in training than their German-speaking counterparts. Our analyses show that the instruments influence training participation in the expected way. Older firms and German-speaking firms have a higher probability to train. The results from our IV approach are highly robust throughout different specifications. Also when we consider that the sample consists of a selection of firms that survived and exclude those that are most likely to exit, our IV results remain stable.

The use of firm age and firm language as instruments for training participation has its limitations. Although our test statistics support the strength and the validity of the instruments, the exogeneity of the instruments might not be comparable with the standards of a randomized experiment. Firm language might enable firms to access certain innovation networks. German-speaking firms might access German-speaking networks easier than French-speaking networks. The association of firms with different networks might influence their innovativeness. To reduce this source of bias, we restrict our sample to the German-speaking part of Switzerland. Nevertheless French-speaking firms might have a different network structure than their German-speaking counterparts resulting in different innovations. Firm age might represent selectivity in the sample we use for our analyses. If older firms represent a selection of firms that are more innovative our results might be biased. Because we lack information on firms that exited the market, we cannot test for this sort of bias. Despite these limitations we find a highly robust and positive coefficient of training participation for general innovation, a result that indicates at least a strong association of training participation and firms' innovativeness.

Further analyses focus on training intensity and firm size. After controlling for training intensity we can show that training participation is still positively associated with innovation. This association indicates that the participation itself is beneficial for firms. The magnitude of the influence of training intensity on innovation is small compared to the coefficient for training participation. Thus, firms might not necessarily have a high training intensity to benefit from the knowledge included in the training curricula. We furthermore investigate whether a high training intensity boosts firms' innovativeness. Our results show an inverted u-shaped relation of training intensity and innovation. An extensive training intensity might therefore reduce firms' gains from training participation.

We lastly investigate the influence of firm size on firms' gains from training participation. We expect larger firms to have comparably smaller benefit from participating in vocational education than smaller firms. Large firms actively participate in the revision of curricula and suggest potential directions for curricula development based on the latest technology they use. Our results show a negative association for the interaction of firm size and training participation and that supports our expectation.

Our results have several theoretical and practical implications. Secondary vocational education contributes to innovation in highly developed countries like Switzerland. This finding contradicts standard models by Krueger and Kumar (2004a, 2004b) and Aghion and Howitt (2006). From the perspective of these models, vocational education does not have the potential to contribute to innovation in a highly innovative environment. As vocational education differs between countries in terms of generality and quality, recent models could consider both factors of vocational education.

By identifying and analyzing knowledge diffusing mechanisms in the Swiss vocational education system we highlight the importance of knowledge diffusion for innovativeness. This finding is of particular importance of policy makers and professional organizations. To ensure a constant inflow of new technologies into curricula, the cooperation of different stakeholders in the curriculum setting process is necessary. Especially the cooperation with stakeholders that determine the use of new technologies in a certain occupation (e.g., customers, suppliers, research institutions) is important for the definition of future contents of the curricula. These stakeholders might not necessarily operate in a similar industry but their expertise might help to adapt faster to technological changes.

Appendix

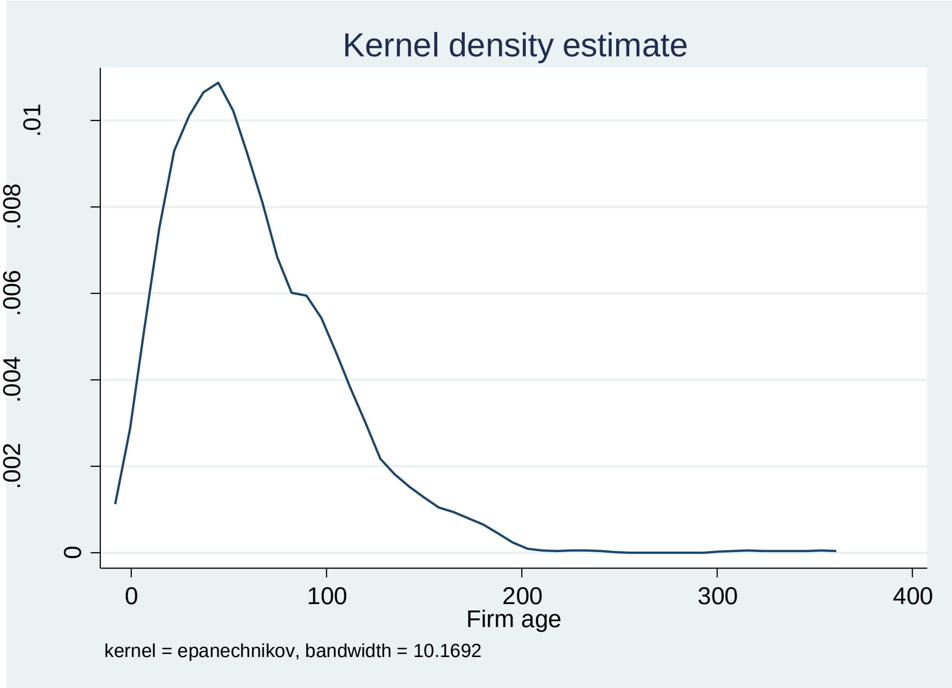

Fig. A1. – Distribution of the founding year

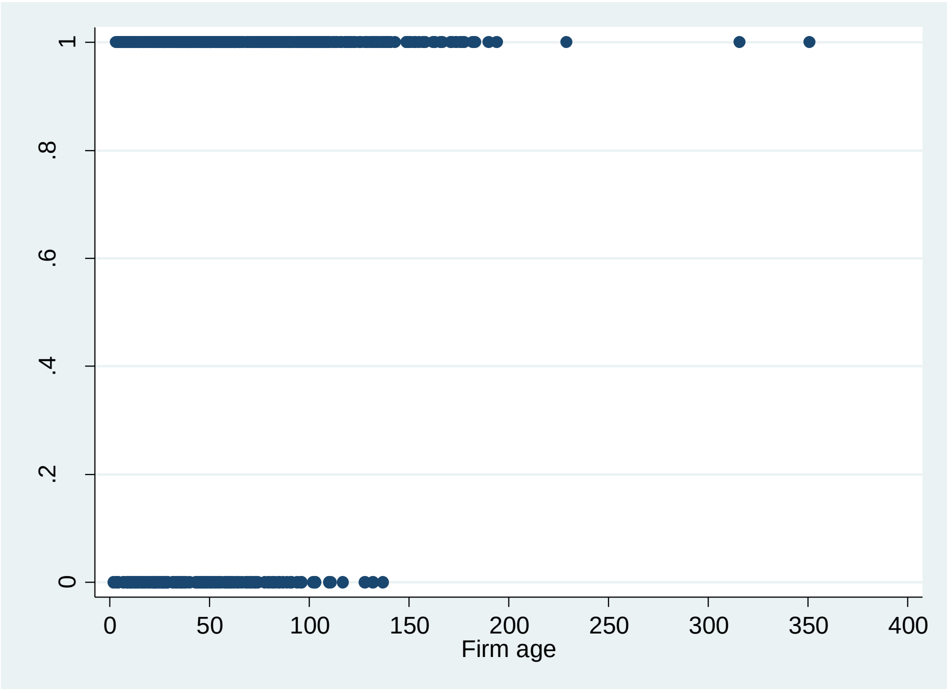

Fig. A2. – Firm age and apprenticeship training

Table A1
Correlations between dependent variables, endogeneous variables and instruments

	1.	2.	3.	4.	5.	6.	7.
1. General innovation	1.000						
2. Product innovation	0.808	1.000					
	(0.000)						
3. Process innovation	0.640	0.408	1.000				
	(0.000)	(0.000)					
4. Patent applications	0.300	0.352	0.183	1.000			
	(0.000)	(0.000)	(0.000)				
5. Training firm	0.072	0.060	0.044	0.073	1.000		
	(0.000)	(0.001)	(0.017)	(0.000)			
6. Firm age	0.033	0.012	0.023	0.033	0.227	1.000	
	(0.074)	(0.527)	(0.213)	(0.079)	(0.000)		
7. German-Speaking firm	0.014	0.031	-0.004	-0.005	0.070	0.007	1.000
	(0.450)	(0.097)	(0.842)	(0.773)	(0.000)	(0.721)	

Note: Significance levels in parentheses.

Table A2
Probit model

Dependent variable	General innovation	Product innovation	Process innovation	Patent applications
Independent variable	Marg. Eff.	Marg. Eff.	Marg. Eff.	Marg. Eff.
Apprenticeship-training firm	0.0726***	0.0683***	0.0447*	0.0717***
	(0.0208)	(0.0229)	(0.0240)	(0.0200)
Workforce education controls	yes			
Sector controls	yes			
Year controls	yes			
Region controls	yes			
Firm controls	yes			
Observations	2936			
Pseudo R-squared	0.1085	0.0997	0.0547	0.2015

Note: Cluster robust standard errors in parenthesis (Cluster level: Firm).

* Statistically significant at the 0.1 level; ** at the 0.05 level; *** at the 0.01 level.

Table A3
Linear probability model, IV estimation (TSLS)

	First Stage	Second Stage			
Dependent Variable	Training firm	General Innovation	Product Innovation	Process Innovation	Patent Applications
Independent Variable	Coef.	Coef.	Coef.	Coef.	Coef.
Training firm		0.1926**	0.1150	0.0705	0.1228
		(0.0978)	(0.1118)	(0.1135)	(0.1001)
Firm Age	0.0021***				
	(0.0003)				
German-Speaking Firm	0.2507**				
	(0.0979)				
Workforce Education Controls	yes	yes			
Sector Controls	yes	yes			
Year Controls	yes	yes			
Region Controls	yes	yes			
Firm Controls	yes	yes			
Observations	2936	2936			
F-Statistic (joint significance of instruments)	34.222				
p-value	0.000				
Hansen J Statistic		0.004	0.875	0.919	2.961
p-value		0.9476	0.3495	0.3376	0.0853
Centered R-squared	0.1132	0.1070	0.1217	0.0715	0.1692

Note: Cluster robust standard errors in parenthesis (Cluster level: Firm).

* Statistically significant at the 0.1 level; ** at the 0.05 level; *** at the 0.01 level.

Table A4
Linear probability model, IV estimation (LIML)

	First Stage	Second Stage			
Dependent Variable	Training firm	General Innovation	Product Innovation	Process Innovation	Patent Applications
Independent Variable	Coef.	Coef.	Coef.	Coef.	Coef.
Training firm		0.1925**	0.1155	0.0707	0.1245
		(0.0978)	(0.1129)	(0.1145)	(0.1031)
Firm Age	0.0021***				
	(0.0003)				
German-Speaking Firm	0.2507**				
	(0.0979)				
Workforce Education Controls	yes	yes			
Sector Controls	yes	yes			
Year Controls	yes	yes			
Region Controls	yes	yes			
Firm Controls	yes	yes			
Observations	2936	2936			
F-Statistic (joint significance of instruments)	34.222				
p-value	0.000				
Hansen J Statistic		0.004	0.875	0.919	2.958
p-value		0.9476	0.3495	0.3376	0.0853
Centered R-squared	0.1132	0.1070	0.1217	0.0715	0.1690

Note: Cluster robust standard errors in parenthesis (Cluster level: Firm).

* Statistically significant at the 0.1 level; ** at the 0.05 level; *** at the 0.01 level.

Table A5
GMM Estimation for firms aged 5 years or older

	First Stage	Second Stage			
Dependent Variable	Training firm	General Innovation	Product Innovation	Process Innovation	Patent Applications
Independent Variable	Coef.	Coef.	Coef.	Coef.	Coef.
Training firm		0.1856*	0.0888	0.0659	0.07922
		(0.1004)	(0.1157)	(0.1163)	(0.1019)
Firm Age	0.0021***				
	(0.0003)				
German-Speaking Firm	0.2402**				
	(0.0990)				
Workforce Education Controls	yes	yes			
Sector Controls	yes	yes			
Year Controls	yes	yes			
Region Controls	yes	yes			
Firm Controls	yes	yes			
Observations	2882	2882			
F-Statistic (joint significance of instruments)	32.143				
p-value	0.000				
Hansen J Statistic		0.060	0.752	1.315	2.618
p-value		0.8064	0.3859	0.2516	0.1056
Centered R-squared	0.1135	0.1065	0.1238	0.0703	0.1743

Note: Cluster robust standard errors in parenthesis (Cluster level: Firm).

* Statistically significant at the 0.1 level; ** at the 0.05 level; *** at the 0.01 level.

References

Acs, Zoltan J., and David B. Audretsch. 1988. Innovation in Large and Small Firms: An Empirical Analysis. American Economic Review 78, no. 4:678–90.

Aghion, Philippe. 2008. Higher Education and Innovation. Perspektiven der Wirtschaftspolitik 9, Special Issue:28–45.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. 2005. Competition and innovation: An inverted-u relationship. The Quarterly Journal of Economics 120, no. 2:701–28.

Aghion, Philippe, Leah Boustan, Caroline Hoxby, and Jerome Vandenbussche. 2005. Exploiting States' Mistakes to Identify the Causal Impact of Higher Education on Growth. Working Paper, Department of Economics, Harvard University.

Aghion, Philippe, and Peter Howitt. 2006. Appropriate Growth Policy: A Unifying Framework. Journal of the European Economic Association 4, nos. 2-3:269–314.

Angrist, Joshua D., and Alan B. Krueger. 2001. Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. Journal of Economic Perspectives 15, no. 4:69–85.

Backes-Gellner, Uschi. 1996. Betriebliche Bildungs- und Wettbewerbsstrategien im deutsch-britischen Vergleich. 1st ed. Munich and Mering: Rainer Hampp Verlag.

Bauder, Tibor, and Fritz Osterwalder. 2008. 75 Jahre eidgenössisches Berufsbildungsgesetz. 1st ed. Bern: h.e.p. verlag ag.

Baum, Christopher F., Mark E. Shaffer, and Steven Stillman. 2007. Enhanced routines for instrumental variables / generalized methods of moments estimation and testing. The Stata Journal 7, no. 4:465-506.

Bierhoff, Helvia, and S. J. Prais 1997. From school to productive work: Britain and Switzerland compared. 1st ed. Cambridge: Cambridge University Press.

Bundesversammlung der Schweizerischen Eidgenossenschaft 2002. Berufsbildungsgesetz: BBG.

Der Schweizerische Bundesrat 2003. Verordnung über die Berufsbildung: BVV.

Dosi, Giovanni. 1988. Sources, Procedures, and Microeconomic Effects of Innovation. Journal of Economic Literature 26, no. 3:1120–71.

Elektromobilität verbindet. 2015a. Standardisiertes Qualifizierungskonzept zur Integration der Hochvolt-Technik in die Duale Berusfausbildung. http://www.elektromobilitaet-verbindet.de/projekte/hochvolt-technik-berufsausbildung.html.

Elektromobilität verbindet. 2015b. Erarbeitung eines Qualifizierungsprozesses aufgrund gestiegener Anforderungen im Rahmen von Elektromobilität für neu zu qualifizierende Fachkräfte. http://www.elektromobilitaet-verbindet.de/projekte/erarbeitung-qualifizierungsprozess.html.

Elektromobilität verbindet. 2015c. Qualifizierung von Fachkräften im Kfz-Handwerk http://www.elektromobilitaet-verbindet.de/projekte/qualifizierung-von-fachkraeften-im-kfz-handwerk.html.

Fagerberg, Jan. 2005. Innovation: A guide to the literature. In The Oxford Handbook of Innovation, eds. Jan Fagerberg, David C. Mowery, and Richard R. Nelson. Oxford: Oxford University Press.

Gonon, Philipp, and Markus Maurer. 2012. Educational Policy Actors as Stakeholders in the Development of the Collective Skill System: The Case of Switzerland. In The Political Economy of Collective Skill Formation, eds. Marius R. Busemeyer and Christine Trampusch. New York: Oxford University Press.

Jensen, Morten Berg, Björn Johnson, Edward Lorenz, and Bengt Ake Lundvall 2007. Forms of knowledge and modes of innovation. Research Policy 36, no 5:680-93.

Kleibergen, Frank, and Mark E. Schaffer. 2010. ranktest: module for testing the rank of a matrix using the Kleibergen-Paap rk statistic. http://ideas.repec.org/c/boc/bocode/s456865.html

Knutti, Peter. 2007. Dokumentation Berufsbildung. http://www.doku.berufsbildung.ch/download/dokubb/index.htm.

Krueger, Dirk, and Krishna B. Kumar. 2004a. Skill-Specific rather than General Education: A Reason for US-Europe Growth Differences? Journal of Economic Growth 9, no. 2:167–207.

Krueger, Dirk, and Krishna B. Kumar. 2004b. US–Europe differences in technology-driven growth: quantifying the role of education. Journal of Monetary Economics 51, no. 1:161–90.

Lucas, Robert E. 2009. Ideas and Growth. Economica 76, no. 1:1–19.

Moulton, Brent. R. 1990. An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. The Review of Economics and Statistics 72, no. 2:334–38.

OPET 2010. Vocational and Professional Education and Training in Switzerland 2010: Facts and Figures. Federal Office for Professional Education and Technology OPET. Bern.

Pakes, Ariel, and Zvi Griliches. 1984. Patents and R&D at the Firm Level: A First Look. In R&D, Patents and Productivity, ed. Zvi Griliches. Chicago: University of Chicago Press.

Pedró, Francesc, Tracey Burns, Katerina Ananiadou, and Henri de Navacelle. 2009. Systemic Innovations in VET: OECD Studie zur Berufsbildung Schweiz. OECD.

Ryan, Paul, Karin Wagner, Silvia Teuber, and Uschi Backes-Gellner. 2010. Trainee Pay in Britain, Germany and Switzerland: Markets and Institutions. SKOPE Research Paper no. 96. ESRC Centre on Skills, Knowledge and Organisational Performance.

Staley, Mark. 2011. Growth and the diffusion of ideas. Journal of Mathematical Economics, 47: 470–478.

Stock, James H, Jonathan H. Wright, and Motohiro Yogo, M. 2002. A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. Journal of Business & Economic Statistics 20, no.4:518–29.

Stock, James H., and Motohiro Yogo. 2002. Testing for Weak Instruments in Linear IV Regression. Technical Working Paper no. 284, National Bureau of Economic Research, Cambridge, MA.

Vandenbussche, Jerome, Philippe Aghion, and Costas Meghir. 2004. Growth, Distance to Frontier and Composition of Human Capital. IFS Working Papers no. WP04/31. The Institute for Fiscal Studies.

Weissenberger-Eibl, Marion A., Rainer Frietsch, Hugo Hollanders, Christian Rammer, Peter Neuhaeusler, and Torben Schubert. 2011. Innovationsindikator. Troisdorf: Broermann.

Wettstein, Emil, and Philipp Gonon. 2009. Berufsbildung in der Schweiz. Bern: h.e.p. verlag ag.

Wolter, Stefan C., and Paul Ryan. 2011. Apprenticeship. In Handbook of the Economics of Education 3, eds. Eric A. Hanushek, Stephen Machin, and Ludger Woessmann. Amsterdam: Elsevier.