Open access • Journal Article • DOI:10.1007/S11573-018-0924-6

How firms' participation in apprenticeship training fosters knowledge diffusion and innovation — Source link

Christian Rupietta, Uschi Backes-Gellner

Institutions: University of Wuppertal, University of Zurich

Published on: 01 Jul 2019 - Journal of Business Economics (Springer Berlin Heidelberg)

Topics: Apprenticeship and Vocational education

Swiss Leading House
Economics of Education • Firm Behaviour • Training Policies

Working Paper No. 74

How firms' participation in apprenticeship training fosters knowledge diffusion and innovation

Christian Rupietta and Uschi Backes-Gellner

Universität Zürich
IBW – Institut für Betriebswirtschaftslehre

UNIVERSITÄT BERN

Working Paper No. 74

How firms' participation in apprenticeship training fosters knowledge diffusion and innovation

Christian Rupietta and Uschi Backes-Gellner

May 2018 (first version: March 2012)

A previous version of this Working Paper No. 74 was first published in March 2012 under the title "High quality workplace training and innovation in highly developed countries".

This is a post-peer-review, pre-copyedit version of an article published in Journal of Business Economics. The final authenticated version is available online at: https://doi.org/10.1007/s11573-018-0924-6

Published as: "How firms' participation in apprenticeship training fosters knowledge diffusion and innovation." *Journal of Business Economics, 89*(2019)5: 569-597. By Christian Rupietta and Uschi Backes-Gellner.

How firms’ participation in apprenticeship training fosters knowledge diffusion and innovation

Christian Rupietta

Schumpeter School of Business and Economics, University of Wuppertal
Gaussstrasse 20
DE 42119 Wuppertal
Tel: +49 (0) 202 439 3178
Email: rupietta@wiwi.uni-wuppertal.de

Uschi Backes-Gellner

Department of Business Administration, University of Zurich
Plattenstrasse 14
CH 8032 Zurich
Tel: +41 (0) 44 63 44 281
Email: backes-gellner@business.uzh.ch

Zurich, May 2018

Acknowledgement

We thank John Addison, Simone Balestra, Bernd Fitzenberger, Simon Janssen, Edward Lazear, Johannes Meuer, Kira Rupietta, Paul Ryan, Rainer Winkelmann and the participants of the annual meetings of the Canadian Economics Association, Swiss Society for Economics and Statistics, the Bildungsökonomischen Ausschusses im Verein für Socialpolitik, the Spring Meeting of Young Economists, the Colloquium on Personnel Economics, the DRUID Society Conference and the research seminars at the University of Zurich for helpful comments and suggestions. We acknowledge and thank Natalia Abrosimova and Jan Hagen for research assistance and Natalie Reid for editorial support. Furthermore we thank the Swiss Economic Institute (KOF) for data provision.

Funding

This study is partly funded by the Swiss State Secretariat for Education, Research and Innovation through its Leading House on the Economics of Education, Firm Behaviour and Training Policies.

How firms' participation in apprenticeship training fosters knowledge diffusion and innovation

Abstract

Previous studies typically relate apprenticeship training or more generally Vocational Education and Training (VET) to training that is highly specific and that uses well-established technologies. Accordingly, apprenticeship training is typically not expected to have positive effects on innovation. In contrast, we argue in this paper that the type of dual apprenticeship training seen in Switzerland (or Germany) does create positive innovation effects due to these VET-systems' built-in and institutionalized *curriculum development and updating processes.* These processes ensure that firms participating in apprenticeship training gain access to knowledge that is close to the innovation frontier and that ultimately fosters innovation. We provide theoretical explanations of how this knowledge diffusion works and how it can help to generate innovation in participating firms. We use the Swiss VET system as one example and derive hypotheses about the relationship between firms' participation in apprenticeship training and their innovation outcomes. Empirical analyses support our hypotheses. In a VET system with a built-in curriculum-updating process like the one in Switzerland (or Germany), firms participating in apprenticeship training have higher innovation outcomes than do non-participating firms.

Keywords: Vocational education, knowledge diffusion, education systems, innovation, empirical analysis, innovation policy

JEL Classification: I20, O31

I. Introduction

In past decades, researchers have associated innovation-relevant knowledge with university education (Aghion 2008; Aghion and Howitt, 2006). In contrast, vocational education and training (VET) was considered to be too firm-specific and to exclusively or mainly use knowledge of old and established technologies (Krueger & Kumar, 2004a, b). The assumption that VET is too firm-specific and outdated may indeed be true if the VET in question is organized and carried out independently by single firms.

We argue, however, that the assumption about high firm-specificity is incorrect in the case of a collectively organized VET system that has built-in curriculum-updating processes – as found in Switzerland or Germany. In such VET systems, occupational training curricula are designed and regularly updated based on national legal frameworks and with the support of industry organizations, social partners and government institutions. As many firms – particularly those at the innovation frontier – are involved in these curriculum-updating processes, the respective training curricula are never firm-specific, and they also cover the latest technologies, including technologies that are not yet used in the mainstream operations of companies. Due to legal regulations, the respective training curricula are also nationally binding, i.e., all firms offering apprenticeship training have to adhere to these curricula and students completing such apprenticeship programs gain a nationally recognized vocational training certificate that is valid in all firms (for more information see Backes-Gellner 1996, 1999, Backes-Gellner et al. 2016).

We argue that the built-in VET curriculum-updating process in Switzerland very much resembles the typical characteristics of any highly innovative institutional system. As innovation research has long shown, it is generally important for innovation to have a system (i.e. institutions and in-built mechanisms) that organizes inter-firm knowledge exchange (e.g., Carlsson & Stankiewicz, 1991; Dosi, 1982; Freeman, 1988; Lynn, Reddy & Aram, 1996;

Sahal, 1985; Watkins et al., 2015). Carlsson and Stankiewicz already argue in 1991 that there are two important institutional layers in highly innovative institutional systems: the super-structure and the sub-structure. The super-structure is an institution or organization that coordinates all involved firms, while the sub-structure consists of a critical number of firms that provide innovative inputs for the coordinating organization. We argue that the VET system such as in Switzerland or Germany provides such an innovation-enhancing super- and sub-structure: it has a curriculum-development-process that coordinates involved firms and it has firms that provide innovative inputs to the process. Accordingly, we expect that firms participating in the VET system by training apprentices and thereby extracting innovative inputs from the curricula and the trained apprentices have innovation advantages over firms that do not participate in apprenticeship training.

We test our hypotheses using Swiss data (the KOF innovation survey) and show that – as expected – firms' participating in apprenticeship training have higher innovation outcomes than do firms not participating. We also show that with respect to the number of apprentices the effect follows an inverted u-shape and that it is relatively stronger for smaller firms. We conduct several robustness tests, including an instrumental variables approach, to try to address potential endogeneity problems. We find that our results remain stable.

II. National education institutions, curriculum-updating, collective learning and innovation

A. Theoretical foundation to analyze innovative institutional frameworks

Since the 1980s several studies have highlighted the importance of institutional frameworks for innovation (e.g., Carlsson & Stankiewicz, 1991; Dosi, 1982; Freeman, 1988, Justman &

Teubal, 1995; Lynn et al., 1996; Lynn, Aram & Reddy, 1997; Watkins et al., 2015). In various theories, the super- and sub-structure relation is a key concept used to show the influence of the institutional framework on innovation. The first are Carlsson and Stankiewicz (1991). They define a technological system that generates innovation as a 'network of agents interacting in a specific economic/industrial area under a particular institutional infrastructure or set of infrastructures and involved in the generation, diffusion, and utilization of technology'. Lynn et al. (1997) characterize the sub- and super-structure as mutually dependent. The super-structure governs and coordinates collective action, while the sub-structure (firms) provides input for the super-structure. Typically, organizations such as trade organizations, employer associations or industrial associations coordinate the action of their members and regulate knowledge exchange.

While several studies apply the sub- and super-structure in the context of the development of new products and processes or the commercialization of innovation, no study explicitly applies the concept to a collectively organized VET system. However, as we show in the following, a collectively organized VET system with a built-in curriculum-updating process (as in Switzerland or Germany) provides a sub- and super-structure between governing organizations and participating firms similar to that found in other innovation frameworks that have been analyzed in previous literature.

B. Institutional framework of apprenticeship training in Switzerland: built-in curriculum-updating processes

Swiss dual apprenticeship training combines vocational schooling with workplace training of three to four years[1]. The training is based on nationally standardized curricula that ensure the transferability of skills between firms (Wolter & Ryan, 2011). These curricula regulate both education in vocational schools and training in the workplace. As qualification requirements constantly change, training curricula are regularly updated according to a well-defined institutional process, taking into account the most recent technological developments and innovation. In this process, government bodies together with employers' associations, firms and employee representatives decide - within a well-defined process - how to update the content of training curricula in their particular industries or occupations[2]. Firms at the innovation frontier are important players in this curriculum-updating process because they are already using new technologies or are the first to envision what new technologies may emerge and what skills will be needed accordingly. In addition, the elaborate updating process also ensures that the resulting training curricula contain not only the skill requirements of firms at the innovation frontier but also other, more mainstream firms in the industry. Thus, through this process, the Swiss VET system has the two important layers of an innovation governance system: the super-structure, in the form of industry organizations and government agents, and the sub-structure, which consists of a critical number of firms that provide innovative inputs for the coordinating organizations. How the two layers are set up and how they work together to collect and diffuse new knowledge within the curriculum-updating process will be explained in the following.

---

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

[2] Future orientation, and thus innovation, is a key component of the Swiss VET system (Bundesversammlung der Schweizerischen Eidgenossenschaft, 2002; Wolter & Ryan, 2011; Bauder & Osterwalder, 2008; Der Schweizerische Bundesrat, 2003)). For an overview of institutions in the Swiss VET system, see Pedró et al. (2009) and Rauner (2008).

C. Knowledge collection, synthesis, transfer and application in the Swiss VET system

To describe how the super-structure and the sub-structure interplay in the Swiss VET context, we use the model of knowledge creation and diffusion developed by Jensen et al. (2007).[3] This model distinguishes between two ideal-type modes of innovation, learning and knowledge diffusion: the Science, Technology, and Innovation (STI) mode and the Doing, Using, and Interacting (DUI) mode. These two modes differ in the types of knowledge processed and in the way in which these different types of knowledge are shared. Different types of knowledge are 'explicit versus implicit knowledge' and 'global versus local knowledge'. The different ways to share them are 'codification vs. personal interaction'.

The STI mode is characterized by the use of global and explicit knowledge. In the STI mode, knowledge is expressed and transferred in written form in scientific articles, manuals, reports and email messages; it is codified, and its primary function is to explain *why* something works in general but not necessarily *how* it works in a specific context. Scientists at universities and in R&D departments of firms typically use and produce such knowledge. The application of codified knowledge to local problems remains part of the STI mode as long as the scientist transfers knowledge in a codified form (Jensen et al., 2007).

In contrast, the DUI mode is characterized by the use of local and implicit knowledge. Employees share knowledge by closely interacting with colleagues, and the understanding of *how* something works is more important than an explicit explanation of *why* something works. The DUI mode also entails learning from and interacting with colleagues from different departments and backgrounds. Communication in these interactions is mainly informal.

Using the different knowledge types and sharing modes also helps to characterize (in a stylized way) the knowledge collection and diffusion mechanisms of the VET system (cf.

[3] Although the original model uses the firm as the level of analysis, the model is also valid for higher aggregation levels such as industries or nations (Jensen et al., 2007).

Figure 1). As we will see, both modes of innovation and learning are built into the Swiss VET-system.

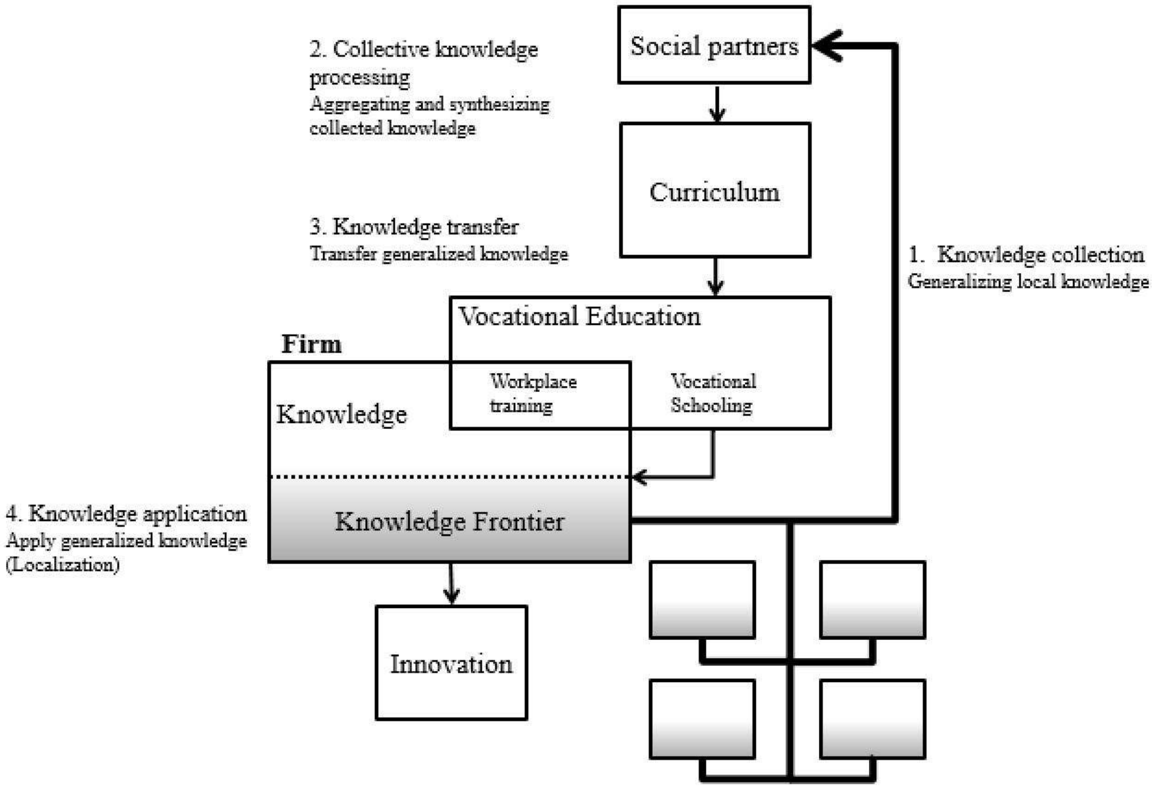

Figure 1: Knowledge-processing in Swiss VET by the built-in curriculum updating processes

First, in the knowledge collection step of the curriculum-updating process, there is implicit and local knowledge collected from innovative and other firms. In the second step, this knowledge is synthesized into updated training curricula, which contain new explicit and global knowledge. In the third step, the updated knowledge is then transferred to and in the fourth step it is applied in all participating firms. These four steps are regulated by the VET legal framework and by institutionalized processes, as explained in the following.

1. Knowledge collection

In Switzerland, there are nationally recognized VET curricula for approximately 220 separate occupations; these occupations cover the entire labor market for VET graduates, i.e., about two thirds of the Swiss working population. Each occupational curriculum includes the knowledge base of the respective industry. The knowledge base is incorporated into the curriculum development process by the respective industry organizations, who in the VET-context are also called social partners (Organisationen der Arbeitswelt).

> *To illustrate this we can look at two examples. For example, the training curriculum for the "polymechanic" apprenticeship program includes the knowledge base of the mechanical engineering industry. In the curriculum development process the knowledge base is represented by the two industry-specific social partners: "Swissmem" and "Swissmechanics". As another example, the training curriculum for the "health professional" apprenticeship program includes the knowledge base of the health sector, which is in the curriculum development process represented by the respective industry organization "OdA Santé".*

To define or update training standards, these social partners require knowledge inputs from the firms in their industry, and most importantly from the industry's most innovative firms because they are the most likely to know which skills will be required in the workplace of the future (given new technological developments and the most recent innovation trends).

The right side of Figure 1 illustrates the knowledge flows into the curriculum-updating process. These knowledge flows contain information on *what* apprentices should know, *how* they should perform certain tasks, and *why* something works. The revision of a curriculum requires a firm to communicate not only codified but also localized knowledge. Detailed descriptions of best practices require explanations of *how* a worker performs a task, *how* a firm organizes its workforce, *how* workers communicate, and *how* the task is embedded in other processes. When revising curricula, firms generalize their local knowledge to communicate best practices to their representatives.

2. Collective knowledge synthesis

In the next step (cf. upper part of figure 1) the social partners aggregate and standardize the collected knowledge to integrate it into an occupation-specific training curriculum (such as polymechanic or health professional).[4] In this step, the social partners have to generalize the originally firm-specific knowledge to occupation-specific knowledge because – according to legal regulations – apprenticeship programs are required to contain occupation-specific knowledge instead of firm-specific knowledge.[5]

3. Knowledge transfer

In the next step, through updated training curricula, the new occupation-specific knowledge is transferred to all firms that train apprentices. To ensure that instructors in vocational schools and in firms conduct training according to the updated curricula, the government mandates that instructors attend preparatory courses. Thus, the knowledge from the updated training curricula is transferred first to the instructors and then from the instructors to the apprentices and thereby also to the firms that were not directly part of the curriculum development process. During the knowledge transfer phase, the STI and the DUI modes are again combined. The training period in the vocational schools (20% of the time) focuses more on the transfer of knowledge that corresponds to the STI mode (science and technological base of innovations), while the on-the-job training period in the firms adds a DUI component (apprentices learn how to apply different types of knowledge by doing, using and interacting).

[4] An example of the revised curricula for commercial employees and knowledge aggregation appears in Pedró et al. (2009).
[5] Bundesversammlung der Schweizerischen Eidgenossenschaft (2002)

4. Knowledge application

Finally, applying the knowledge that is transferred to training firms via updated training curricula leads to an increased potential for innovation because of the benefits offered by the novel qualifications of the freshly trained young workers and of their instructors. With these novel qualifications, firms can more quickly absorb new knowledge, novel technological developments and other innovations. Thus, once again, the STI mode of learning and innovation is combined with the DUI mode: the highly generalized knowledge flowing in via updated curricula and the new knowledge of instructors and young workers are combined with local knowledge and thereby used to generate more learning and innovation in the training firm.

D. Implications: participating in apprenticeship training and innovation in firms

As shown above, firms that participate in apprenticeship training in the Swiss context indirectly have an additional knowledge inflow of new scientific knowledge and best practices. The further application of this new knowledge, i.e. the combination of codified innovative knowledge and learned practices that show how employees can apply this new knowledge, enables training firms to generate new local knowledge that fosters their innovative output. This leads to our first hypothesis:

H1: Firms that participate in apprenticeship training (train at least one apprentice) are more innovative than firms that do not participate (have no apprentices).

Given that every apprentice learns and understands differently and broadens the general knowledge base of a firm, we additionally expect that the innovation effect increases with training intensity, i.e., with the number of apprentices a firm trains. However, as the additional knowledge that each additional apprentice can add is limited, we expect decreasing marginal returns. Our second hypothesis (H2) thus states:

H2: Training intensity has a positive but diminishing effect on innovation output.

Our third and last implication concerns different types of firms. Firms that operate with the latest technologies and that have already established novel practices are the ones that provide inputs to the updating process of the respective occupational training curricula. These firms are often large or medium-sized, and they are leaders in the global or national markets. In contrast, the large number of small firms, on average, are less likely to be at the innovation frontier, and they operate less regularly with the latest technologies. We therefore expect curricula to affect small firms more strongly than large firms. Nevertheless, we still expect large firms to benefit from participating in apprenticeship training because they also have access to additional knowledge that other leading firms have transferred to the curriculum-updating process. Our third hypothesis (H3) thus states:

H3: The positive effect of training participation on innovation is stronger for small firms than for large firms.

III. Data

A. Sample and descriptive statistics

For our empirical analysis, we use the Swiss Innovation Survey, a triennial panel, of the Swiss Economic Institute (KOF). The data set contains several innovation indicators, such as binary indicators for process and product innovation and patent applications. We use these indicators as our dependent variable.[6] We find that more than 50% of the firms in our sample report either a process innovation or a product innovation, compared to a smaller number of firms that applied for a patent (19.2%). While process and product innovations might include innovations that are only new to the firm, innovations that are patentable might be either new to the world or new to the industry and are thus expected to be less likely in general.

To construct our main explanatory variable, we generate a binary variable entitled "training firm," which takes the value of one if a firm employs at least one apprentice and zero otherwise. We built the variable "training firm" from information provided about the educational composition of the workforce. The educational composition uses the following categories: 1. workers with university degrees, 2. workers with degrees higher than vocational education, 2. workers with vocational education degrees, 3. workers with degrees lower than vocational education, and 4. apprentices. In addition, we use a number of usual control variables. Descriptive statistics for all variables are given in Table 1 for our estimation sample.[7] Firms in our sample are, on average, larger than the average firm in Switzerland because the KOF survey applies a sampling scheme that oversamples large and medium-size firms. 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' innovation outcomes in general, not on a representative description of Switzerland.

[6] 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 it is an indicator for general innovation activity. Thus, we give this constructed measure the title of general innovation.

[7] To investigate the influence of vocational education on firms' innovation outcomes, 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. Furthermore, we restrict our sample to German-speaking regions, as vocational education is more widespread in these regions and because firms in the French-speaking part often follow a more consecutive VET approach that is closer to the French than the German training tradition.

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

B. Instrumental variables

As the training decision may be endogenous, we will also use an instrumental variable approach to try to solve potential endogeneity problems. The data set allows us to define two instruments, both of which are not ideal but better than nothing: first, a firm's location and firm language, which represents different apprenticeship training traditions in the different linguistic regions of Switzerland, and second, a firm's age, which represents a longer or shorter training tradition. We will briefly explain these two instruments in the following. First, because training traditions in the different language regions are similar to the training traditions in the respective neighboring countries of Germany, France or Italy, the training traditions of firms also differ in the three linguistic regions (French, German and Italian) in Switzerland. Firms in the German-speaking tradition are typically more likely to offer apprenticeship training than are firms in the French- or Italian-speaking tradition (Gonon & Maurer, 2012; OPET, 2010). As the data set contains information on the firms' language, we use this information to construct an instrumental variable for firms with a stronger vs. weaker apprenticeship training tradition.[8] If a firm's language is German we expect that - all else equal - they are more likely to train apprentices than if a firm's language is not German. Therefore we use company language as our first instrument. Second, as apprenticeship training in Switzerland has a long tradition and gained its highest popularity in the middle of the 20th century (Knutti, 2007), we expect older firms (that grew with these traditions) to have a training tradition that relies more heavily on apprenticeship training, compared with younger firms (that are more influenced by recent discussion on the advantage of academic education). Therefore, we choose firm age as a second instrument for firms' training participation.

[8] 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 can request a different linguistic version of the questionnaire. Using this information, we can construct a binary variable that indicates whether a firm's language is German or not. As our sample contains 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 who have French or Italian as their firm language.

IV. Estimation Strategy

For our empirical analysis of firm's innovation outcomes, we augment the knowledge production function proposed by Pakes and Griliches (1984). Their original model contains two equations: The first equation explains knowledge growth by a firm-specific productivity shifter, a time trend and research expenditures of past periods. The second equation uses patents as an outcome variable and explains it with knowledge growth and a time trend as input factors[9]. We augment this framework by introducing knowledge of VET training curricula as an additional input factor. 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 innovations are also considered outcomes of firms' knowledge growth.

We estimate a linear probability model because we have binary innovation measures taking the value of one if the firm innovated successfully during the respective period and zero otherwise. The advantage of the use of the linear probability model over probit models is the direct calculation of the marginal effects of the training decision on innovation. If we take endogenous training decisions into account, this feature allows the comparison of the results between the IV specification and the OLS specification.[10]

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

Equation 1 is our basic estimation equation. We use 4 different measures for innovation outcomes (general innovation, product innovation, process innovation and patent applications). The explanatory variable of interest is the binary indicator for training participation of firm j at time t: tr_{jt}. We also include a set of control variables. These control

[9] Due to missing observations for firms' R&D expenditures, we cannot include their lags in our model.
[10] Due to the panel structure of our data set, we risk getting biased standard errors if we 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.

variables include firm size, educational composition of a firm's workforce, competition measures such as price and non-price competition (Aghion, Bloom, Blundell, Griffith, & Howitt, 2005), lack of qualified workers, an indicator for foreign firms, and sector, year, and regional dummies.[11]

V. Results

A. OLS estimates

To test our hypothesis, we estimate equation 1 for our four innovation measures. According to our first hypothesis, we expect a firm's participation in training to have a positive impact on its innovation outcomes. 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 firm's decision to train apprentices. The results indicate that firms participating in training have a 7.8 percentage point-higher probability of innovating than firms not participating. Columns 2 and 3 show the estimation results for product innovation and process innovation. 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. 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 between a firm's training participation and its patent applications. The marginal effect of training on patent applications is 6.7 percentage points.

[11] Endogenous training decisions 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 test the robustness of the OLS estimates against endogenous training decisions with an instrumental variable strategy (Angrist & Krueger, 2001). As instrumental variables, we use the firm's age and firm's language to measure its training tradition.

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.

B. Effects of training intensity

The impact of participating in apprenticeship training on innovation might depend not only on the binary decision of whether a firm participates but also on the training intensity, i.e., number of apprentices employed in relation to the total workforce. According to our second hypothesis, we expect the relation between innovation and training intensity to follow an inverted u-shape. Table 3 thus provides estimation results regressing our four innovation indicators on the number of apprentices and the squared number of apprentices. All specifications include the full set of control variables.

We first find that participating in apprenticeship training is positively associated with all four innovation indicators and follows 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. 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 of their training intensity. These results are all in line with our theoretical expectations.

Table 3 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.

C. Effects of firm size

According to hypothesis three, we expect the positive effect of training participation on innovation to be larger for small firms than for large firms. Table 4 therefore presents the results for estimations with interaction terms between training participation and firm size. We find that the interaction of vocational education and firm size is negatively associated with innovation. However, these estimates are only statistically significant for patent applications (5 percent level). The results also show that participating in training is still positively associated with the four innovation measures. In general, larger firms thus seem to gain from participating in apprenticeship training, but their gain is comparatively weaker than the effect for smaller firms. These results support our hypothesis.

Table 4 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.

D. Instrumental variable estimates

As our estimations so far may suffer from endogeneity problems, we use our instrumental variable approach to try to account for this problem. Table 5 shows the estimation results of the first and second stage of a GMM estimation.[12] The first stage contains our two instruments: *Firm Age* and *German-speaking Firm*.

[12] We present the results of a GMM estimation in this section, as GMM is an efficient estimator if heteroscedasticity and clustering occur. We also run the IV estimations with TSLS and limited information maximum likelihood (LIML). LIML is more robust to weak instruments than the procedures mentioned above (Stock, Wight, & Yogo, 2002). The two alterative IV estimations generate results that are in line with those in table 5.

Table 5 shows statistically significant estimates at the 1 percent and 5 percent levels, respectively, and the expected signs for all of our instruments in the first stage. 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 of participating in vocational education than do 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 5 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.[13]

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 5, we find p-values above 0.33. Thus, the hypothesis that the instruments are valid cannot be rejected in these cases. Only in the specification shown in column 5, the patent applications, can we reject this hypothesis at the 10 percent level. Overall, the results of the second stage estimations in table 5 are in line with the results in table 2. They indicate a positive impact of firms' participation in training on general innovation. This effect is statistically significant at the 5 percent level. For the other innovation outcomes, the effects are not statistically significant but with the expected sign. Thus, these results also indicate that

[13] 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 less than 10 percent of the time. This test result supports the strength of our instruments.

our hypothesis cannot be rejected. In sum, the instrumental variable estimates support our previous finding: firm's participating in apprenticeship training have higher innovation outcomes than non-participating firms.

Table 5 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.

However, 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, and 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 years old. This restriction allows the exclusion of firms that are most likely to exit the market.

Table A1 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 years or older. This table shows similar results compared to table 5, where we use the entire sample. The first-stage estimates in table A1 show a positive effect of all instruments on firms' training participation at a statistical significance level similar to that reported in table 5. Furthermore, the results in table A1 support the second-stage results in table 5. 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' innovation outcomes.

VI. Discussion and conclusion

In this paper, we showed that firms' participation in dual apprenticeship training, as in Switzerland, creates positive innovation effects for these firms. This is because the Swiss VET system (similar to that in Germany) has a built-in *curriculum updating process.* This institutionalized updating process ensures that training curricula, as well as trainers and vocational schools, have knowledge close to the innovation frontier. The curricula diffuse this knowledge to all firms participating in apprenticeship training.

We provided theoretical explanations on how the whole knowledge collection, synthesis and diffusion process works and how it can help to generate innovation in training-firms.

Important characteristics are national legal frameworks, the engagement of industry organizations, social partners and government institutions, and – a particularly important characteristic - the involvement of firms close to the innovation frontier in the curriculum updating processes. With these characteristics, the built-in curriculum updating processes of the Swiss VET system resemble the characteristics of a typical highly innovative institutional system with a super- and sub-structure of knowledge exchange.

Our empirical analyses based on the KOF innovation survey confirm our theoretical expectations; our analyses show that firms participating in apprenticeship training have higher innovation outcomes than firms not participating. We also find that the effects are stronger for smaller firms but still exist for large firms. Estimations with an instrumental variable approach and a number of robustness tests confirm our results. Our model and our results have implications for education theory and educational policy makers and for a firms' innovation strategies.

First, regarding educational policy, by distinguishing different types of VET and their surrounding VET systems (firm organized firm-specific trainings vs. dual apprenticeship training in collectively organized VET systems), we can show that the impact of vocational education on innovation depends on the institutional design of the vocational education system. Furthermore, we show that the institutional design of collectively organized VET systems, such as those in Switzerland or Germany, shares important similarities with the organization of innovation communities (e.g., Carlsson & Stankiewicz, 1991; Dosi, 1982; Justman & Teubal, 1995; Lynn et al., 1996; Watkins et al., 2015). As vocational education has become more popular in recent years and in many countries, our results help to design VET systems that also support innovation in participating firms. Our model highlights important characteristics of the governance structure of such a VET system: it should consist of a super-structure and a sub-structure that support knowledge collection, synthesis and diffusion. The model explains how industry associations, employees' organisations and government agents together with firms (particularly innovative firms) can collaborate to initiate knowledge collection from firms at the innovation frontier, synthesize it in updated training curricula and help diffuse it to other firms by their participation in apprenticeship training.[14]

Second, our results have implications for firm's innovation strategies. We show that in a VET system with an effective curriculum-updating process, firms can gain innovation advantages by participating in apprenticeship training. This produces an additional and stand-alone incentive for firms to participate in apprenticeship training on top of any other benefits that may occur during and after training[15]. Our results also show that firms do not have to train at a very high intensity to gain such benefits. A large part of the innovation effect occurs due to participation itself but not so much from having a relatively large number of apprentices. Thus, firms considering participating in apprenticeship training should also take such innovation effects into account and not just short-term costs and immediate productivity contributions of apprentices. This can strengthen a firm's willingness to train and, at the same time, secure a critical mass of firms participating in the system.

To summarize, our paper shows that an adequate governance structure is key to an innovation-enhancing VET system, which most importantly supports knowledge diffusion across firms through an institutionalized systematic curriculum-updating process. Our paper thereby also adds to a currently limited strand of literature that analyses knowledge diffusion in detail and reveals mechanisms of how firms can generalize their implicit and local knowledge to synthesize it and diffuse it across all participating firms.

[14] Vocational education systems can also be an important part of a national innovation system and contribute in combination with other national institutions to the innovation outcomes of firms (Meuer, Rupietta & Backes-Gellner, 2015).

[15] Non-monetary benefits of (vocational) education for firms include for example effects on the productivity of co-workers (e.g., Backes-Gellner, Rupietta & Tuor Sartore, 2017) or on the diversity of a firm's knowledge base that improve a firm's innovation performance (e.g., Bolli, Renold & Wörter, 2018).

Appendix

Table A1 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.

References

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, 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. 1999. Humanressourcen im internationalen Wettbewerb. *Zeitschrift für Betriebswirtschaft*, 69(1999)1: 8-10.

Backes-Gellner, Uschi, Christian Rupietta, and Simone N. Tuor Sartore. 2017. Reversed educational spillovers at the firm level. *Evidence-based HRM: a Global Forum for Empirical Scholarship* 5, no. 1:80-106.

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

Bolli, Thomas, Ursula Renold, and Martin Wörter. 2018, Vertical educational diversity and innovation performance. *Economics of Innovation and New Technology* 27, no. 2:107-131.

Bundesversammlung der Schweizerischen Eidgenossenschaft 2002. *Berufsbildungsgesetz*: *BBG*.

Carlsson, Bo and R. Stankiewicz 1991. On the nature, function and composition of technological systems. *Journal of Evolutionary Economics* 1, no. 2:93-118.

Dosi, Giovanni. 1982. Technological paradigms and technological trajectories. *Research Policy* 11, no. 3:147-162.

Freeman, Christopher. 1988. Diffusion: The spread of new technology to firms, sectors and nations. In *Innovation, Technology and Finance*, ed. Arnold Heertje. Oxford: Basil Blackwell.

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.

Justman, Moshe, and Morris Teubal. 1995. Technological infrastructure policy (TIP): creating capabilities and building markets. *Research Policy* 24, no. 2:259-281.

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.

Lynn, Leonard H., John D. Aram, and N. Mohan Reddy. 1997. Technology communities and innovation communities. *Journal of Engineering and Technology Management* 14, no. 2:129-145.

Lynn, Leonard H., N. Mohan Reddy, and John D. Aram. 1996. Linking technology and institutions: the innovation community framework. *Research Policy* 25, no. 1:91-106.

Meuer, Johannes, Christian Rupietta, and Uschi Backes-Gellner. 2015. Layers of co-existing innovation systems. *Research Policy* 44, no. 4:888-910.

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

Rauner, Felix. 2008. *Steuerung der beruflichen Bildung im internationalen Vergleich.* Kurzfassung. Gütersloh: Bertelsmann Stiftung.

Sahal, Devendra. 1985. Technological guideposts and innovation avenues. *Research Policy* 14, no. 2:61-82.

Watkins, Andrew, Theo Papaioannou, Julius Mugwagwa, and Dinar Kale. 2015. National innovation systems and the intermediary role of industry associations in building institutional capabilities for innovation in developing countries: A critical review of the literature. *Research Policy* 44, no. 8:1407-1418.

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.