Training and Innovation

Stefan Bauernschuster+, Oliver Falck*, Stephan Heblich‡

Accepted at the Journal of Human Capital

Abstract

Research analyzing the importance of human capital for innovation usually focuses on formal secondary and tertiary education. This paper takes a different perspective and focuses on human capital arising from in-firm training. We argue that continuous training guarantees access to leading-edge knowledge and thus increases a firm's propensity to innovate. To test this hypothesis empirically, we use German establishment-level data. Our results show a strong association between lagged continuous training and innovation. Based on the results of an instrumental variable approach, we cautiously argue that the association between lagged continuous training and innovation is indeed a causal effect. Our instrumental variable approach exploits provisions of the German Works Constitution Act, allowing us to use works councils as a relevant and valid instrument for continuous training conditional on a well-specified set of covariates derived from legal regulations.

JEL Classification: J 24, L 11, O31
Keywords: Innovation, training, works council

+ Ifo Institute for Economic Research, Poschingerstr. 5, D-81679 Munich (Germany), Phone: +49 89 9224 1368, Fax: +49 89 9224 1604, Email: bauernschuster@ifo.de.

* Ifo Institute for Economic Research, Poschingerstr. 5, D-81679 Munich (Germany), Phone: +49 89 9224 1370, Fax: +49 89 9224 1460, Email: falck@ifo.de, CESifo.

‡ Max Planck Institute of Economics, Entrepreneurship, Growth, and Public Policy Group, Kahlaischestr. 10, D-07745 Jena (Germany), Phone: +49 3641 686 733, Fax: +49 3641 686 710, Email: heblich@econ.mpg.de.

Acknowledgments: We are indebted to John Addison, David Audretsch, Lutz Bellmann, Ron Boschma, Oliver Kirchkamp, Chiara Monfardini, Steffen Mueller, Guido Schwerdt, Rainer Winkelmann, Stefan Wolter, the participants of the Swiss Leading House conference on "Economics of Education, Firm Behavior and Training Policies" in Zurich, 2008, the participants of the DRUID Conference in Copenhagen, 2008, and the EALE Conference in Tallinn, 2009, the editors Isaac Ehrlich and Francis T. Lui, and two anonymous referees for insightful comments that helped improve the paper. The provision of data by the Institute for Employment Research is gratefully acknowledged.

1. Introduction

Innovation is a well-accepted driver of economic growth and development and the key determinant underlying the innovation process is assumed to be human capital.[1] Accordingly, theories of endogenous growth (cf. Lucas 1988; Romer 1990; Aghion and Howitt 1998) do not limit the effects of human capital to increasing labor productivity only, but also view them as increasing the innovative capacity of the economy as a whole, in the form of new processes and products. The most common indicators of human capital are the amount and quality of schooling. Hanushek and Woessmann (2008) analyze a large variety of countries and find a positive relation between long-term growth and the quality of secondary schooling. Going one step further, Aghion *et al.* (2009a, 2009b) find that it is tertiary education, that is, education that takes place in colleges and universities, rather than primary and secondary education that is most supportive of leading-edge innovation and growth. The common basis of both studies is their focus on human capital investments undertaken *before* entering the labor market.

However, as pointed out by Arrow (1962), many skills are best learned on the job, i.e., *after* entering the labor market. Taking into consideration that due to the rapidly changing environment of today's world in which human capital derived from formal education (schooling, vocational education) depreciates quickly, learning by doing, in the form of in-firm training, may be an additional way to continue to accumulate leading-edge knowledge.[2] And, indeed, we observe that a large portion of human capital investment occurs within firms in the form of training (Acemoglu 1997), which is usually, at least to some extent, co-financed by employers (Pischke 2001). Based on a large German establishment-level dataset, this paper analyzes the impact of training on innovation, which to date, to our knowledge, has

[1] While human capital refers to the key input variable in the innovation process, the ultimate commercialization of the innovation largely depends on the market structure. Cohen and Levin (1989) provide a rich over view of literature in this field. More recently, contributions following the ideas of Aghion and Howitt (1998) have further formalized and tested the Schumpeterian view on innovation and market structure. In this paper, we concentrate on human capital as major asset in the innovation process within a competitive environment.
[2] For a formal synthesis of learning and training, see Killingsworth (1982).

been neglected in the literature on human capital and innovation but could be of particular importance for certain kinds of innovation. Note that we are not engaged in analyzing the impact of innovation on the necessity to train workers in new technologies or processes; this direction of causality is already addressed in literature on skill-biased technological change (e.g., Autor *et al.* 2003; Bresnahan *et al.* 2002).

In the training literature, Becker's (1964) initial contribution argues that firms will invest in training only if they can appropriate its future rent, i.e., the workers' higher productivity. Acemoglu and Pischke (1999a) extend Becker's reasoning and argue that noncompetitive labor markets, in combination with a compressed wage structure, can provide an incentive for firm-sponsored training because firms can appropriate parts of the expected rent.[3] In this paper, we extend this line of work by considering the product market and argue that innovations are a way to keep up with the technology frontier in a competitive environment and that firm-sponsored training increases a firm's likelihood to come up with certain kinds of innovation.[4] This is because trained workers who have leading-edge knowledge understand complex products and production processes and thus are more likely to come up with technological improvements. Our argument further suggests that training is especially important in the case of so-called routine innovations, that is, those that involve significant improvements to existing products or processes, whereas the creation of something radically

[3] Possible explanations for a compressed wage structure include transaction costs, such as search and matching frictions (Mortensen 1982; Diamond 1982); asymmetric information about the worker's true level of training (Katz and Zidermann 1990; Chang and Wang 1996); asymmetric information about an applicant's—particularly a young applicant without a comprehensive work record—motivation to apply for a new job (Is the applicant one of low ability who has been fired from a previous job or is he or she an underpaid high-ability worker?) (Acemoglu and Pischke 1998); complementarities between the training of specific and general skills (Acemoglu and Pischke 1999b); and given labor market institutions, such as minimum wages or labor unions (Acemoglu and Pischke 1999b, 2003; Freeman and Medoff 1984).

[4] Some related literature analyses the connection between product market competition and training (cf. Bassanini and Brunello 2010; Gersbach and Schmutzler 2003; Goerlitz and Stiebale 2008). This literature argues that product market competition increases training incentives because training is a way to increase workers' productivity while we argue that trained workers are more likely to come up with an innovation. Our idea is to analyze the effect of training within the framework of the literature on competition and innovation by arguing that innovation is the weapon to fight competition while training is the ammunition (cf. Aghion *et al.* 2005).

new might require additional skills like creativity and inventive talent, which cannot be taught.

In identifying the causal effect of training on innovation, our chief concern is omitted variable bias because we cannot, for example, completely control for the competitive environment, firms' organizational structures, and management practices, which also might drive firm training.[5] To overcome this problem, we apply an instrumental variable approach and instrument training with the existence of a works council. Works councils are worker representatives within the establishment; they have no bargaining power, but are legally entitled to encourage and take part in the decision-making as to training for workers. We exploit provisions of the German Works Constitution Act, which lead to works councils being independent of potential outcomes in firms with at least five workers conditional on a well-specified set of covariates derived from legal regulations. Workers of every firm with at least five workers may, but are not required to, have a works council. A works council can be set up on the initiative of three workers. Assuming that the share of works-council-prone workers is equally distributed across firms, firms with a works council can be regarded as a random sample of all firms, conditional on firm size and some other controls arising from the Works Constitution Act. By instrumenting training with the existence of a works council, we can estimate the local average treatment effect (LATE) of training on innovation.

The results of our empirical analyses show a strong association between lagged continuous training and innovation. Applying instrumental variables (IV) techniques to overcome any potential omitted variable bias, we find that IV estimates do not significantly differ from simple probit and OLS specifications, possibly indicating that omitted variables are only weakly correlated or uncorrelated with lagged continuous training and innovation. Consequently, we cautiously conclude that our probit and OLS coefficients might indeed be

[5] See Bloom and van Reenen (2007) for a recent survey on the difficulties in measuring firms' management practices.

interpreted as being of a causal nature, i.e., continuous training of workers increases an establishment's propensity to innovate.

The remainder of the paper is organized as follows. Section 2 introduces our empirical strategy and discusses the validity of our instrument in more detail. Section 3 introduces the data and Section 4 presents the results, along with several robustness checks. Section 5 concludes with some implications for further research.

2. Identification Strategy for the Causal Effect of Continuous Training on Innovation

To empirically analyze the relationship between training and a firm's propensity to innovate, we start out with simple probit and linear probability models. We are interested in the effect of training on innovation. Since it is equally plausible to assume that causality runs in the other direction, that is, it could be that it is innovation that is driving the need for training (cf. Autor *et al.* 2003; Bresnahan *et al.* 2002), we lag the right-hand-side training variable. Additionally, we focus on continuous training, i.e., training offered constantly over the years, instead of training at a single point in time. This two-fold strategy should help overcome any reverse causality problems.

Although we might be able to control for a wide range of firm-level and industry-level characteristics in a multivariate regression framework, we are still concerned about an omitted variable bias. Unobserved characteristics that are correlated with the training variable and at the same time with our outcome variable, innovation, might lead to a correlation between training and innovation that cannot be interpreted as an unbiased causal effect.[6] This is why, in a next step, we apply instrumental variable techniques in order to identify the causal effect

[6] For example, Bloom and van Reenen (2007) show that management matters for firm profitability, productivity, Tobin's Q, and firm survival. In a similar line of argument, we might suspect that unobserved management characteristics also matter for the provision of training and at the same time for a firm's propensity to innovate, which would leave us with omitted variable bias. Consequently, the obtained coefficient is just a biased estimate of the real causal effect.

of interest. In particular, we instrument training by the existence of a works council in the firm. Works councils are worker representatives elected by a firm's workers. Once a works council is established, the works council has several information and consultation rights, as well as veto and co-determination rights in personnel matters, e.g., in the case of layoffs. However, works councils do not have any wage bargaining rights.

For IV to have a causal interpretation, this instrument has to be relevant and independent of potential outcomes (conditional on certain covariates). Relevance means that the instrument has a clear effect on the endogenous variable. Sections 96–98 of the German Works Constitution Act (*Betriebsverfassungsgesetz*) state that works councils are legally entitled to encourage and take part in decision-making as to worker training. Thus, the existence of a works council should affect the provision of training in the firm, an argument that is formalized in Acemoglu and Pischke (1999b). Moreover, empirical studies for Germany confirm the theoretical prediction that works councils have a strong impact on a firm's decision to train (cf. Bellmann and Leber 2005; Neubaeumer and Kohaut 2008). Thus, works councils should be a relevant instrument for training.

Independence is found if the instrument is as good as randomly assigned, i.e., independent of potential outcomes, conditional on certain covariates. Here, again, the legal regulations on works councils are of great interest because they give rise to a quasi-experimental setting. Section 1 of the German Works Constitution Act states that works councils can be founded in establishments with five and more workers who have voting rights, i.e., the employed are at least 18 years old. To found a works council, at least three workers must announce a worker meeting at which an election committee is formed. Of course, the bigger an establishment, the more likely it is that there will be three employed who are eager to fulfill this task. However, the worker meeting can also be announced by a labor union that is present at the establishment (§17(3)). Alternatively, the general works council of the corporation (if it exists) can set up the election committee (§17(1)). Once a works council exists at the establishment level, §16 makes it is very likely that it will continue to exist because that section provides that the works council can, in effect, reestablish itself. One may therefore conclude that the older an establishment, the more likely it is, *ceteris paribus*, that a works council is established. For our identification strategy to work, it is crucial that employers must not hinder the founding of a works council. Indeed, employers are subject to severe penalties if they attempt to interfere with works council formation. All these regulations lead us to the conclusion that the existence of a works council is a random event, once we control for a well-specified set of covariates derived from legal regulations, namely, establishment size and age, the existence of a labor union, and the branch plant status, i.e., whether the establishment is part of multi-establishment firm or is a single-establishment firm. And, in fact, empirical evidence for Germany supports our theoretical arguments as it shows that apart from establishment size and age, branch plant status, and union density, virtually no variable can be found that systematically determines the existence of works councils (cf. Addison *et al*. 1997, 2003). Thus, we are confident that for establishments with at least five workers, works councils are indeed independent of potential outcomes, conditional on the discussed covariates.

In a recent study, Jirjahn (2009) shows that works councils are often established during bad economic times when the employed are fearful of being laid off. In such a situation, a credit-constrained establishment should also have difficulty financing R&D that could translate into future innovations, resulting in a negative association between the existence of a works council and future innovation that works through a channel other than training. Since this would downward-bias our estimates, our IV approach will actually estimate a lower bound of the causal impact of training on innovation. Apart from this, we assume that works councils do not work through other unknown channels to affect a firm's propensity to innovate.[7]

The continuous training variable and the innovation variable are both binary. Therefore, we might be tempted to use nonlinear models to analyze the determinants of a firm's propensity to innovate and to train continuously. Thus, continuous training could be the independent variable of the innovation probit model and the dependent variable of the second probit model, i.e., continuous training is endogenized in this system of equations. However, such a nonlinear model cannot be solved in a two-stage instrumental variable framework. A feasible way to handle this problem is to employ a recursive bivariate probit model where the error terms of the two probit models are allowed to be correlated (Evans and Schwab 1995). In this seemingly unrelated bivariate probit model, the probit equations on training and innovation are estimated simultaneously, as described in the equations below, where I(.) is the indicator function taking the value 1 if its argument is true and the value 0 otherwise. *IN* stands for the innovation dummy, *CT* for a dummy indicating continuous training, *S* is establishment size, and *A* is establishment age. *BP* represents a variable that captures the branch plant status, *U* is a dummy variable signifying whether the establishment is bound to a union contract, and *WC* shows the existence of a works council in the establishment; e_1 and e_2 are the error terms of the specific equation.

$$\mathrm{IN} = \mathrm{I}\ \left(1 \middle|\ \mathrm{CT, S, A, BP, U}, e_1\right) \qquad (1)$$

$$\mathrm{CT} = \mathrm{I}\ \left(1 \middle|\ \mathrm{WC, S, A, BP, U}, e_2\right) \qquad (2)$$

$$\operatorname{cov}\left[e_1, e_2\right] = \rho \qquad (3)$$

[7] Addison *et al*. (2004) provide an extensive overview of the research into the economic consequences of works councils. Comparing older research that finds rather negative effects with new research discovering some positive effects, Addison *et al*. argue that these results have to be taken "with more than a pinch of statistical salt" and suggest that, on average, the effects of works councils are zero.

Angrist and Krueger (2001) suggest estimating a linear instrumental variable two-stage least squares regression even if the endogenous regressor is a dummy variable. Using probit or logit to generate first-stage predicted values is not necessary and may even do some harm. Kelejian (1971) shows that consistency of second-stage estimates is not dependent on the functional form of the first stage being correct. What is more, computing predicted values in a nonlinear first stage, which are then plugged in at the second stage, does not result in consistent estimates unless the nonlinear model happens to be exactly right (Angrist and Krueger 2001). To avoid problems arising from misspecification of the first stage, we prefer a linear instrumental variable specification where innovation is used as a dependent variable and the endogenous variable, continuous training, is instrumented by the existence of a works council. As discussed above, we control for establishment size and age, branch plant status, and union contract.

We do not assume homogenous treatment effects; rather, what we estimate in our IV approach is a local average treatment effect (LATE). For IV to give us LATE, we assume monotonicity (Angrist and Imbens 1994), i.e., we have no establishments that have works councils and do not train, but would conduct training in the absence of a works council. Note that causal inference is driven by our instrument works council while the variable of interest remains training. We might think of this strategy as a causal chain where a works council affects training, which in turn affects innovation. Put differently, we only use the variation in training that is induced by the exogenous variation in the presence of works councils. Consequently, we identify the causal effect of training for those firms that would have trained their workers in the presence of a works council and would not have done so without a works council (Angrist and Imbens 1994). Without making further assumptions (e.g., constant causal effects), LATE cannot give us information about causal effects for subpopulations other than this complier subpopulation (Angrist and Krueger 2001). Different valid instruments for the same causal relation may provide similar or different results depending on special characteristics of the exogenous variation in training employed, and we thus reiterate that we have a strong claim for internal validity, i.e., for the causal effect of the kind of training that is induced by works councils. We solve the first-order problem of omitted variable bias for this well-defined subpopulation. However, we do not claim the same degree of external validity. The existence of heterogeneous treatment effects calls for more IV approaches to estimate the effect of training on innovation.

3. Data on Innovation and Training

Information on innovative activity, continuous training, and additional firm-level characteristics is generated from the IAB establishment panel (*Betriebspanel*), waves 1997–2001. For a detailed description of this data source, see Bellmann (2002). Access to the data was granted via on-site use at the Research Data Centre (*FDZ*) of the Research Data Centre of the Federal Employment Agency (*BA*) at the Institute for Employment Research (*IAB*) and via controlled data teleprocessing at the *FDZ*. As the name "establishment panel" implies, it is the establishment, not the company, that is the unit of measurement. Thus, we have two categories of entities: firm headquarters and subsidiaries. Establishments contained in the German Social Insurance Statistics form the population of the IAB establishment panel. The establishments are selected according to the principle of optimum stratification of the random sample. The stratification cells are defined by establishment size categories and industries. The establishment panel data encompass the results of annual surveys of establishments that have been carried out in West Germany since 1993 and in East Germany since 1996. The annual surveys cover questions on a series of characteristics. Additional complexes of questions dealing with special topics, such as working time flexibility, elder workers, or innovative activities, are included in selected annual catalogues.

To analyze the impact of continuous training on an establishment's propensity to innovate, we use data for the period 1997–2001. Only those establishments that answered the questionnaire in every year of this period are included in our dataset. Furthermore, the whole public sector is excluded, resulting in a balanced panel of 3,198 private-sector establishments for the period 1997–2001. This represents a uniquely rich source of data for our analysis. For our IV estimations, we drop those 632 establishments with on average less than five workers, since no law explicitly entitles workers of these establishments to set up works councils, and thus their inclusion would invalidate our random assignment assumption.

Information on innovative activity is available for the year 2001. In this year, the establishments were asked whether they introduced a completely new product/service during the past two years, whether they newly adopted a product/service, or whether they enhanced an existing product/service. Strictly speaking, only the first category (introduction of a completely new product/service) can be called a true innovation. However, for our analysis, innovation is more broadly defined and the innovation variable is given a value of unity if an establishment carried out any of the above-mentioned innovative activities; 0 otherwise.

Information on training is drawn from the 1997, 1999, and 2000 surveys. The interviewed establishments were asked whether or not training for their workers was encouraged either by (partly) financing the training or by releasing the employed from work to attend training. The question referred to the first half of every year. If an establishment promoted training, *T*, in all the years *t*, the variable *CT* (i.e., continuous training) takes the value of unity, otherwise it is 0.

$$CT = \begin{cases} 1 & \text{if } T_t = 1 \ \forall\, t \in \{1997, 1999, 2000\} \\ 0 & \text{otherwise} \end{cases} \qquad (4)$$

Information on our instrumental variable, i.e., the existence of a works council, is not available for all establishments during the entire period of observation; however, these data are available for the years 1998 and 2000. We took data from these two years to create a variable that takes the value of unity if a works council existed in both years and 0 otherwise.

Information on establishment size and age, branch plant status, and union contract is available from the establishment panel and is introduced in our model in the form of control variables. The average number of workers in the period 1997–2000 is computed and transformed into 10 firm-size classes to capture establishment-size effects. A dummy variable is used to capture the age of the establishment, which is 0 if it was established before 1990 and 1 if it was established in 1990 or later. Additionally, we use a dichotomous variable to capture the branch plant status, i.e., whether the establishment was part of a multi-establishment firm or was a single firm from 1998 to 2000. Another control variable takes on the value of unity if the establishment was tied to a union contract for at least three years between 1997 and 2000 (cf. Neubaeumer and Kohaut 2008). The data would allow distinguishing between sectoral and firm-level union contracts, but we decided not to use this variation because our estimation results turned out to be unaffected by these alternative measures. Summary statistics for all control variables can be found in the Appendix.

Table 1 gives a quick overview of the association between continuous training and an establishment's innovative activities. A simple computation of relative frequencies suggests that continuous training of workers is positively correlated with a firm's innovation. In particular, those establishments that continuously trained their workers during the period 1997–2000 exhibited more innovative activity from 1999 to 2001. While only 28 percent of the establishments that did not continuously train reported innovative activity, this number more than doubles and rises to 59 percent for the establishments that continuously train their workers. This correlation is supported across the single establishment-size classes. Fisher's exact test confirms that innovation is not independent from continuous training (p-value 0.000).

<< Insert Table 1 about here >>

In the remainder of this paper, we first test whether this relationship continues to hold in multivariate settings and then analyze the causal effect of training on a firm's propensity to innovate by using instrumental variable techniques.

4. Evidence for the Effect of Continuous Training on Innovation

4.1 Association Between Continuous Training and Innovation in a Multivariate Setting

Following the procedure outlined in Section 2, we start with estimating a probit model and a linear probability model with innovation as the dependent variable that signifies whether the establishment undertook any kind of innovative activity between 1999 and 2001. As the main regressor of interest we use continuous training in 1997, 1999, and 2000. All models in this section are estimated using cluster robust standard errors, where each federal state x industry cell forms one cluster.

Column 1 of Table 2 presents the results of a simple probit model in a baseline specification, where we include only those control variables that will later be important for our instrumental variables approach. We report the marginal effects at the sample mean and find a positive association between lagged continuous training and an establishment's propensity to innovate, an association that is highly significant. Furthermore, it can be seen that the bigger and, interestingly, younger an establishment is, the more likely it is to innovate. Whether an establishment is part of a multi-establishment firm or a single firm has no impact on the propensity to innovate, and neither does the existence of a union contract. Similarly, we find no significant direct impact of a works council on innovation. Similar results are obtained with a linear probability regression using the same specification (cf. Column 5 of Table 2). Indeed, comparing marginal effects from the probit model to the OLS coefficients shows that there is hardly any difference between the two methods. In an extended specification, we include 17 federal state and 23 industry dummies. Neither the probit estimates (cf. Column 2 of Table 2), nor the OLS estimates (cf. Column 6 of Table 2) are significantly affected by the inclusion of these additional controls.

<< Insert Table 2 about here >>

We are interested in establishing the role of training in innovation independent of prior educational attainment. If training is offered to highly skilled workers that are more likely to come up with an innovation then not controlling for educational attainments prior to training, i.e., schooling, would upward bias our results. However, we believe that in our setting not controlling for workers' schooling should be a minor issue because we analyze whether an establishment offers training in general but not to whom. Nevertheless, we are cautious and acknowledge that controlling for workers' schooling enhances the reliability of our estimates suggesting an independent role of training in innovation. Unfortunately, our establishment-level data do not include direct information about schooling of every single worker. Still, to come as close as possible to controlling for workers' schooling, we can include measures for workers' occupational status. The occupational status should at least to some extend be correlated with schooling because different occupations demand different minimum requirements of schooling.[8] Therefore, we include the average share of all blue and white collar workers with vocational training, university degree and/or relevant job specific experience over the total number of workers in an establishment in the years 1997 to 2000. Additionally, we include the average ratio of part-time workers in the years from 1997 to 2000 since part-time workers are on average less skilled than full-time workers (Hirsch 2005). Since innovative employed with a university degree can likely be found in the R&D department of an establishment (Falck et al. 2008), we add a dummy variable that indicates whether an establishment had an R&D department in 1998. This variable not only proxies workers with a university degree, but it also constitutes an indication for continuously having

[8] The correlation between schooling and occupational status should be especially strong in Germany for at least two reasons: First, Germany has a highly tracked secondary school system and mobility between different secondary school types is quite limited (Juerges and Schneider 2007). Along this line, Dustmann (2004) shows that different secondary school tracks translate into substantial wage differentials (probably due to different occupational status) later in life. Second, labor and product markets are highly regulated. Consequently, several tasks are only fulfilled by individuals with a certain amount or quality of schooling and, eventually, occupational mobility is low (Prantl 2009).

R&D projects. Since one could argue that firms that frequently innovate and/or expect to innovate in the near future because of their R&D projects are more likely to offer training, the inclusion of this control should further strengthen our point and speak out against reverse causality concerns.

To account for further heterogeneity among establishments, we also include dummies indicating the technological condition of an establishment's machines in 1997, whether an establishment invested in information and telecommunication technologies whether it invested in production technologies in the years 1997-2000, and whether it undertook any organizational changes in 1998. The results of the specifications including all these controls are presented in Columns 3 and 7 of Table 2. Finally, in Columns 4 and 8, we add a dichotomous variable indicating whether an establishment introduced an innovation in the period from 1996 to 1998; this should again diminish any remaining reverse causality concerns. It should be noted that many of these establishment-level controls might again suffer from endogeneity and consequently distort a clear causal analysis. Nevertheless, it is at least encouraging to see that the positive association between training and innovation is confirmed throughout all specifications (see Table 2). Though not presented in the table, the coefficient on the average ratio of skilled workers comes out insignificant whereas the dummy variable indicating the presence of an R&D department is positive and highly significant. A negative and highly significant association is found for the average ratio of part-time workers while the dichotomous variable capturing past innovative activity is positively associated with an establishment's propensity to innovate. Despite the vast amount of controls, the training coefficient remains positive and highly significant. In particular, introducing the indicator variable for past innovation in Columns 4 and 8 hardly affects the training coefficient at all. However, note that all these results might still suffer from omitted variable problems if there are any other unobservable covariates that are correlated with both the provision of continuous training and an establishment's innovations. Therefore, any problem arising from unobserved heterogeneity is directly addressed by instrumental variable techniques.

4.2. Instrumental Variable Results

Since the estimates from Table 2 could still suffer from omitted variable bias, we apply instrumental variable methods to examine the unbiased causal effect of training on a firm's propensity to innovate. As explained above, we instrument continuous training by the existence of a works council and control for establishment size and age, the existence of a labor union, and branch plant status. Workers are not legally entitled to set up a works council in establishments with less than five workers. In these establishments, employers might hinder the founding of a works council depending on unobserved employer characteristics. This challenges our assumption that the existence of works councils is random, given some observed control variables. Accordingly, we drop those 632 establishments that employ on average less than five workers. Using this subsample and conditioning on the discussed covariates, we are confident that works councils are indeed a random event, i.e., independent of potential outcomes. The results of the instrumental variable two-stage least squares estimations are presented in Table 3, where Columns 1 and 2 present results of the basic specification and Columns 3 and 4 show the results of a specification including federal state and industry dummies. Looking at the results of the first stage, we see that works councils have a strong positive impact on the provision of training. Indeed, F-statistics of the excluded instrument of 38.38 (p-value .000) and 23.91 (p-value .000), respectively, indicate that there is not a weak-instrument problem (Stock *et al.* 2002). Thus, we meet one important assumption of any instrumental variable approach, namely, that the instrument is relevant. Additionally, we conclude from the first stage that bigger establishments are more likely to train their workers. Furthermore, establishments that are part of a multi-establishment firm do more training than their counterparts. Moreover, and unsurprisingly, establishments bound to a union contract are more likely to train their workers; nearly every German union contracts includes provisions for worker training. The first-stage fitted values are then plugged into the second-stage equation to obtain the causal effect of lagged continuous training on a firm's propensity to innovate. From Column 2 of Table 3 we see that training has a significant and positive causal effect on innovation. Indeed, the coefficient even appears to have increased compared to the respective OLS results from Table 2. After adding federal state and industry controls to the two-stage procedure, the coefficient of training becomes smaller and insignificant, yet remains positive.

Since we do not impose the strict assumption of homogeneous treatment effects, we argue that what we identify in our IV approach is a LATE, i.e., the effect of training on innovation for the complier subpopulation. In our case, the complier subpopulation comprises those establishments that train their workers in the presence of a works council but would not do so otherwise. To arrive at a more precise picture of the compliant subpopulation, we can compute the proportion of the treated who are compliers. This figure is given by the first stage, times the probability the instrument *WC* (works council) is switched on, divided by the proportion of those undergoing the treatment *CT*, i.e., the proportion of establishments that continuously train their workers.

$$P\left[CT_{1i} > CT_{0i} \mid CT_1 = 1\right] = \frac{P\left[WC_i = 1\right]\left(E\left[CT_i \mid WC_i = 1, X\right] - E\left[CT_i \mid WC_i = 0, X\right]\right)}{P\left[CT_i = 1\right]} \quad (5)$$

Taking the first-stage works council coefficient of the basic specification presented in Column 1 of Table 3, we find that 15.96 percent of those establishments that continuously provide training for their workers do so due to the existence of a works council.

In a next step, we run Durbin-Wu-Hausman chi² tests, which basically allow us to compare the IV coefficients with simple OLS coefficients. Doing this reveals that the IV estimates do not differ significantly from simple OLS regressions of training on innovation in our data. The null hypothesis of an exogenous regressor cannot be rejected in either specification.

<< Insert Table 3 about here >>

Despite our endogeneity concern with further firm-level controls, we estimate IV regressions including all the additional covariates included in Table 2. The results are not presented here since the general picture is not affected by this modification: Again, F-statistics confirm that our instrument is strong (Stock *et al.* 2002). The training coefficient remains insignificant; yet, Durbin-Wu-Hausman chi² tests reinforce the previous finding that the null hypothesis of an exogenous regressor cannot be rejected, i.e., the IV estimates do not statistically differ from the respective OLS estimates. This finding suggests that there is little bias from omitted variables in the OLS estimate, probably because omitted variables in the innovation equation are only weakly correlated or uncorrelated with continuous training, at least for our complier subpopulation. Given this result of the Durbin-Wu-Hausman chi² tests and acknowledging the fact that IV estimates are almost always less efficient than OLS estimates, we argue that we can cautiously interpret the original uninstrumented estimates to show a positive causal effect of lagged continuous training on an establishment's propensity to innovate.

4.3 Robustness Checks

As suggested by Angrist and Krueger (2001), we employ a fully linear IV specification even though the endogenous regressor is a dummy variable. Nevertheless, to check the robustness of our results, we also estimate an IV probit model, where we linearize the first stage in order to avoid severe problems arising from potential misspecification but estimate a probit second-stage model. Additionally, we run a seemingly unrelated bivariate probit model, as described above, where we allow for correlation of the error terms of the two equations. The results are presented in Table 4. Qualitatively, our earlier findings are confirmed. In both models, the coefficient of lagged training on innovation is positive, yet only significant in the IV probit regression. Running a Wald test, we find that the hypothesis of exogeneity cannot be rejected (p-value: 0.336). Similarly, for the seemingly unrelated bivariate probit model, a Wald test (p-value: 0.627) as well as a likelihood ratio test (p-value: 0.611) cannot reject the hypothesis that the errors of the two equations are uncorrelated. This again can be seen as evidence that omitted variables are only weakly correlated or uncorrelated with training and innovation (Monfardini and Radice 2008). Consequently, we have additional support for our argument that the training coefficients of our original uninstrumented models might indeed be cautiously interpreted as being of a causal nature.

<< Insert Table 4 about here >>

In a next step, we make use of information available in our dataset on training intensity in addition to the simple training dummies. This training intensity variable, which is available only for 1997 and 1999, indicates the ratio of workers trained (or how many instances of training occurred) in the first half of the respective year. We take the average training intensity over these two years and include this variable in place of simple training dummies in our regressions. Our results from the continuous training dummy can also be transferred to training intensity (see Table 5). Training intensity exhibits a highly significant positive correlation with an establishment's propensity to innovative in simple probit and OLS models. Again, we observe that the resulting probit marginal effects hardly differ from the OLS coefficients. Applying instrumental variable techniques, where training intensity is instrumented by the existence of a works council, makes the training coefficient even larger. The estimated coefficient of the lagged continuous training variable is 1.11, i.e., a 10 percentage point increase in training intensity translates into an 11 percentage point higher propensity to innovate. An F-statistic for the excluded instrument of 15.55 indicates that works council can again be regarded as have a strong influence. Running a Durbin-Wu-Hausman chi² test reveals that we can reject the hypothesis that OLS and IV coefficients do not differ (p-value: 0.082). This indicates that, if anything, we *underestimate* the causal effect of training intensity on an establishment's propensity to innovate in simple probit and OLS estimations. Next, we drop those establishments that do not offer any training and analyze whether increasing training intensity has a positive effect on an establishment's propensity to innovate in this subsample. Interestingly, our instrument now becomes weak, as indicated by an F-statistic for the excluded instrument of 4.81. Based on the Works Constitution Act and earlier empirical work, we suggested that works councils can foster the provision of training. However, whether an individual worker actually takes part in a provided training program is beyond the control of a works council. Consequently, conditional on the general provision of training, the correlation between the existence of a works council and training intensity is weak, which is why in this paper we prefer the dichotomous training variable over the training intensity variable. Moreover, the dichotomous variable should be more reliable and suffer less from measurement error than the training intensity variable since the former is easier for firms to report.

<< Insert Table 5 about here >>

Finally, we examine whether continuous training favors radical innovations, which, in our context, means the introduction of completely new products or services. A binary-coded variable is generated that takes the value of unity only if the establishment reported having undertaken such a “real” innovation in the period 1999–2001. The reference group is comprised of establishments that introduced minor innovations during this same period, which, in our context, means imitation or enhancement of an existing product/service.

<< Insert Table 6 about here >>

It turns out that neither in the probit and OLS estimations nor in the two-stage least squares instrumental variables estimation presented in Table 6 can continuous training explain the creation of radical innovations as compared to minor innovations. It seems that a firm cannot systematically increase the propensity to radically innovate by the provision of training once it is a minor innovator. What is more, the bad fit of the models indicates that we could not identify factors at work for radical innovations as compared to product enhancement or imitation. We suggest that in order to increase the propensity of radical innovations, establishments have to rely on the creativity, skill, and genius of their workers, as well as their willingness to cooperate in teams, all of which might require outside-the-box-thinking. This sort of “soft” factor is difficult to analyze in our observational data; however, this line of argument could inspire further research into the determinants of radical innovation.

5. Discussion

This paper’s goal is to test the importance of training on a firm’s innovativeness. We suggest that trained workers with leading-edge knowledge understand complex products and production processes. This, in turn, increases their probability of coming up with innovations. Testing this empirically, we find support for the hypothesis that a firm’s investment in continuous training raises the firm’s probability to innovate. Moreover, we claim that we made some progress in estimating effects of training on innovation that are plausibly causal. However, our results do not allow for distinguishing between the impact of continuous training on radical innovations as compared to product enhancement or imitation.

Our findings contribute to the literature on the effects of human capital on innovation and growth. This strand of literature distinguishes investments in primary/secondary education from those in tertiary education but has to date neglected investment in training (Aghion *et al*. 2009a, 2009b). The main result of this research is that in advanced economies, investment in tertiary education increases patenting of inventions. We extend the literature by examining education that takes place *after* entering the labor market, instead of that occurring prior thereto. In doing so, we argue that firms that constantly train their workforce are more likely to maintain their position at the technological frontier because the leading-edge knowledge gained or increased from in-firm training supports additional innovative activity. To measure innovation, we rely on firms’ self-reports, which is arguably a somewhat crude measure of innovation. However, given our assumption that training is especially supportive of nonpatentable product enhancement and imitation, our measure seems more appropriate than a patent-based measure that might fail to recognize the effect of training on a firm's nonpatentable innovation.

Our study has two important policy implications. First, it shows the value of works councils, which have an indirect effect on innovation via continuous training. This indirect effect is neglected in previous literature that finds no direct effect of works councils on productivity growth and innovation. Second, it illuminates the association between education and innovation (growth) and thus adds to the discussion about which education investments affect innovation and hence economic growth.

References

Acemoglu, Daron, and Joern S. Pischke. 1998. "Why Do Firms Train? Theory and Evidence." *Quarterly Journal of Economics* 113 (1): 79–119.

Acemoglu, Daron, and Joern S. Pischke. 1999a. "Beyond Becker: Training in Imperfect Labour Markets." *Economic Journal* 109 (453): F112–F142.

Acemoglu, Daron, and Joern S. Pischke. 1999b. "The Structure of Wages and Investment in General Training." *Journal of Political Economy* 107 (3): 539–572.

Acemoglu, Daron, and Joern S. Pischke. 2003. "Minimum Wages and On-the-Job Training." *Research in Labor Economics* 22: 159–202.

Acemoglu, Daron. 1997. "Training and Innovation in an Imperfect Labour Market." *Review of Economic Studies* 64 (3): 445–464.

Addison, John, Claus Schnabel, and Joachim Wagner. 1997. "On the Determinants of Mandatory Works Councils in Germany." *Industrial Relations* 36 (4): 419–445.

Addison, John, Claus Schnabel, and Joachim Wagner. 2004. "The Course of Research into the Economic Consequences of German Works Councils." *British Journal of Industrial Relations* 42 (22): 255–281.

Addison, John, Lutz Bellmann, Claus Schnabel, and Joachim Wagner. 2003. "German Works Councils Old and New: Incidents, Coverage, and Determinants." *Schmollers Jahrbuch* 123 (3): 339–358.

Aghion, Philippe, and Peter Howitt. 1998. *Endogenous Growth Theory*. Cambridge: MIT Press.

Aghion, Philippe, Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. 2005. "Competition and Innovation: An Inverted-U Relationship." *Quarterly Journal of Economics* 120 (2): 701–728.

Aghion, Philippe, Mathias Dewatripont, Caroline M. Hoxby, Andreu Mas-Colell, and Andre Sapir. 2009a. "The Governance and Performance of Research Universities: Evidence from Europe and the U.S." NBER Working Paper 14851.

Aghion, Philippe, Caroline M. Hoxby, and Jerome Vandenbussche. 2009b. "The Causal Impact of Education on Economic Growth: Evidence from U.S." Brookings Papers on Economic Activity.

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 (4): 69–85.

Angrist, Joshua D., and Guido Imbens. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467–475.

Arrow, Kenneth. 1962. "The Economic Implications of Learning by Doing." *Review of Economic Studies* 29 (3): 155–173.

Autor, David, Frank Levy, and Richard Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118 (4): 1279–1333.

Bassanini, Andrea and Giorgio Brunello. 2010. "Barriers to Entry, Deregulation and Workplace Training: A Theoretical Model with Evidence from Europe." CES ifo Working Paper 2945.

Becker, Gary S. 1964. *Human Capital*. Chicago: University of Chicago Press.

Bellmann, Lutz, and Ute Leber. 2005. "Betriebliche Weiterbildung im regionalen Kontext." In *Bildungsökonomische Analysen mit Mikrodaten*, edited by Lutz Bellmann and Dieter Sadowski. Nuernberg: Beiträge zur Arbeitsmarkt- und Berufsforschung.

Bellmann, Lutz. 2002. "Das IAB-Betriebspanel: Konzeption und Anwendungsbereiche." *Allgemeines Statistisches Archiv* 86 (2): 177–188.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices Across Firms and Countries." *Quarterly Journal of Economics* 122 (4): 1351–1408.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *Quarterly Journal of Economics* 117 (1): 339–376.

Chang, Chun, and Yijiang Wang. 1996. "Human Capital Investment Under Asymmetric Information: The Pigovian Conjecture Revisited." *Journal of Labor Economics*, 14 (3): 505–519.

Cohen, Wesley M., and Richard C. Levin. 1989. "Empirical Studies of Innovation and Market Structure." In *Handbook of Industrial Organization*, edited by Richard Schmalensee and Robert D. Willig. New York: North-Holland.

Diamond, Peter A. 1982. "Wage Determination and Efficiency in Search Equilibrium." *Review of Economic Studies* 49 (2): 217–227.

Dustmann, Christian. 2004. "Parental Background, Secondary School Track Choice, and Wages." *Oxford Economic Papers*, 56(2): 209-230.

Evans, William N., and Robert M. Schwab. 1995. "Finishing High School and Starting College: Do Catholic Schools Make a Difference?" *Quarterly Journal of Economics* 110 (4): 941–974.

Falck, Oliver, Stefan Kipar, and Ludger Woessmann. 2008. "Humankapital und Innovationstätigkeit von Unternehmen.“ *Ifo Schnelldienst*, 61(7): 10-16.

Freeman, Richard B., and James L. Medoff. 1984. *What Do Unions Do?* New York: Basic Books.

Gersbach, Hans, and Armin Schmutzler. 2001. "A product market theory of training." IZA Discussion Paper 327.

Goerlitz, Katja and Joel Stiebale. 2008. "Does Product Market Competition Decrease Employers' Training Investments? Evidence from German Establishment Panel Data." Ruhr Economic Paper No. 41

Hanushek, Eric A., and Ludger Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature* 46 (3): 607–668.

Hirsch, Barry T. 2005. "Why Do Part-Time Workers Earn Less? The Role of Worker and Job Skills." *Industrial and Labor Relations Review* 58(4): 525–551.

Jirjahn, Uwe. 2009. "The Introduction of Works Councils in German Establishments—Rent Seeking or Rent Protection?" *British Journal of Industrial Relations* 47 (3): 521–545.

Juerges, Hendrik, and Kerstin Schneider. 2007. "What Can GoWrong Will GoWrong: Birthday Effects and Early Tracking in the German School System." Tech. Rep. 138, Mannheim Research Institute for the Economics of Aging.

Katz, Eliakim, and Adrian Zidermann. 1990. "Investment in General Training: The Role of Information and Labour Mobility." *Economic Journal* 100 (403): 1147–1158.

Kelejian, Harry H. 1971. "Two-Stage Least Squares and Econometric Systems Linear in Parameters but Nonlinear in the Endogenous Variables." *Journal of the American Statistical Association* 66 (334): 373–374.

Killingsworth, Mark R. 1982. "'Learning by Doing' and 'Investment in Training': A Synthesis of Two 'Rival' Models of the Life Cycle." *Review of Economic Studies* 49 (2): 263–271.

Lucas, Robert E. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (1): 3–42.

Monfardini, Chiara, and Rosalba Radice. 2008. "Testing Exogeneity in the Bivariate Probit Model: A Monte Carlo Study." *Oxford Bulletin of Economics and Statistics* 70 (2): 271–282.

Mortensen, Dale T. 1982. "Property Rights and Efficiency in Mating, Racing and Related Games." *American Economic Review* 72 (5): 968–979.

Neubaeumer, Renate, and Susanne Kohaut. 2008. "A Double Hurdle Approach for Company Further Training Behaviour." *Zeitschrift für ArbeitsmarktForschung* 40 (2-3): 251–269.

Pischke, Joern S. 2001. "Continuous Training in Germany." *Journal of Population Economics* 14 (3): 523–548.

Prantl, Susanne, and Alexandra Spitz-Oener. 2009. "How Does Entry Regulation Influence Entry into Self-employment and Occupational Mobility?" *Economics of Transition,* 17(4): 769-802.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98 (5): S71–S102.

Stock, James H., Jonathan H. Wright, and Motohiro Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business and Economic Statistics* 20 (4): 518–529.

Table 1: Cross tables on continuous training and innovations across size classes

Size class	Continuous training	Innovation 1999–2001 No	Yes	Total
Total	No	72.45	27.55	100.00
	Yes	41.07	58.93	100.00
	Total	59.10	40.90	100.00
0–9	No	78.60	21.40	100.00
	Yes	63.01	36.99	100.00
	Total	76.51	23.49	100.00
10–49	No	67.36	32.64	100.00
	Yes	53.99	46.01	100.00
	Total	63.04	36.96	100.00
50–249	No	60.91	39.09	100.00
	Yes	43.89	56.11	100.00
	Total	50.31	49.69	100.00
250–999	No	/	/	100.00
	Yes	33.23	66.77	100.00
	Total	38.14	61.86	100.00
1,000 and more	No	/	/	100.00
	Yes	15.17	84.83	100.00
	Total	16.20	83.80	100.00

Notes: *Figures are percentage shares; / signifies data anonymized due to low observation numbers.*

Table 2: Determinants of innovations: Simple probit and OLS regressions

	Probit marginal effect on INNOVATION				OLS coefficient on INNOVATION			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lagged continuous training	.164 ***	.165 ***	.098 ***	.093 ***	.159 ***	.148 ***	.086 ***	.077 ***
	(.025)	(.026)	(.026)	(.027)	(.025)	(.024)	(.022)	(.022)
Average no. of workers (Baseline: 0–4 workers)								
5–9 workers	.087 ***	.076 **	.034	.014	.067 **	.061 **	.023	.010
	(.034)	(.034)	(.036)	(.036)	(.026)	(.026)	(.025)	(.025)
10–24 workers	.130 ***	.126 ***	.061	.034	.104 ***	.101 ***	.045	.025
	(.038)	(.039)	.(044)	(.042)	(.031)	(.030)	(.031)	(.029)
25–49 workers	.217 ***	.215 ***	.091 *	.074	.185 ***	.179 ***	.074 **	.058 *
	(.042)	(.043)	(.047)	(.046)	(.037)	(.036)	(.035)	(.034)
50–99 workers	.232 ***	.233 ***	.095 *	.060	.201 ***	.196 ***	.078 **	.050
	(.045)	(.045)	(.050)	(.051)	(.042)	(.039)	(.038)	(.038)
100–249 workers	.255 ***	.229 ***	.057	.029	.225 ***	.195 ***	.052	.030
	(.056)	(.056)	(.060)	(.059)	(.054)	(.048)	(.046)	(.044)
250–499 workers	.309 ***	.295 ***	.121 **	.088	.283 ***	.254 ***	.102 **	.075 *
	(.053)	(.055)	(.061)	(.061)	(.055)	(.050)	(.045)	(.044)
500–999 workers	.328 ***	.290 ***	.075	.037	.305 ***	.250 ***	.069	.039
	(.062)	(.067)	(.077)	(.079)	(.067)	(.061)	(.057)	(.057)
1,000–1,999 workers	.475 ***	.447 ***	.254 ***	.221 ***	.468 ***	.382 ***	.174 ***	.143 ***
	(.042)	(.051)	(.076)	(.084)	(.056)	(.056)	(.053)	(.054)
2,000 and more workers	.534 ***	.545 ***	.374 ***	.301 ***	.545 ***	.482 ***	.239 ***	.181 ***
	(.034)	(.035)	(.078)	(.092)	(.056)	(.056)	(.056)	(.056)
Founded after 1990	.059 ***	.095 ***	.075 ***	.060 ***	.051 ***	.081 ***	.062 ***	.047 ***
	(.020)	(.022)	(.023)	(.022)	(.018)	(.018)	(.017)	(.016)
Branch plant status	.044	.025	.035	0.029	.039	.018	.022	.017
	(.031)	(.030)	(.030)	(.031)	(.028)	(.025)	(.023)	(.023)
Works council	.028	-.006	-.053	-.057 *	.027	-.003	-.039	-.042
	(.032)	(.033)	(.034)	(.034)	(.031)	(.029)	(.028)	(.026)
Union contract	-.040	-.025	-.020	-.005	-.036	-.022	.019	-.008
	(.025)	(.027)	(.028)	(.028)	(.022)	(.023)	(.021)	(.020)
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Federal state dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Industry dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Further firm level controls	No	No	Yes	Yes	No	No	Yes	Yes
Past innovation dummy	No	No	No	Yes	No	No	No	Yes
No. of observations	3,067	3,056	2,960	2,954	3,067	3,067	2,971	2,965
Wald chi^2	310.07	780.92	851.41	1040.35				
Prob > chi^2	.000	.000	.000	.000				
Pseudo R^2	.1145	.1839	.2534	.2878				
R^2					.1483	.2322	.3052	.3441

Notes: *The table reports probit and OLS regressions of lagged continuous training on innovation. For the probit model, marginal effects are reported at the sample mean. Further firm level controls include the average ratio of skilled workers in the years 1997-2000, the average ratio of part-time workers in the years 1997-200, the technological condition of the machines in 1997, whether an establishment invested in information and telecommunication technologies and whether it invested in production technologies in the years 1997-2000, whether it undertook any organizational changes in 1998 and whether it had an R&D department in 1998. In columns (2), (3), and (4) the numbers of observations from the probit models differ from the respective numbers of the OLS regressions because all observations from the shipbuilding and aircraft construction industry were dropped due to a lack of variation in our outcome variable for this industry.*

**** 1% level of significance, ** 5% level of significance, * 10% level of significance; clustering robust standard errors in parentheses.*

Table 3: IV regressions

	2SLS		2SLS	
	First stage TRAINING (1)	Second stage INNOVATION (2)	First stage TRAINING (3)	Second stage INNOVATION (4)
Works council	.197 ***		.160 ***	
	(.026)		(.027)	
Lagged continuous training		.328 **		.138
		(.163)		(.185)
Average no. of workers (Baseline: 5–9 workers)				
10–24 workers	.055 **	.027	.065 **	.042
	(.027)	(.032)	(.027)	(.031)
25–49 workers	.132 ***	.093 **	.127 ***	.120 ***
	(.030)	(.040)	(.029)	(.039)
50–99 workers	.211 ***	.093	.216 ***	.136 **
	(.033)	(.059)	(.032)	(.061)
100–249 workers	.328 ***	.094	.333 ***	.129
	(.035)	(.086)	(.035)	(.090)
250–499 workers	.432 ***	.134	.421 ***	.185 *
	(.041)	(.098)	(.040)	(.102)
500–999 workers	.474 ***	.148	.471 ***	.183
	(.047)	(.116)	(.047)	(.119)
1,000–1,999 workers	.502 ***	.304 ***	.472 ***	.303 **
	(.048)	(.116)	(.048)	(.118)
2,000 and more workers	.512 ***	.378 ***	.457 ***	.398 ***
	(.059)	(.115)	(.060)	(.113)
Founded after 1990	.017	.034	-.016	.071 ***
	(.018)	(.022)	(.020)	(.023)
Branch plant status	.119 ***	.020	.093 ***	.016
	(.022)	(.038)	(.022)	(.034)
Union contract	.041 **	-.043 *	.040 *	-.018
	(.020)	(.026)	(.020)	(.027)
Constant	Yes	Yes	Yes	Yes
Federal state dummies	No	No	Yes	Yes
Industry dummies	No	No	Yes	Yes
No. of observations	2,453	2,453	2,453	2,453
R^2	.3238	.1104	.3699	.2307
Durbin-Wu-Hausman test of exogeneity:				
chi^2		1.090		.010
p-value		.297		.921

Notes: *The table reports 2SLS regressions of lagged continuous training on innovation where training is instrumented by the existence of a works council. Columns 1 and 2 present the basic specification without federal state and industry dummies, whereas these are included in the results reported in Columns 3 and 4.*
**** 1% level of significance, ** 5% level of significance, * 10% level of significance; clustering robust standard errors in parentheses.*

Table 4: IV probit regression and seemingly unrelated bivariate probit

	IV probit marginal effects		Seemingly unrelated biprobit
	First stage	Second stage	Second stage marginal effect on
	TRAINING	INNOVATION	INNOVATION
	(1)	(2)	(3)
Works council	.197 ***		
	(.032)		
Lagged continuous training		.327 **	.060
		(.144)	(.078)
Average no. of workers			
(Baseline: 5–9 workers)			
10–24 workers	.055 *	.032	.061 **
	(.030)	(.037)	(.029)
25–49 workers	.131 ***	.105 **	.162 ***
	(.035)	(.044)	(.032)
50–99 workers	.211 ***	.103	.219 ***
	(.036)	(.063)	(.042)
100–249 workers	.328 ***	.103	.302 ***
	(.039)	(.091)	(.059)
250–499 workers	.432 ***	.142	.410 ***
	(.040)	(.103)	(.063)
500–999 workers	.474 ***	.156	.455 ***
	(.041)	(.120)	(.072)
1,000–1,999 workers	.502 ***	.325 ***	.614 ***
	(.039)	(.106)	(.041)
2,000 and more workers	.512 ***	.409 ***	.672 ***
	(.039)	(.085)	(.031)
Founded after 1990	.017	.037	.027
	(.021)	(.023)	(.018)
Branch plant status	.119 ***	.024	.109 ***
	(.022)	(.041)	(.025)
Union contract	.041 *	-.048 *	.005
	(.024)	(.027)	(.022)
Constant	Yes	Yes	Yes
Federal state dummies	No	No	No
Industry dummies	No	No	No
No. of observations	2,453		2,453
Wald chi²	221.38		618.66
Prob>chi²	.000		.000
Wald test of exogeneity:			
chi²	.93		.235
Prob>chi²	.336		.627
Likelihood ratio test of exogeneity:			
chi²			.258
Prob>chi²			.612

Notes: *The table reports IV probit and seemingly unrelated biprobit regressions of lagged continuous training on innovation where training is instrumented by the existence of a works council. The training equation of the seemingly unrelated biprobit model has the same specification as the IV probit first stage; we therefore do not report it in the table. Marginal effects at the sample mean are reported for the second stages of the IV probit and biprobit models.*

**** 1% level of significance, ** 5% level of significance, * 10% level of significance; clustering robust standard errors in parentheses.*

Table 5: Probit, OLS, and IV regressions: Training intensity and innovation

	Probit marginal effect on INNOVATION	OLS coefficient on INNOVATION	2SLS First stage TRAINING	2SLS Second stage INNOVATION
	(1)	(2)	(3)	(4)
Works council	.048	.048 *	.061 ***	
	(.032)	(.032)	(.016)	
Lagged training intensity	.276 ***	.246 ***		1.110 **
	(.042)	(.035)		(.552)
Average no. of workers				
5–9 workers	.102 ***	.081 ***		
	(.033)	(.026)		
10–24 workers	.162 ***	.131 ***	-.020	.066 **
	(.037)	(.030)	(.016)	(.031)
25–49 workers	.260 ***	.224 ***	-.007	.147 ***
	(.040)	(.036)	(.017)	(.036)
50–99 workers	.287 ***	.252 ***	-.020	.183 ***
	(.042)	(.040)	(.019)	(.040)
100–249 workers	.327 ***	.295 ***	-.026	.229 ***
	(.050)	(.051)	(.018)	(.044)
250–499 workers	.378 ***	.352 ***	.020	.244 ***
	(.046)	(.051)	(.033)	(.058)
500–999 workers	.379 ***	.359 ***	.008	.262 ***
	(.056)	(.066)	(.032)	(.069)
1,000–1,999 workers	.518 ***	.540 ***	.051 *	.403 ***
	(.035)	(.055)	(.029)	(.077)
2,000 and more workers	.556 ***	.604 ***	.065 *	.453 ***
	(.031)	(.059)	(.035)	(.087)
Founded after 1990	.049 **	.043 **	.045 ***	-.008
	(.020)	(.018)	(.013)	(.032)
Branch plant status	.026	.025	.118 ***	-.076
	(.029)	(.027)	(.019)	(.079)
Union contract	-.046 *	-.042 *	.037 ***	-.071 **
	(.025)	(.022)	(.013)	(.034)
Constant	Yes	Yes	Yes	Yes
Federal state dummies	No	No	No	No
Industry dummies	No	No	No	No
No. of observations	3,009	3,009	2,397	2,397
Wald chi²	273.61			
Prob>chi²	.000			
R^2		.1389		
Pseudo R^2	.1081			
Uncentered R^2			.4697	.4091
Durbin-Wu-Hausman test of exogeneity:				
chi²				3.026
Prob>chi²				.082

Notes: *The table reports probit, OLS, and 2SLS regressions of lagged training intensity on innovation where training intensity (the avg. ratio of workers trained within the first half of the years 1997 and 1999) is instrumented by the existence of a works council in the 2SLS specification. The baseline size class is "1–4* workers*" for the OLS and probit model and "5–9* workers*" for the 2SLS model. Marginal effects at the sample mean are reported for the probit model.*
**** 1% level of significance, ** 5% level of significance, * 10% level of significance; clustering robust standard errors in parentheses.*

Table 6: Determinants of radical innovations: Probit, OLS, and IV regressions

	Probit marginal effect on RADICAL INNOVATION	OLS coefficient on RADICAL INNOVATION	2SLS First stage TRAINING	2SLS Second stage RADICAL INNOVATION
	(1)	(2)	(3)	(4)
Works council	.005	.006	.154 ***	
	(.040)	(.042)	(.039)	
Lagged continuous training	.016	.014		.033
	(.029)	(.028)		(.275)
Average no. of workers				
5–9 workers	-.005	-.003		
	(.049)	(.040)		
10–24 workers	.027	.023	.040	.023
	(.054)	(.044)	(.047)	(.043)
25–49 workers	.074	.064	.056	.065
	(.056)	(.045)	(.048)	(.051)
50–99 workers	.058	.050	.187 ***	.048
	(.068)	(.054)	(.052)	(.078)
100–249 workers	.064	.055	.317 ***	.051
	(.072)	(.060)	(.054)	(.129)
250–499 workers	.095	.083	.422 ***	.077
	(.081)	(.067)	(.060)	(.162)
500–999 workers	.038	.032	.456 ***	.025
	(.084)	(.072)	(.067)	(.169)
1,000–1,999 workers	.181 **	.163 **	.461 ***	.156
	(.092)	(.071)	(.064)	(.171)
2,000 and more workers	.032	.027	.447	.021
	(.085)	(.073)	(.072)	(.169)
Founded after 1990	.029	.029	.034	.028
	(.029)	(.029)	(.027)	(.029)
Branch plant status	-.012	-.012	.080 ***	-.018
	(.026)	(.027)	(.028)	(.037)
Union contract	.010	.010	.063 **	.018
	(.027)	(.027)	(.030)	(.035)
Constant	Yes	Yes	Yes	Yes
Federal state dummies	No	No	No	No
Industry dummies	No	No	No	No
No. of observations	1,249	1,249	1,123	1,123
R^2		.0147	.3210	.0137
Pseudo R^2	.0148			

Notes: *The table reports probit, OLS, and 2SLS regressions of lagged continuous training on radical innovation (i.e., no imitation or enhancement of an existing product/service) where training is instrumented by the existence of a works council in the 2SLS specification. The baseline size class is "1–4* workers*" for the OLS and probit model and "5–9* workerss*" for the 2SLS model. Marginal effects at the sample mean are reported for the probit model.*

**** 1% level of significance, **5 % level of significance, * 10% level of significance; clustering robust standard errors in parentheses.*

Appendix

A.1: Distribution of establishments across industries

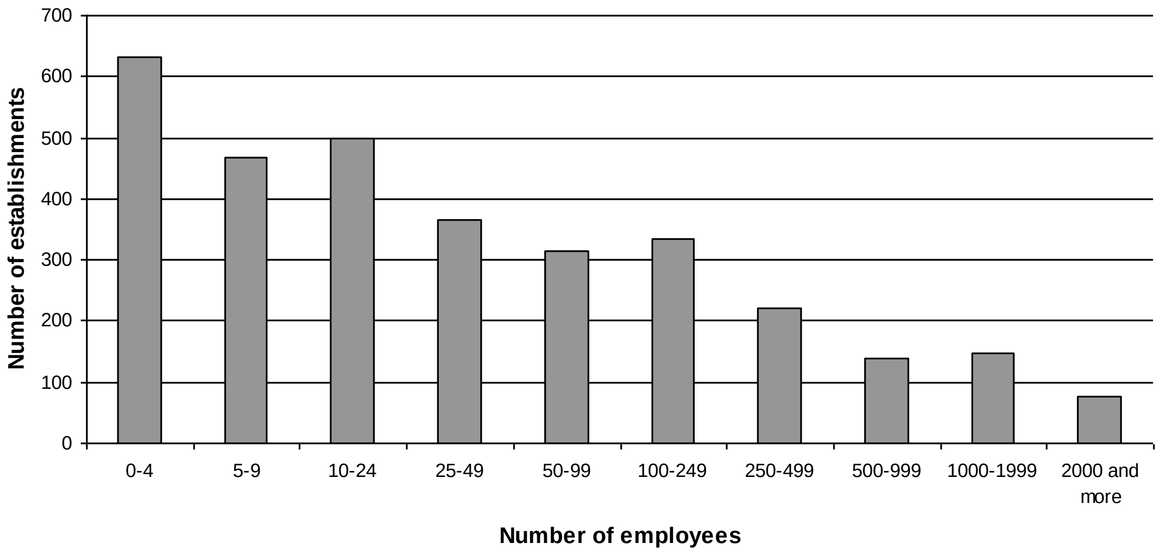
Industry	Freq.	Percent
Agriculture and forestry	156	4.88
Energy, mining, water supply	83	2.60
Chemical industry, petroleum processing	51	1.59
Plastics, rubber industry	/	/
Earths, stones, and fine ceramics industry	63	1.97
Iron, steel, and metal industry	96	3.00
(Light) metal construction	182	5.69
Electrical engineering, data processing machines	117	3.66
Road vehicle manufacturing, garages	90	2.81
Shipbuilding, aircraft construction	/	/
Fine mechanics, toys industry	65	2.03
Wood working	69	2.16
Paper and printing industry	47	1.47
Textile industry	48	1.50
Food, beverages, and tobacco industry	122	3.81
Building industry	427	13.35
Trade	552	17.26
Communications and information transmission	172	5.38
Credit institutions	98	3.06
Insurance industry	48	1.50
Real estate services	63	1.97
Restaurants, accommodation services	167	5.22
Other services	442	13.82
Total	3.198	100.00

Notes: */ signifies data anonymized due to low observation numbers.*

A.2: Distribution of establishments across federal states

Federal state	Freq.	Percent
Berlin West	90	2.81
Schleswig Holstein	55	1.72
Hamburg	55	1.72
Lower Saxony	161	5.03
Bremen	21	0.66
Northrhine-Westphalia	446	13.95
Hesse	128	4.00
Rhineland Palatinate	74	2.31
Baden Wuerttemberg	227	7.10
Bavaria	290	9.07
Saarland	25	0.78
Berlin East	128	4.00
Brandenburg	298	9.32
Mecklenburg Western Pomerania	248	7.75
Saxony	296	9.26
Saxony Anhalt	324	10.13
Thuringia	332	10.38
Total	3,198	100.00

A.3: Distribution of establishments across size classes



A.4: Summary descriptive statistics on dummy independent variables

Full sample		Freq.	Percent
Works council	No	2,102	66.71
	Yes	1,049	33.29
	Total	3,151	100.00
Union contract	No	1,649	51.56
	Yes	1,535	48.00
	Total	3,184	100.00
Founding year	Before 1990	1,930	60.62
	1990 or after	1,254	39.38
	Total	3,184	100.00
Branch plant status	Single-establishment firm	2,562	80.11
	Multi-establishment firm	636	19.89
	Total	3,198	100.00
Subsample of establishments with more than 5 workers			
Works council	No	1,478	58.65
	Yes	1,042	41.35
	Total	2,520	100.00
Continuous training	No	1,254	48.95
	Yes	1,308	51.05
	Total	2,562	100.00