

Goldbach, Stefan

**Conference Paper**

## Innovation and Education: Is there a 'Nerd Effect'?

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2012: Neue Wege und Herausforderungen für den Arbeitsmarkt des 21. Jahrhunderts - Session: Innovation and Network Industries, No. E03-V2

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Goldbach, Stefan (2012) : Innovation and Education: Is there a 'Nerd Effect'? , Beiträge zur Jahrestagung des Vereins für Socialpolitik 2012: Neue Wege und Herausforderungen für den Arbeitsmarkt des 21. Jahrhunderts - Session: Innovation and Network Industries, No. E03-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:

<https://hdl.handle.net/10419/62307>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Innovation and Education: Is there a ‘Nerd Effect’?

Stefan Goldbach<sup>‡</sup>

First version: 23.01.2012

This version: 14.08.2012

## Abstract

This paper investigates whether entrepreneurs with technical education are more innovative in high-tech industries than economists ('nerd effect'). To analyze this question, we examine the KfW/ZEW Start-Up Panel between 2007 and 2008. Two independent OLS regressions are conducted for entrepreneurs, one for university degree and one for apprenticeship as highest qualification. Having controlled for entrepreneurial and firm characteristics, we do not find a 'nerd effect'.

**Keywords:** entrepreneurship, innovation, education, start-ups

**JEL Classification Codes:** A20, L26, I21, O32

---

\*Address: Chair of International Economics, Darmstadt University of Technology, Marktplatz 15, 64283 Darmstadt, goldbach@vwl.tu.darmstadt.de

†I want to thank Marcus Dejardin, Volker Nitsch, Florian Klohn, Stefan Pichler, Dirk Bursian, Markus Roth and participants from the Darmstadt research column 'innovation and growth' for helpful comments.

# 1 Introduction

Policy makers are interested in fostering economic growth and employment. A central driver of economic growth are innovative products (e.g., Hasan and Tucci (2010)). Entrepreneurship is one potential channel through which these goals are achievable. However, not every start-up is innovative (e.g., Reynolds (2005)). Therefore, it is important to know how to boost innovation in an effective way. One important factor is education (e.g., Cooray (2010) or McMahon (1998)). This paper focuses on the transmission effect of education on innovation within start-ups. Small ventures meet with growing interest because recent empirical studies show that start-ups have a comparative advantage in fostering innovation, as Acs and Audretsch (2005) argue. This view contradicts earlier contributions (e.g., Schumpeter (1942)) where large corporations are the only agents who can sustain R&D expenditures through market power. The central question is whether entrepreneurs with technical education are more innovative in high-tech industries compared to economists. This potential interrelation is defined as the ‘nerd effect’ throughout this paper. To analyze this question, we examine a novel data set (KfW/ZEW Start-Up Panel) which contains a random sample of German start-up companies between 2007 and 2008. Two independent OLS regressions are conducted for entrepreneurs, one for university degree and one for apprenticeship as highest qualification. Having controlled for several entrepreneurial and firm characteristics, we do not find a ‘nerd effect’. Our results indicate that the transmission channel within start-ups is different compared to other organizational structures, since entrepreneurs with technical orientation have no advantage in high-tech industries. On the contrary, prior literature (e.g., Toivanen and Väänänen (2011)) shows that individuals who possess an engineering degree and work in a corporation have indeed a positive impact on the registration of patents.

The paper proceeds as follows: Section 2 reviews the relevant literature. Section 3 describes the data set, definitions and provides summary statistics. Section 4 presents the regression results and some robustness checks. Section 5 summarizes the main results and concludes.

## 2 Literature

A (large) theoretical and empirical literature deals with the interaction of education, personality traits and incentives on entrepreneurial entry decision and innovation.

Education itself can be defined in diverse manner. One variant is the choice of career, which is affected by different psychological and socio-economic criteria. Caner and Okten (2010) investigate this relationship for Turkish students in 2002 and conclude that especially male students with high ability have a higher probability of studying engineering.

Innovation is influenced by different characteristics. Several papers concentrate on education as a possible determinant. Toivanen and Väänänen (2011) investigate whether an engineering degree has a positive influence on the registration of patents, which is confirmed for Finnish data between 1988 and 1996. Furthermore, de Mel et al. (2009) find that owner characteristics, especially ability (measured as IQ) and education (measured as years of schooling), play an important role for explaining innovation, using data from Sri Lanka for 2008. Another aspect discussed in the literature is the role of incentives on innovation, as in Sauermann and Cohen (2010). Motives are important but they differ in their effects: intellectual challenge and independence show a strongly positive effect, while job security and responsibility seem to have a negative effect. Innovation within start-ups is highlighted by Praag and Versloot (2007). They review 19 different empirical contributions which differ in how they measure innovation: some concentrate on quantity, others on quality, commercialization or adoption. Entrepreneurs do not invest more in R&D than their competitors and produce fewer innovations. However, they have a comparative advantage in the production of high-quality innovations and in commercialization of innovations.

Another strand deals with the entrepreneurial entry decision. Education or skills are used as the main explanatory variable but the definitions vary. One central contribution for skills is given by Lazear (2005), who introduces the Jack-of-all-trades hypothesis. According to the latter, entrepreneurs have to possess many different skills compared to a specialist. The hypothesis

is confirmed using alumni data from Standford Business School. However, Lazear (2005) neglects the role of education on innovation. His hypothesis is also tested for other countries. Two versions are given by Wagner (2003) and Wagner (2006) who examine a German random sample for 1998 and 1999 and confirm the hypothesis. Institutional settings in schools can also play a role for entry. Using US county-level data from 2000, Sobel and King (2008) show that voucher systems foster entrepreneurial activity through implicit learning in competition and the need for innovation.

Further literature discusses the role of education on start-up performance. There is a possible correlation between performance and innovation but it is not one-to-one. Being innovative does not always coincide with being successful in a monetary sense. Gompers et al. (2005) show that the R&D elasticity of output is less than one. This means that there are patents with zero business value. Moreover, Gilbert and Newbery (1982) argue that companies can register a patent without ever using it. This decision can be strategically motivated because these firms prohibit competition and maintain their market power. Education is defined in this context with reference to different dimensions. One possibility is years of schooling, as used in Parker and van Praag (2006) for Dutch start-ups in 1994. Higher education leads to fewer capital constraints and therefore to better performance. In addition, more schooling also leads directly to more profitability. Human capital can be distinguished by explicit (formal education) and tacit knowledge (know-how) as in Davidsson and Honig (2003) who use Swedish data. Performance is not influenced by these two types of education<sup>1</sup>. Education can also be seen as a quality signal. Backes-Gellner and Werner (2007) explore the effect of eduction as such a signal for banks and employees for German start-ups in 1998 and 1999. The disparity between high-tech and non-high-tech start-ups is emphasized. The authors indeed find such positive effects in the high-tech industry but not in the traditional industries. Dutta et al. (2011) provide another distinction with specialized and diversified education for entrepreneurship alumni data between 1988 and 2008 from public universities in Northeast USA. Specialized knowledge is defined as entrepreneurship courses that are explicitly

---

<sup>1</sup>But the entry decision is affected by both variables.

designed for nascent entrepreneurs. Diversified education is the attendance of courses that are not necessarily related to entrepreneurship. Both variables have no impact on profits.

Education has a notable effect on economic growth as Murphy et al. (1991) show. They introduce two sectors in their theoretical growth model including human capital: one sector is rent-seeking, the other growth enhancing. They define college students enrolled in engineering as persons who initiate technological progress, while those enrolled in law are characterized as rent seekers. More engineers boost economic growth, while more lawyers decrease it. As argued before, a potential channel for economic growth is entrepreneurship. Based on the sector that is growth-promoting, entrepreneurs with an engineering degree should be more innovative than others. This paper can identify and test this potential channel using micro data.

It is especially the latter theoretical and empirical arguments that lead to the so-called ‘nerd effect’ in this paper which is defined as comparative advantage of individuals with technical education in innovation.

**Hypothesis 1** *Entrepreneurs with technical education are more innovative in high-tech industries than entrepreneurs with an economics degree.*

### 3 Data and Summary Statistics

The data used in this paper is a novel data set (KfW/ZEW Start-Up Panel)<sup>2</sup>. It is a random sample that contains yearly data for German start-up companies and contains complete information for 2007 and 2008. Further details are provided by Fryges et al. (2010). An entrepreneur is defined here as someone who belongs to the group of persons who establish a start-up. First, the definitions are described.

#### 3.1 Definitions of Basic Variables

The literature discusses different methods and strategies for measuring ‘innovation’. Acs and Audretsch (2005) emphasize that innovation and technolog-

---

<sup>2</sup>The start-ups are identified by the database of Creditreform.

ical change is a process that is not easily measured. They mention attempts to measure innovation more accurately by using independent experts in the technological field who are able to weight the innovations. Typically, input and output variables are used in empirical studies. Bearing these potential problems of measuring innovation in mind, innovation is approximated in several dimensions. As the basic measures for innovation, a binary input variable indicating whether R&D was conducted (*r&d*) and a binary output variable that indicates whether something new has been released on the market since the foundation (*mrel*) are used. For robustness checks, R&D expenditures per worker (*expend*), the scope of the market release (*new*), a dummy variable showing whether patents are used today or in the future (*pat\_use*) and a dummy variable showing whether a product (*prod*) or process (*proc*) innovation is achieved are employed. *expend* is one further input variable for R&D that is used by other studies examining innovation. The advantage is that innovation activity is measured more objectively, as opposed to a potential bias resulting from more subjective measures and can be evaluated on a metric scale. *new* takes value one if there is no new market innovation, for value two the innovation is at regional level, value three at national level and value four at worldwide level. The variable *pat\_use* describes an output variable which is a dummy. In contrast to other measures, it includes a time dimension. The variables *prod* and *proc* are output dummy variables. They concentrate on the type of innovation. The main explanatory variable in this analysis is education. It is measured in two dimensions: the amount is scaled as dummy variable *uni*, which takes value one if the entrepreneur has a degree from a university and zero if the person completed vocational training. The second dimension is the field that is studied. Dummy variables are generated for business or economics (*econ*), natural science (*nat*), mathematics or informatics (*mathinf*), engineering (*eng*) and other subjects (*other*). These are only available for entrepreneurs with a university degree. Practical education uses a different notation. Having an apprenticeship in commerce is *comm*. The other fields are technical (*tech*), social (*social*), other services (*othserv*) and other professions (*other\_job*).

### 3.2 Definitions of Control Variables

To control for entrepreneur's personal traits, nationality (*german*), sex (*male*), experience, prior employment situation, foundation motivation and ownership structure are included. Experience is measured in intervals: less than seven years, more than seven and less than 13 years, more than 13 and less than 20 years and more than 20 years. The employment situation immediately before the start of the venture is approximated by dummy variables. An entrepreneur was either self-employed (*sit\_e*), employed (*sit\_em*), registered as unemployed (*sit\_unem*) or not working (*sit\_ne*). Motivation is classified as working independently, realising a business idea, improper employment opportunities, escape from unemployment, encouragement by former employer or tax incentives. Ownership structure is measured as the share that is financed by the entrepreneur himself (*fin\_sh*) and by external investors (*fin\_ext\_sh*). The higher the entrepreneur's share, the greater the incentive to be successful in innovation because he gains from being successful. Beside these personal traits, firm characteristics are also considered as additional control variables. Firm size is determined by the number of different types of employees: amount of full-time, part-time, mini, family members, trainees, freelancer, interns and temporary employees. The sum of all these types is illustrated in *employment*. Another component is the quality of this employment pool: the number of employees with no apprenticeship (*sh\_l*), an apprenticeship (*sh\_m*) or a university degree (*sh\_h*) is embedded. Competition structure may affect innovative activity as well. *lcomp* describes low competition when the start-up faces less than six other companies as competitors, *mcomp* identifies between six and twenty companies as competitors and *hcomp* more than twenty companies. ZEW categorizes industries into high-tech and non-high-tech industries. This definition is adopted in the following. The classification is described in Table 1.

### 3.3 Summary Statistics

This subsection starts with the provision of some stylized facts based on the sample. We have an unbalanced panel for German start-ups founded in

2007 and 2008, with approximately 7,000 observations. Table 2 provides a description for our proxies for innovation<sup>3</sup>. The basic measures show that 23% of the start-ups are temporarily or permanently engaged in R&D and 21% have released a market innovation since foundation. These two variables illustrate only a part of the whole innovation process. Other variables are investigated that can illustrate further aspects. Each start-up invests 2,920 euro per employee on average. The high standard deviation indicates that there is a remarkable fraction of start-ups investing no money. *new* describes the average innovation as relatively small in scope. According to *pat\_use*, only a fraction of start-ups (approximately 8%) are engaged in patenting today or in the future. 37% of start-up innovations are connected to products, 27% exhibit innovation in processes. Table 3 presents personal traits and firm characteristics.

First, the distribution of education is described. Approximately 38% have a university degree. Of the latter, 46% studied engineering, 27% economics, 12% mathematics or informatics, 12% natural science and 15% another subject. According to these statistics, most start-ups in the sample were founded by persons with a technical background. These numbers are compared with individuals who have a practical education: most have either a technical (62%) or commercial (26%) education. Some studied social science (8%), other services (7%) or completed an apprenticeship in other jobs (4%). Figure 1 describes the distribution among both industries for those with a university degree. There is a higher fraction of engineers, natural scientists and mathematicians/informaticists in the high-tech industry. In contrast, entrepreneurs with a business/economics degree are better represented in the non-high-tech industry. The same holds true for apprenticeship, as Figure 2 illustrates. More entrepreneurs with a background in commercial or other services are active in the non-high-tech industry, while those with technical and social sciences are engaged in the high-tech industry. Next, the other personal traits are examined. 95% of the entrepreneurs are German and 80% male. Most entrepreneurs (61%) were employed in a firm prior to the start-

---

<sup>3</sup>Minimum and maximum values are not reported due to provision restrictions. However, these values are very similar compared to the reported intervals.

up, only 22% were self-employed. Finally, the firm characteristics are shown. 41% of the ventures are engaged in the high-tech industry, 27% founded by teams. A team dummy variable is included to control for potential collaboration effects. The average entrepreneur contributes 20% of the assets by himself and receives 12% from outside financiers. Many start-ups have only few employees (4-5) and, if so, the number of employees with vocational training is highest. 56% face high competition in their environment.

## 4 Empirical Results

### 4.1 Baseline Regressions

In the following, we want to test the hypothesis previously stated in the theoretical section: entrepreneurs with technical education are more innovative in high-tech industries than entrepreneurs with an economics degree. Accordingly, we use  $r\&d$  and  $mrel$  as dependent variables. They describe different parts of the overall innovation process.  $r\&d$  can be interpreted as the input,  $mrel$  as the output variable. Subsequently, other innovation proxies are used for robustness checks. Table 4 presents the correlation among the dependent variables. It shows that the variables are correlated to some extent but multicollinearity can be ruled out.  $mrel$  and  $new$  are highly correlated because the first variable is approximated by the second one. The correlation matrix indicates that other proxies do not capture the same aspect compared to both baseline measures. To establish a relationship between innovation and education the following equations are estimated:

$$r\&d_i = \alpha + \beta x_i + \gamma z_i + u_i \quad (1)$$

$$mrel_i = \alpha + \beta x_i + \gamma z_i + u_i \quad (2)$$

where  $x_i$  is the vector of explanatory variables (in this case the variables for education) and  $z_i$  the vector of control variables (other personal traits and firm characteristics). We estimate two independent OLS regressions for each equation, one for university degree and one for apprenticeship as highest

qualification. Although the dependent variable is a dummy variable, we do not estimate a probit model as baseline regression for one particular reason: the probit model structure imposes normality as restrictive assumption for the cumulative distribution function. When a saturated model is at hand, Angrist and Pischke (2008) suggest that using OLS is superior for identifying causality. This is only true when there is a random sample treatment in the data<sup>4</sup>. Ideally, we would be able to reveal the relation between education and innovation experimentally, meaning that the entrepreneurs should be randomly endowed with different types of education. Since the implementation of such an experiment is obviously impossible, we have to approximate such a situation as best we can. Our identification strategy is to control for the maximum number of variables that are correlated with innovation and education. Socio-demographic and personality traits that impair the career choice are especially taken into account in order to rule out a possible selection bias. All estimations include robust standard errors. We start with the analysis of having a university degree.

Table 5 presents the estimation results for  $r\&d$  and  $mrel$  as dependent variable. The first column uses the education dummies and their interaction with high-tech industry as explanatory variables, while the team dummy, the high-tech dummy and time fixed effects (for potential time trends) serve as controls. The interaction terms can be interpreted as the additional effect of having a certain type of education in the high-tech industry on innovating activity. Possessing a degree in natural science is weakly significant and positive. In addition, an economics and natural science degree have a positive and statistically significant effect for entrepreneurs in the high-tech industry. In the second column other personal traits are included as additional controls. It contains gender, nationality, experience, the situation prior to the foundation of the start-up, motives for foundation and ownership structure of the start-up. The last variable is measured by the share of assets that is provided by the entrepreneur himself. It could be that the entrepreneur is more innovative just because of a better financial situation. More equipment can be bought which can be used for innovation. The third column appends

---

<sup>4</sup>The underlying data-generating process in the ZEW/KfW Start-Up Panel is random.

firm characteristics, such as firm size, the quality of the employment pool and the competition structure. Since the education effect on innovation is the focus of this study, the other regression coefficients are not presented<sup>5</sup>. As can be seen in columns (2) and (3), the relation between an economics degree and innovation is negative and significant. In contrast, for economists in the high-tech industry the effect is significant and positive. The overall net effect (economics plus interaction of economics and high-tech industry) is positive for entrepreneurs in the high-tech industry. The impact is robust among all specifications. Furthermore, we do not find a ‘nerd effect’. Holding a natural science degree is associated with more innovation in the high-tech industry but this effect is only weakly significant for columns (1) and (2). Therefore, the hypothesis can be rejected. Columns (4) to (6) illustrate the findings for *mrel* as dependent variable. The columns have the same structure as before in the sense that further control variables are included in each step to control for potential biases. Almost all education variables for start-ups in the non-high-tech industry are insignificant, the only exception being column (4). Natural science exhibits a weakly significant and positive influence. Possessing an economics or engineering degree in the high-tech industry is positively correlated with innovation. This effect vanishes when personal traits or firm characteristics are included as control variables. Next, the hypothesis is tested for practical education. Equations (1) and (2) are re-estimated for persons with apprenticeship as highest qualification.

The results are completely different, as Table 6 illustrates. Technical education is insignificant in all three columns. The interaction term reveals a negative and significant association which becomes weaker when controlling for more characteristics. This strong negative influence is also true for social education in the high-tech industry. The other education variables are mostly insignificant. The findings are even weaker for columns (4) to (6). All in all, there is neither a ‘nerd effect’ for entrepreneurs with a university degree nor for those with vocational training as highest qualification. Possessing an economics degree can increase innovation in the high-tech industry as some specifications suggest, but there is no effect for vocational training. As

---

<sup>5</sup>Additional regression results can be provided upon request.

argued before, the defined innovation variables are not able to capture the complete innovation process. For this reason, further dimensions are tested to investigate whether the findings depend on the underlying definitions.

## 4.2 Robustness Checks

The linear probability model approach has the drawback that fitted values of the dependent variable can lie outside the range between zero and one. To check whether this is indeed the case here, equations (1) and (2) are re-estimated with probit. The signs and significance levels do not change. Fitted values from our OLS regressions show that only few values are outside the range between zero and one<sup>6</sup>. This could be a possible explanation for the appearance of almost identical findings.

For further robustness checks, OLS is used. Other measures for innovation are employed. As the other input variable which exhibits a metric scale, R&D expenditures per employee is typically used. They can be interpreted as importance of R&D to the firm. *new* provides information about the innovation level on an ordinal scale with higher values indicating more scope (whether the innovation is worldwide or only regional). An ordered probit approach is used for evaluation. Table 7 reports the results for entrepreneurs who possess a university degree.

Almost all specifications are characterized by insignificant education variables for *expend*. The only exception is natural science in column (1), where no entrepreneurial and firm characteristics are considered as controls. Therefore, education does not seem to play an important role for R&D expenditures per employee. *new* is positively influenced by mathematics/informatics. In contrast, the additional effect (interaction term) of this subject is negative. Again, the hypothesis can be rejected. Table 8 shows the results for practical education. Again, most educational variables are insignificant. Technical education is significant only in column (4) with a negative sign.

*pat\_use* includes a time dimension showing whether patents play a role

---

<sup>6</sup>One potential reason for this is that we include mostly dummy variables as explanatory variables in the regressions.

for the start-up today or in the future. Table 9 shows the regression results for university degree, Table 10 for vocational training. Possessing a degree in economics or natural science in the high-tech industry is associated with more patent use. The significance of economics is larger, while the magnitude of natural science is slightly higher. In contrast, all education variables are insignificant for practical education. The last two variables *prod* and *proc* distinguish between the type of innovation that is conducted. It is conceivable that different types of entrepreneurs focus on various aspects of innovation. Table 11 illustrates the comparison of product and process innovation for entrepreneurs with a university degree.

Among all specifications only column (4) exhibits a weakly positive association between economics degree in the high-tech industry and process innovation. All other factors are irrelevant. Table 12 presents the findings for vocational training. Again, many variables are insignificant. Only other services for product innovation and social science in the high-tech industry for process innovation show positive and significant results. All findings using different proxies for innovation show that the results vary according to the underlying definition of innovation. This seems reasonable because the definitions can only illustrate some part of the whole innovation process. Every indicator focuses on different dimensions that are not identical. Nevertheless, one central finding is observed among all specifications: entrepreneurs with technical education do not seem to have a comparative advantage in the high-tech industry compared to economists. Therefore, we do not find a ‘nerd effect’ in our sample.

## 5 Conclusion

Entrepreneurship is one potential channel through which economic growth and employment are achievable. This paper focuses on the transmission effect of education on innovation within start-ups using a novel data set (ZEW/KfW Start-Up Panel) from 2007 to 2008. This set contains various dimensions of innovation and many entrepreneurial (especially type of education) and firm characteristics. The findings suggest that there is no

‘nerd effect’. Individuals with an economics degree are more innovative in the high-tech industry. According to *r&d* and *mrel*, they are more capable of conducting R&D and selling the innovation to the market. These conclusions cannot be drawn for persons with vocational training. Robustness checks with other proxies for innovation are conducted to capture more dimensions of the whole innovation process. Although these results provide different significance levels, the central conclusion that there is no ‘nerd effect’ is maintained.

## References

- Acs, Z. and Audretsch, D. (2005). Entrepreneurship, innovation and technological change. *Foundations and Trends in Entrepreneurship*, 1(4):1–65.
- Angrist, J. and Pischke, J. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Backes-Gellner, U. and Werner, A. (2007). Entrepreneurial signaling via education: A success factor in innovative start-ups. *Small Business Economics*, 29(1):173–190.
- Caner, A. and Okten, C. (2010). Risk and career choice: Evidence from turkey. *Economics of Education Review*, 29(6):1060–1075.
- Cooray, A. (2010). The role of education in economic growth. Economics Working Papers wp10-14, School of Economics, University of Wollongong, NSW, Australia.
- Davidsson, P. and Honig, B. (2003). The role of social and human capital among nascent entrepreneurs. *Journal of Business Venturing*, 18(3):301–331.
- de Mel, S., McKenzie, D., and Woodruff, C. (2009). Innovative firms or innovative owners? determinants of innovation in micro, small, and medium enterprises. IZA Discussion Papers 3962, Institute for the Study of Labor (IZA).
- Dutta, D., Li, J., and Merenda, M. (2011). Fostering entrepreneurship: impact of specialization and diversity in education. *International Entrepreneurship and Management Journal*, 7(2):163–179.
- Fryges, H., Gottschalk, S., and Kohn, K. (2010). The kfw/zew start-up panel: Design and research potential. *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 130(1):117–132.

- Gilbert, R. J. and Newbery, D. M. G. (1982). Preemptive patenting and the persistence of monopoly. *American Economic Review*, 72(3):514–26.
- Gompers, P., Lerner, J., and Scharfstein, D. (2005). Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999. *Journal of Finance*, 60(2):577–614.
- Hasan, I. and Tucci, C. L. (2010). The innovation-economic growth nexus: Global evidence. *Research Policy*, 39(10):1264–1276.
- Lazear, E. P. (2005). Entrepreneurship. *Journal of Labor Economics*, 23(4):649–680.
- McMahon, W. (1998). Education and growth in east asia. *Economics of Education Review*, 17(2):159–172.
- Murphy, K., Shleifer, A., and Vishny, R. (1991). The allocation of talent: Implications for growth. *The Quarterly Journal of Economics*, 106(2):503–530.
- Parker, S. C. and van Praag, C. M. (2006). Schooling, capital constraints, and entrepreneurial performance: The endogenous triangle. *Journal of Business & Economic Statistics*, 24:416–431.
- Praag, C. and Versloot, P. (2007). What is the value of entrepreneurship? a review of recent research. *Small Business Economics*, 29(4):351–382.
- Reynolds, P. (2005). Understanding business creation: Serendipity and scope in two decades of business creation studies. *Small Business Economics*, 24(4):359–364.
- Sauermann, H. and Cohen, W. (2010). What makes them tick? employee motives and industrial innovation. *Management Science*, 56(12):2134–2153.
- Schumpeter, J. (1942). *Socialism, capitalism and democracy*. Harper and Bros.

- Sobel, R. and King, K. (2008). Does school choice increase the rate of youth entrepreneurship? *Economics of Education Review*, 27(4):429–438.
- Toivanen, O. and Väänänen, L. (2011). Education and invention. CEPR Discussion Papers 8537, CEPR Discussion Papers.
- Wagner, J. (2003). Testing lazear's jack-of-all-trades view of entrepreneurship with german micro data. *Applied Economics Letters*, 10(11):687–689.
- Wagner, J. (2006). Are nascent entrepreneurs 'jacks-of-all-trades'? a test of lazear's theory of entrepreneurship with german data. *Applied Economics*, 38(20):2415–2419.

# Appendix

High-technology industries
Cutting-edge technology manufacturing
High-technology manufacturing
Technology-intensive services
Software
Non-high-tech industries
Non-high-tech manufacturing
Skill-intensive services (non-technical, consulting services)
Other business-oriented services
Consumer-oriented services
Construction
Wholesale and retail market

Table 1: Industry classifications

Variable	Obs	Mean	Std. Dev.	1%	99%
r&d	7,028	0.2314	0.4217	0	1
mrel	7,028	0.2107	0.4079	0	1
expend	7,028	2,920.89	20,910.88	0	50,000
new	7,028	1.39	0.8353	1	4
pat_use	5,006	0.0759	0.2649	0	1
prod	4,048	0.3752	0.4842	0	1
proc	4,077	0.2727	0.4454	0	1

Table 2: Summary statistics of dependent variables (innovation)

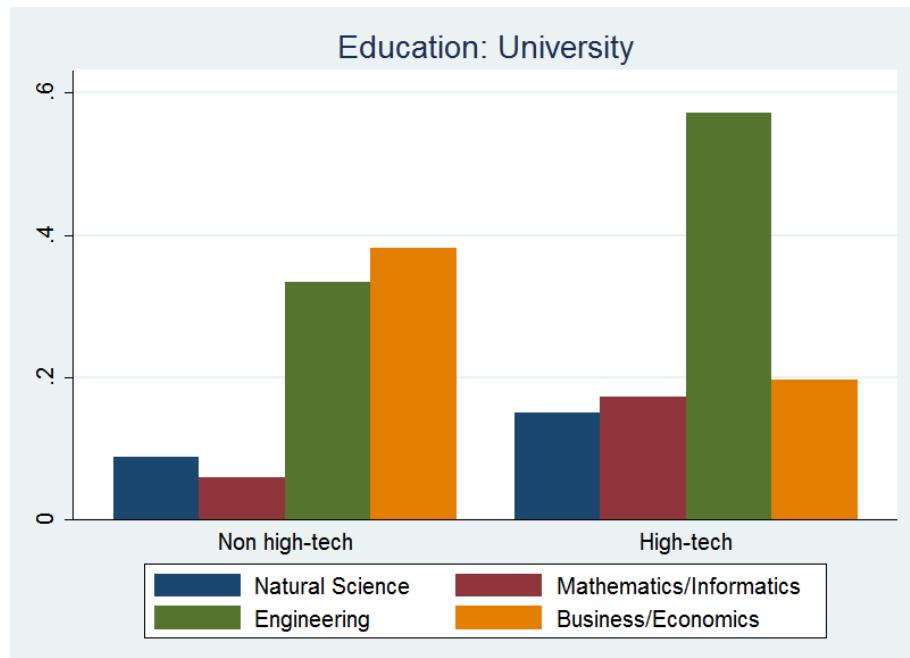


Figure 1: Share of education (university) in non high-tech and high-tech industries

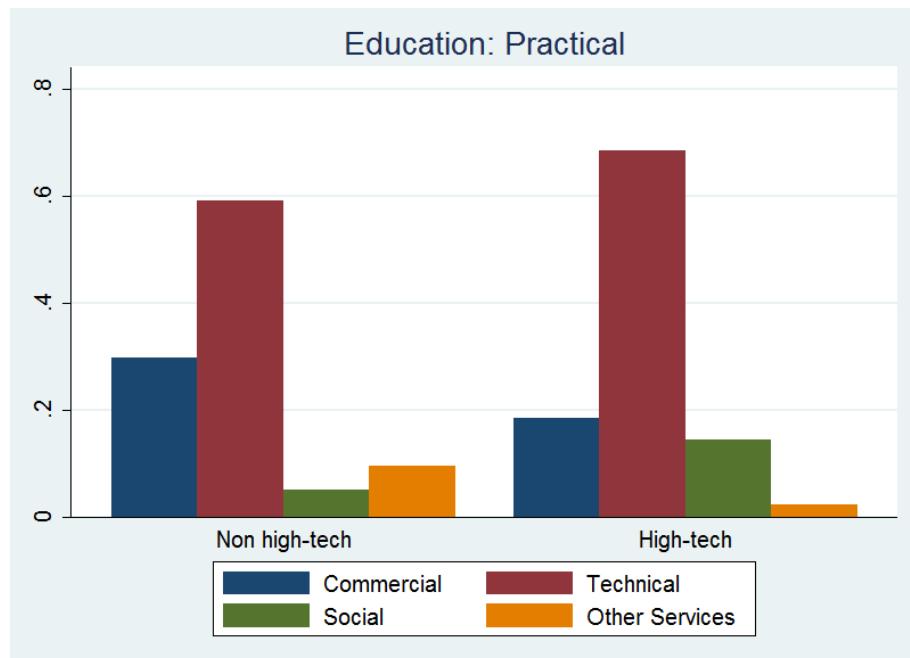


Figure 2: Share of education (practical) in non high-tech and high-tech industries

Variable	Obs	Mean	Std. Dev.	1%	99%
<b>A. Education</b>					
uni	7,028	0.3840	0.4864	0	1
nat	2,660	0.1218	0.3271	0	1
mathinf	2,660	0.1214	0.3267	0	1
eng	2,660	0.4654	0.4989	0	1
econ	2,660	0.2789	0.4486	0	1
other	2,660	0.1526	0.3597	0	1
comm	4,493	0.2597	0.4385	0	1
tech	4,493	0.6227	0.4848	0	1
social	4,493	0.0828	0.2756	0	1
othserv	4,493	0.0701	0.2554	0	1
other_job	4,493	0.0445	0.2063	0	1
<b>B. Other personal traits</b>					
ger	7,020	0.9352	0.2462	0	1
male	7,028	0.8029	0.3978	0	1
sit_e	7,007	0.2179	0.4129	0	1
sit_em	7,007	0.6158	0.4864	0	1
sit_unem	7,007	0.1590	0.3657	0	1
sit_ne	7,007	0.0918	0.2887	0	1
<b>C. Firm characteristics</b>					
ht	7,028	0.4119	0.4922	0	1
team	7,028	0.2694	0.4437	0	1
fin_sh	4,745	20.66	30.72	0	100
fin_ext_sh	4,820	12.49	25.49	0	100
employment	7,028	4.64	6.61	1	29
sh_l	6,013	0.8202	2.97	0	14
sh_m	6,016	1.63	3.44	0	16
sh_h	6,020	0.3802	1.65	0	8
lcomp	4,886	0.2386	0.4263	0	1
mcomp	4,886	0.2016	0.4012	0	1
hcomp	4,886	0.5598	0.4965	0	1

Table 3: Summary statistics of explanatory variables

Variable	r&d	mrel	expend	new	pat_use	prod	proc
r&d	1.0000						
mrel	0.3022	1.0000					
expend	0.2595	0.0998	1.0000				
new	0.3570	0.9009	0.1418	1.0000			
pat_use	0.3178	0.2609	0.2000	0.3369	1.0000		
prod	0.2480	0.2618	0.0299	0.2520	0.0988	1.0000	
proc	0.2357	0.1877	0.0579	0.1927	0.1198	0.3715	1.0000

Table 4: Correlation matrix

Variables	r&d	r&d	r&d	mrel	mrel	mrel
team	0.0947*** (0.0187)	0.0283 (0.0269)	0.0122 (0.0265)	0.0872*** (0.0184)	0.0306 (0.0260)	0.0126 (0.0256)
ht	0.102** (0.0438)	0.0961* (0.0545)	0.0839 (0.0533)	-0.0140 (0.0426)	-0.0334 (0.0520)	-0.00222 (0.0507)
nat	0.0932* (0.0503)	0.0636 (0.0640)	0.0341 (0.0607)	0.0904* (0.0515)	0.0576 (0.0660)	0.0396 (0.0625)
mathinf	0.0492 (0.0574)	0.0527 (0.0703)	0.0292 (0.0678)	0.0770 (0.0586)	0.0882 (0.0703)	0.0971 (0.0719)
eng	-0.0184 (0.0325)	-0.0424 (0.0419)	-0.0441 (0.0415)	-0.0421 (0.0317)	-0.0541 (0.0408)	-0.0308 (0.0398)
econ	-0.0448 (0.0310)	-0.0701* (0.0397)	-0.0856** (0.0390)	-0.0116 (0.0312)	-0.0337 (0.0404)	-0.0239 (0.0399)
ht_nat	0.120* (0.0652)	0.139* (0.0827)	0.115 (0.0794)	0.0572 (0.0660)	0.0823 (0.0832)	0.0695 (0.0797)
ht_mathinf	0.0555 (0.0704)	0.0599 (0.0858)	0.0785 (0.0830)	-0.0745 (0.0703)	-0.0995 (0.0839)	-0.113 (0.0846)
ht_eng	0.0545 (0.0483)	0.0846 (0.0598)	0.0821 (0.0590)	0.0862* (0.0471)	0.0950 (0.0578)	0.0629 (0.0563)
ht_econ	0.171*** (0.0482)	0.152** (0.0602)	0.147** (0.0591)	0.0986** (0.0477)	0.0843 (0.0598)	0.0574 (0.0581)
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Entr. Char.	No	Yes	Yes	No	Yes	Yes
Firm Char.	No	No	Yes	No	No	Yes
Observations	2,660	1,673	1,643	2,660	1,673	1,643
R-squared	0.088	0.160	0.204	0.028	0.093	0.156

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 5: Baseline regressions for university degree. Standard errors are corrected for heteroscedasticity.

Variables	r&d	r&d	r&d	mrel	mrel	mrel
team	0.114*** (0.0147)	0.0816*** (0.0214)	0.0624*** (0.0213)	0.0967*** (0.0150)	0.0483** (0.0210)	0.0332 (0.0213)
ht	0.278*** (0.0491)	0.252*** (0.0616)	0.237*** (0.0605)	0.0917* (0.0495)	0.0880 (0.0606)	0.0806 (0.0579)
comm	-0.00471 (0.0220)	-0.00164 (0.0297)	-0.00755 (0.0291)	-0.0115 (0.0240)	-0.00294 (0.0319)	-0.00313 (0.0321)
tech	0.00557 (0.0221)	0.000194 (0.0305)	0.00673 (0.0301)	-0.0593** (0.0241)	-0.0396 (0.0318)	-0.0273 (0.0318)
social	-0.0528* (0.0270)	-0.0402 (0.0396)	-0.0467 (0.0397)	-0.0204 (0.0345)	0.0101 (0.0473)	-0.00227 (0.0497)
othserv	0.0358 (0.0278)	0.0672* (0.0387)	0.0486 (0.0393)	-0.00834 (0.0307)	-0.000668 (0.0404)	-0.0260 (0.0414)
ht_comm	-0.0190 (0.0474)	-0.0448 (0.0603)	-0.0513 (0.0588)	-0.00745 (0.0478)	-0.0109 (0.0606)	-0.0238 (0.0578)
ht_tech	-0.147*** (0.0488)	-0.133** (0.0615)	-0.116* (0.0602)	-0.0145 (0.0488)	-0.0268 (0.0601)	-0.0144 (0.0575)
ht_social	-0.175*** (0.0543)	-0.179** (0.0715)	-0.149** (0.0713)	-0.0985* (0.0589)	-0.117 (0.0733)	-0.0732 (0.0729)
ht_othserv	-0.0428 (0.0940)	0.00674 (0.117)	0.0429 (0.117)	0.0969 (0.0957)	0.167 (0.123)	0.207* (0.119)
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Entr. Char.	No	Yes	Yes	No	Yes	Yes
Firm Char.	No	No	Yes	No	No	Yes
Observations	4,493	2,870	2,798	4,493	2,870	2,798
R-squared	0.069	0.085	0.127	0.026	0.067	0.105

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 6: Baseline regressions for practical education. Standard errors are corrected for heteroscedasticity.

Variables	expend	expend	expend	new	new	new
team	-359.2 (1,080)	-627.7 (1,008)	-673.9 (1,026)	0.270*** (0.0519)	0.122 (0.0772)	0.0691 (0.0804)
ht	3,540 (2,543)	2,574 (2,477)	2,322 (2,483)	0.0747 (0.121)	-0.0254 (0.150)	0.0468 (0.158)
nat	3,089* (1,601)	1,449 (1,056)	514.8 (998.1)	0.281** (0.132)	0.225 (0.167)	0.154 (0.170)
mathinf	2,020 (1,402)	2,906 (1,917)	2,341 (1,959)	0.335** (0.162)	0.393** (0.200)	0.456** (0.214)
eng	431.2 (584.2)	-167.9 (620.8)	-111.6 (647.9)	-0.0630 (0.0960)	-0.116 (0.122)	-0.0547 (0.128)
econ	443.3 (642.3)	583.1 (734.2)	370.1 (744.6)	0.00328 (0.0905)	-0.0592 (0.116)	-0.0499 (0.125)
ht_nat	2,251 (2,835)	3,386 (2,410)	2,767 (2,387)	0.172 (0.172)	0.251 (0.220)	0.240 (0.223)
ht_mathinf	-3,040 (2,857)	-3,669 (2,652)	-3,441 (2,671)	-0.327* (0.194)	-0.409* (0.240)	-0.492* (0.255)
ht_eng	274.2 (2,723)	1,335 (2,604)	938.7 (2,650)	0.195 (0.136)	0.278 (0.169)	0.198 (0.177)
ht_econ	6,016* (3,498)	2,049 (3,065)	1,901 (3,099)	0.215* (0.130)	0.190 (0.169)	0.123 (0.178)
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Entr. Char.	No	Yes	Yes	No	Yes	Yes
Firm Char.	No	No	Yes	No	No	Yes
Observations	2,660	1,673	1,643	2,660	1,673	1,643
R-squared	0.022	0.052	0.085			

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 7: OLS and ordered probit regressions for university degree. Standard errors are corrected for heteroscedasticity.

Variables	expend	expend	expend	new	new	new
team	900.9** (449.9)	1,094* (633.7)	1,124* (657.4)	0.393*** (0.0494)	0.214*** (0.0731)	0.179** (0.0762)
ht	2,256* (1,350)	1,690 (1,506)	1,497 (1,538)	0.384** (0.160)	0.395** (0.201)	0.403** (0.201)
comm	-635.3 (670.4)	92.64 (978.6)	279.9 (1,027)	0.0121 (0.0894)	0.0708 (0.119)	0.0801 (0.125)
tech	-594.2 (683.1)	-999.4 (825.9)	-927.2 (861.5)	-0.198** (0.0896)	-0.108 (0.119)	-0.0592 (0.123)
social	1,800 (2,970)	4,679 (5,146)	4,553 (5,177)	-0.0867 (0.132)	0.0465 (0.180)	0.0337 (0.193)
othserv	-591.6 (953.5)	-1,213 (896.8)	-1,229 (918.7)	-0.0463 (0.110)	-0.0227 (0.146)	-0.126 (0.158)
ht_comm	1,204 (1,234)	-64.23 (1,468)	-216.7 (1,478)	-0.0658 (0.158)	-0.105 (0.202)	-0.179 (0.202)
ht_tech	-979.6 (1,374)	-385.4 (1,530)	-270.2 (1,557)	-0.0205 (0.161)	-0.0917 (0.201)	-0.0800 (0.201)
ht_social	-4,512 (3,153)	-6,619 (5,357)	-6,173 (5,334)	-0.351* (0.205)	-0.465* (0.262)	-0.361 (0.268)
ht_othserv	-1,565 (1,634)	-359.6 (1,458)	-402.0 (1,533)	0.227 (0.258)	0.374 (0.304)	0.527* (0.295)
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Entr. Char.	No	Yes	Yes	No	Yes	Yes
Firm Char.	No	No	Yes	No	No	Yes
Observations	4,493	2,870	2,798	4,493	2,870	2,798
R-squared	0.004	0.012	0.018			

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 8: OLS and ordered probit regressions for practical education. Standard errors are corrected for heteroscedasticity.

Variables	pat_use	pat_use	pat_use
team	0.0533*** (0.0158)	0.0229 (0.0187)	0.0218 (0.0190)
ht	-0.0106 (0.0364)	-0.00379 (0.0385)	-0.00460 (0.0382)
nat	0.0894** (0.0448)	0.0788 (0.0499)	0.0570 (0.0479)
mathinf	-0.0561** (0.0240)	-0.0515* (0.0271)	-0.0629*** (0.0244)
eng	0.0140 (0.0226)	0.00457 (0.0261)	0.00647 (0.0266)
econ	-0.00343 (0.0221)	-0.0195 (0.0241)	-0.0208 (0.0240)
ht_nat	0.135** (0.0636)	0.134** (0.0674)	0.130* (0.0663)
ht_mathinf	0.0617 (0.0419)	0.0607 (0.0443)	0.0658 (0.0424)
ht_eng	0.0736* (0.0405)	0.0641 (0.0425)	0.0557 (0.0428)
ht_econ	0.128*** (0.0436)	0.126*** (0.0449)	0.120*** (0.0452)
Time Effects	Yes	Yes	Yes
Entr. Char.	No	Yes	Yes
Firm Char.	No	No	Yes
Observations	1,844	1,673	1,643
R-squared	0.065	0.107	0.128

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 9: OLS regressions for university degree. Standard errors are corrected for heteroscedasticity.

Variables	pat_use	pat_use	pat_use
team	0.0460*** (0.0109)	0.0199 (0.0131)	0.0177 (0.0134)
ht	0.0828** (0.0412)	0.0773* (0.0422)	0.0658 (0.0420)
comm	0.00775 (0.0146)	0.0131 (0.0170)	0.0135 (0.0176)
tech	4.52e-05 (0.0150)	0.00834 (0.0175)	0.00885 (0.0182)
social	-0.0190 (0.0182)	-0.00576 (0.0216)	-0.0123 (0.0220)
othserv	-0.0102 (0.0166)	-0.0117 (0.0176)	-0.0233 (0.0188)
ht_comm	0.0273 (0.0414)	0.0261 (0.0427)	0.0250 (0.0428)
ht_tech	-0.0410 (0.0414)	-0.0377 (0.0423)	-0.0270 (0.0422)
ht_social	-0.0483 (0.0454)	-0.0426 (0.0469)	-0.0217 (0.0471)
ht_othserv	-0.0306 (0.0604)	-0.0406 (0.0535)	-0.0194 (0.0548)
Time Effects	Yes	Yes	Yes
Entr. Char.	No	Yes	Yes
Firm Char.	No	No	Yes
Observations	3,251	2,870	2,798
R-squared	0.029	0.054	0.071

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 10: OLS regressions for practical education. Standard errors are corrected for heteroscedasticity.

Variables	prod	prod	prod	proc	proc	proc
team	0.0609** (0.0267)	0.0307 (0.0364)	0.0126 (0.0373)	0.0722*** (0.0251)	0.0248 (0.0347)	0.00547 (0.0348)
ht	0.0757 (0.0615)	0.141* (0.0729)	0.145** (0.0731)	0.0707 (0.0567)	0.111* (0.0668)	0.122* (0.0668)
nat	-0.0256 (0.0738)	0.0116 (0.0870)	-0.00285 (0.0879)	0.0473 (0.0692)	0.0402 (0.0804)	0.0355 (0.0821)
mathinf	-0.0199 (0.0828)	0.0125 (0.0916)	0.00923 (0.0962)	0.0342 (0.0757)	0.0972 (0.0893)	0.0804 (0.0927)
eng	-0.0625 (0.0474)	-0.0182 (0.0554)	-0.0255 (0.0565)	-0.0197 (0.0409)	0.00777 (0.0492)	-0.000867 (0.0499)
econ	-0.0365 (0.0470)	-0.00959 (0.0547)	-0.0214 (0.0555)	0.0281 (0.0416)	0.0512 (0.0495)	0.0348 (0.0504)
ht_nat	0.0621 (0.0932)	-0.0139 (0.111)	-0.00933 (0.111)	-0.00601 (0.0881)	-0.0589 (0.103)	-0.0824 (0.103)
ht_mathinf	-0.0728 (0.100)	-0.120 (0.112)	-0.114 (0.115)	-0.0554 (0.0933)	-0.137 (0.108)	-0.136 (0.110)
ht_eng	-0.0107 (0.0684)	-0.0858 (0.0811)	-0.0871 (0.0811)	0.0292 (0.0625)	-0.00680 (0.0741)	-0.0134 (0.0738)
ht_econ	0.104 (0.0692)	0.0276 (0.0825)	0.0293 (0.0827)	0.126* (0.0647)	0.0649 (0.0758)	0.0457 (0.0758)
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Entr. Char.	No	Yes	Yes	No	Yes	Yes
Firm Char.	No	No	Yes	No	No	Yes
Observations	1,522	1,129	1,114	1,531	1,134	1,119
R-squared	0.019	0.046	0.067	0.031	0.046	0.074

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 11: OLS regressions for university degree. Standard errors are corrected for heteroscedasticity.

Variables	prod	prod	prod	proc	proc	proc
team	0.0813*** (0.0230)	0.0706** (0.0307)	0.0530* (0.0314)	0.113*** (0.0217)	0.0991*** (0.0289)	0.0786*** (0.0297)
ht	0.103 (0.0727)	0.177** (0.0853)	0.171** (0.0836)	0.0628 (0.0654)	-0.0128 (0.0756)	-0.0534 (0.0760)
comm	0.0353 (0.0372)	0.0612 (0.0427)	0.0593 (0.0436)	-0.00352 (0.0330)	0.0135 (0.0382)	-0.00533 (0.0379)
tech	-0.0395 (0.0371)	-0.00514 (0.0428)	0.0114 (0.0440)	-0.0354 (0.0328)	-0.0377 (0.0388)	-0.0426 (0.0389)
social	0.00520 (0.0592)	-0.0144 (0.0670)	-0.0144 (0.0674)	0.00767 (0.0531)	-0.00808 (0.0621)	-0.0294 (0.0613)
othserv	0.146*** (0.0499)	0.174*** (0.0594)	0.141** (0.0614)	0.00506 (0.0432)	0.0113 (0.0521)	-0.00528 (0.0521)
ht_comm	0.0251 (0.0711)	-0.0580 (0.0848)	-0.0590 (0.0834)	0.0447 (0.0654)	0.0530 (0.0773)	0.0629 (0.0771)
ht_tech	-0.0246 (0.0727)	-0.105 (0.0850)	-0.103 (0.0834)	0.00128 (0.0658)	0.0763 (0.0764)	0.110 (0.0762)
ht_social	-0.00223 (0.0915)	0.00232 (0.107)	0.0202 (0.105)	0.0894 (0.0839)	0.176* (0.0993)	0.230** (0.100)
ht_othserv	-0.0638 (0.124)	-0.228 (0.157)	-0.180 (0.148)	-0.00106 (0.117)	-0.000291 (0.140)	0.0383 (0.146)
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Entr. Char.	No	Yes	Yes	No	Yes	Yes
Firm Char.	No	No	Yes	No	No	Yes
Observations	2,632	1,954	1,901	2,653	1,968	1,911
R-squared	0.027	0.044	0.066	0.030	0.045	0.067

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 12: OLS regressions for practical education. Standard errors are corrected for heteroscedasticity.